

## Supporting student well-being with AI agents: A pilot study in Catalan secondary schools

### Suport al benestar dels estudiants amb agents d'intel·ligència artificial: un estudi pilot a les escoles secundàries catalanes

Joonas Merikko<sup>a</sup>

Jana Moral Banal<sup>b</sup>

<sup>a</sup>Departament de Ciències de la Computació, Universitat d'Hèlsinki i Annie Advisor Ltd., Hèlsinki (Finlàndia)

A/e:

[joonas.merikko@annieadvisor.com](mailto:joonas.merikko@annieadvisor.com)

<https://orcid.org/0000-0003-4166-4762>

<sup>b</sup>Annie Advisor Ltd., Hèlsinki (Finlàndia)

A/e:

[jana.moral@annieadvisor.com](mailto:jana.moral@annieadvisor.com)

Data de recepció de l'article:  
23 de setembre de 2025

Data d'acceptació de l'article:  
13 de novembre de 2025

Data de publicació de l'article:  
4 de maig de 2026

DOI: 10.2436/20.3007.01.239



Copyright © 2026

Joonas Merikko i Jana Moral Banal. Aquest article és d'accés lliure subjecte a la llicència Creative Commons Reconeixement – No Comercial–Sense Obra Derivada 4.0 Internacional. Per a més informació consulteu: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>

#### Abstract

Secondary education schools face persistent challenges of disengagement and early school leaving. AI agents may offer novel, low-threshold opportunities to tackle this problem by enhancing student support. The present pilot study in five Catalan secondary schools during the spring of 2025 aims to explore students' engagement with AI agents for student well-being, including the nature of students' needs and concerns in AI conversations, and student perceptions of the AI agents. The agents were co-designed by the authors and the schools, which defined their functions, boundaries, and referral pathways, and deployed them to students ( $N = 363$ ) via WhatsApp nudge messages. Research data was gathered from the support bot system (engagement, conversation themes), and a WhatsApp student survey (satisfaction, willingness to continue use). Engagement ranged from one in five to one in two students per school. The majority (58.0%) of interactions indicated "everything is OK". The most significant expressed needs concerned life plans and identity (8.0%), study planning (5.8%), and psychological well-being (4.1%). Slightly over half (52.5%) of the students wished to continue use, but satisfaction varied between schools. The findings suggest that the design of AI agents should adapt to students' varying receptiveness and local implementation contexts.

#### Keywords

AI agents, student well-being, secondary education, help-seeking, early school leaving.

## Resum

Les escoles d'educació secundària s'enfronten a reptes persistents de desvinculació i abandó escolar prematur. Els agents d'intel·ligència artificial (IA) poden oferir oportunitats noves i de baix llindar per abordar aquest problema i millorar el suport dels estudiants. El present estudi pilot a cinc instituts catalans durant la primavera de 2025 té com a objectiu explorar el compromís dels estudiants amb els agents d'IA per al seu benestar, incloent la naturalesa de les seves necessitats i preocupacions en les converses d'IA, i les percepcions dels estudiants dels agents d'IA. Els agents van ser codissenyats pels autors i les escoles, que van definir-ne les funcions, els límits i les vies de derivació, i els van desplegar als estudiants (N = 363) a través de missatges de WhatsApp. Les dades de recerca es van recollir per mitjà del sistema que gestiona els bots de suport (participació, temes de conversa) i una enquesta d'estudiants per WhatsApp (satisfacció, disposició a continuar fent servir els agents). La participació va variar d'un de cada cinc a un de cada dos estudiants per escola. La majoria (58,0 %) d'interaccions indicava «tot està bé». Les necessitats expressades més significatives es referien a plans de vida i identitat (8,0 %), planificació d'estudis (5,8 %) i benestar psicològic (4,1 %). Una mica més de la meitat (52,5 %) dels estudiants desitjaven continuar fent servir els agents, però la satisfacció variava entre escoles. Els resultats suggereixen que el disseny d'agents d'IA s'hauria d'adaptar a la diferent receptivitat dels estudiants i als contextos d'implementació locals.

## Paraules clau

Agents d'IA, benestar estudiantil, educació secundària, cerca d'ajuda, abandonament escolar prematur.

Com fer referència a aquest article / How to cite this article:

Merikko, J. & Moral Banal, J. (2026). Supporting student well-being with AI agents: A pilot study in Catalan secondary schools. *Revista Catalana de Pedagogia*, 29, 166-185. <https://doi.org/10.2436/20.3007.01.239>

## 1. Introduction

The landscape of education is facing escalating challenges that profoundly affect student outcomes: a surge in adolescents' mental health needs (Soriano et al., 2025), young people's declining faith in the value of education (Jones, 2024) and, most recently, the introduction of novel AI technologies calling for a fundamental re-evaluation of the objectives and methodologies of education (Kasneci et al., 2023; Markauskaite et al., 2022; Tuomi, 2023). These issues are compounded by a lack of belongingness within educational environments (Arslan, 2021), student disengagement (Af Ursin et al., 2023) and, ultimately, high rates of early school leaving (Berral-Ortiz et al., 2022; López et al., 2017).

Early school leaving often follows a gradual process of disengagement, where education systems fail to adequately address students' diverse needs (Nevala et al., 2011). Although schools often have various kinds of student support systems and initiatives, barriers like stigma, negative beliefs towards support services, difficulties in recognition of symptoms and a preference for self-reliance often hinder students' help-seeking (Aguirre Velasco et al., 2020; Gulliver et al., 2010). Due to these barriers, students seek help too late or not at all, which leads to escalating issues that negatively affect their study progress.

Recent advancements in artificial intelligence (AI) show promise in helping schools address challenges in student engagement and well-being. First, AI can lower barriers to help-seeking: research indicates that people are often more comfortable disclosing sensitive concerns to conversational agents than to adults, perceiving AI as less judgmental and more approachable (Branley-Bell et al., 2023; Bojd et al., 2024; Ta et al., 2020). Second, AI can complement human professionals by managing routine check-ins, providing reflection support, and triaging low-intensity needs, thus enabling school staff to focus on students requiring intensive or specialized care (Anonymous, 2021; Anonymous, 2024).

In Catalonia, as in many other regions, secondary education schools grapple with early school leaving and student disengagement. This paper reports exploratory findings from a pilot study conducted in five Catalan schools, where students engaged with AI agents designed to encourage reflection and offer guidance ranging from everyday matters to psychosocial well-being. In addition to offering low-threshold support, the system generated anonymized statistics from student-AI conversations, providing school staff with new insights into emerging patterns of well-being across their student populations.

### 1.1. *Artificial intelligence as a support for learning*

Visions of artificial intelligence in education range from dystopian warnings to utopian promises. Sharples (2022) cautioned that if students turn to AI to produce essays and teachers turn to AI to assess them, "nobody learns, nobody gains". In contrast, Andreessen (2023), a Silicon Valley venture capitalist, imagines every learner accompanied by an endlessly patient, compassionate, and knowledgeable AI tutor, "the machine version of infinite love". These contrasting viewpoints reflect a broader

tradition: for decades, new technologies have inspired both radical optimism and deep skepticism about the future of education (Cuban, 2001; Selwyn, 2010). As Selwyn (2019) argues, however, the integration of technology into education is never inevitable but reflects deliberate social and political choices.

The recent introduction and rapid uptake of large language models (LLMs) has made this debate very concrete. These models generate natural language and are expected to change learning in profound ways (Floridi & Chiriatti, 2020; Kasneci et al., 2023). Additionally, the field of learning analytics (LA) has advanced from dashboards showing learner progress (Verbert et al., 2013) to scaffolds that give learners prompts and nudges attempting to support and develop learners' self-regulation (Heikkinen et al., 2023; Lim et al., 2023; Lodge et al., 2018; van der Graaf et al., 2021). Still, evidence of the influence of these technologies on learning and teaching is mixed. For example, relying on AI can make learners less active in their own thinking, leading to "metacognitive laziness" (Fan et al., 2024). Furthermore, AI in education research has mostly focused on content learning, ignoring broader goals such as identity, social skills, and critical thinking (Tuomi, 2023).

From a teacher perspective, an interesting topic is how these technologies might shape the teaching profession in the future. According to Frey and Osborne (2017), teaching is among the professions least susceptible to automation. Instead of automation, the more plausible trajectory is augmentation, where AI extends rather than replaces teachers' work (An et al., 2020; Langer & Landers, 2021). Selwyn (2019) describes two very different possible future scenarios: AI might free teachers to focus on meaningful teaching, or it might limit their autonomy and reduce their work to technical functions. To support meaningful integration of AI into the teaching profession, building trust, respecting teachers' agency, and integrating professional judgment into system design are essential (Luckin et al., 2022; Nazaretsky et al., 2022). For example, Cukurova et al. (2023) found that teacher engagement with adaptive learning platforms was predicted less by product quality or technical competence than by workload, ownership, support structures, and trust. Furthermore, Viberg et al. (2025) showed that teacher attitudes toward AI and EdTech are shaped by institutional and cultural factors as much as personal beliefs. This suggests that the impact of AI will depend on how it is introduced and supported in practice.

### 1.2. *Artificial intelligence as a support for psychosocial well-being*

The idea of using computers for psychological support dates back to the 1960s, when Weizenbaum's ELIZA (1966) simulated the persona of a psychotherapist by reflecting users' input back to them, sparking debate about whether machines could play a role in emotional support. Today, such debates are urgent as companionship and therapeutic dialogue have emerged as some of the most prevalent use cases of generative AI (Zao-Sanders, 2025). Several studies have shown that conversations with generative AI may offer a low-stakes, non-judgmental space to overcome stigma about discussing sensitive topics (Bojd et al., 2024; Branley-Bell et al., 2023; Ta et al., 2020). On the other hand, research has identified risks such as overreliance on AI, factual inaccuracies ("hallucinations"), sycophancy (the tendency of the model to uncritically

agree with the user), and unclear accountability for the advice provided (Malfacini, 2025; Sharma et al., 2024; Zhai et al., 2024).

It is important to distinguish between general-purpose generative AI tools (such as ChatGPT, Gemini, and Copilot) and specialized AI agents. The latter are defined as artificial entities capable of perceiving their environment, making decisions, and taking actions, and are typically designed with specific contexts, use cases, and risks in mind (Xi et al., 2023). For instance, an AI agent designed to support student well-being should meet specific design requirements, such as the ability to escalate cases to human professionals and alignment with the ethical standards of the student well-being field.

Several design strategies have been explored to mitigate the risks of AI in the context of psychological well-being. Prompt engineering can structure and strengthen therapeutic dialogue (Souza et al., 2024). Fine-tuning methods can help models sustain longer supportive conversations while respecting professional boundaries (Madan & Subramanian, 2025). Value alignment efforts examine whether LLMs adhere to human values that are essential in mental health contexts (Hadar-Shoval et al., 2024). In addition, frameworks for risk evaluation and response monitoring aim to detect unsafe content and protect users (Marrapese et al., 2024).

A growing body of research has begun to explore whether conversational AI agents can deliver meaningful psychological support in practice. Maples et al. (2024) surveyed over one thousand students using the chatbot Replika and found that many students perceived high levels of social support from the agent, with notably 3% reporting that engagement with the chatbot had interrupted suicidal ideation. Hatch et al. (2025) showed that language models can produce therapeutic responses rated by participants as on par with those of professional therapists. Heinz et al. (2025) conducted the first randomized controlled trial of a generative AI therapy bot, finding medium-to-large effects on depression, anxiety, and eating disorder symptoms, alongside high ratings of therapeutic alliance. Furthermore, in the context of online acceptance and commitment therapy intervention for higher education students, Anonymous (2025) found that students' willingness to engage with AI guidance was predicted by both their satisfaction with the AI agent, and their help-seeking barriers, with students more comfortable with emotional expression being more likely to prefer guidance by AI agents.

While these early findings underscore the potential of AI in psychological support, they also highlight the importance of further research on safety, boundaries, and long-term impact, including possible adverse effects.

### 1.3. *The present study*

The current work builds on earlier work by Merikko and Silvola (2024), which demonstrated how AI agents can act as low-threshold first listeners for students while simultaneously generating anonymized well-being data for schools. Most existing studies regarding AI in school context focus on content learning, whereas studies on digital mental health rarely specialize in school contexts. In particular, there is little

evidence on how students in secondary education interact with universally offered well-being focused conversational AI, what kinds of needs they voice in these conversations, and how they perceive the experience. In response to this gap, the present study addresses the following research questions:

RQ1. To what extent do students in secondary schools engage with conversational AI agents designed for guidance, reflection and well-being?

RQ2. What kinds of needs and concerns do students bring forward in these interactions?

RQ3. How do students perceive the usefulness and acceptability of engaging with such agents?

Together, these questions explore how conversational AI agents may play a role in expanding access to support, contributing to early detection of needs, and complementing staff resources in ways that could promote student well-being and school retention.

## 2. Material and methods

### 2.1. Context

The pilot was conducted across five educational institutions in Catalonia, representing a diverse mix of contexts. These included two *centres de formació d'adults* (CFA, adult education centers) serving adults and young people returning to studies, one *institut-escola* (IE, combined primary and secondary public school), one *institut d'educació secundària* (INS, public secondary school) offering compulsory and post-compulsory education, and one semi-private secondary school (PC), operating within the publicly funded but privately managed sector. This diversity ensured that the pilots captured a broad picture of student needs and institutional settings. The acronyms CFA, IE, INS, and PC will be used in subsequent tables to indicate institution type. Informed consent was obtained from the participating students before starting the pilot, stating their willingness to participate in the pilot as well as accepting the reporting of the results.

### 2.2. Designing AI agents

The AI agents were developed through a collaborative process with the participating schools. Each center began by completing a structured preparation task in which staff defined the agent's role and limits. This involved specifying what the agent could and could not do, identifying the materials it should draw on (such as school policies, curriculum documents, or support resources), and deciding which human professionals (e.g., tutors, teachers, psychologists) the agent should refer students to when needed.

The needs of each school were identified, affecting the design of each AI agent. School A (CFA) considered the possibility of including the external organizations it collaborates with (municipal services and NGOs related to health and sexual identity, psychological

support, women's services, and youth services) among its support providers. However, they eventually decided to start the pilot without including them, to first observe how the tool worked. School B (CFA) did not want to cover topics related to students' emotional well-being, only academic matters, being a small school with limited resources. School C (IE), School D (INS) and School E (PC) wanted to cover all areas and specified that almost all requests should go directly to the tutor. This is because in compulsory secondary education the tutor plays a key role in supporting students and is usually the person they trust the most in the school. In these cases, the tutor assessed the need and referred to another professional if necessary.

Using this information, an initial version of the agent was built with the GPT-4o language model on the OpenAI platform separately for each school. Test versions were given to school staff for feedback. In addition to teacher testing, a simulation was run with 48 example student queries covering a wide range of typical issues. Staff then reviewed the responses to judge their relevance, accuracy, and compliance with ethical guidelines. Several iterations to the agents were carried out based on observations and feedback by school staff. For example, the reference materials schools wanted the AI agent to utilize initially included documentation about the school (courses, practical information, educational project, school management plan, regulations, etc.). Later, more resources were added (study techniques, information about higher education options, and materials to clarify subject content). The staff of School B (CFA) and School C (IE) were especially active in the iteration cycles, testing and giving ideas for improvements.

Once the schools approved their AI agent's performance, it was introduced to students through WhatsApp, the platform most commonly used by them. Each student ( $N = 363$ ) received a short nudge message, prepared by school staff and sent on behalf of the school, which explained the purpose of the agent and invited them to start a conversation. During the spring semester of 2025, each school sent three to four nudges, with wording differing slightly in each round. Examples are shown in table 1.

### 2.3. *Data collection and analysis*

All the data was collected from the support bot system in a pseudonymized form.

Regarding RQ1, we collected separately for each round 1) the number of students receiving the introductory message, 2) the number of students engaging in a conversation with the AI agent, and 3) the number of students creating a support request to a human professional with the assistance of the AI agent. In addition, we collected the daily number of incoming student messages separately for each school.

For RQ2, we collected thematic statistics separately for each round. The platform applies LLM-based thematic coding (see e.g., Simon, 2025) to the student-agent conversations, using the classification of student support needs developed by Anonymous (2024). While conversations could include several topics, each was assigned a single theme, determined by the automated system as the most dominant one.

For RQ3, we surveyed the students about their experience via WhatsApp. They were asked to report their satisfaction with the AI agent on a scale from 1 (“Not satisfactory”) to 4 (“Very satisfactory”), and whether they would like to continue using the tool in the future (“Yes” or “No”). Optionally, students could also give open feedback about the agent.

TABLE 1  
*Examples of nudge messages*

School	Example message
School A (CFA)	<p>Hello {firstname}, I’m {agentname}, your digital support from {schoolname}. How’s everything going this term? Do you need help or support with any of the following? Please reply with A, B, C, or D.</p> <p>A. Studies B. Work and career C. Emotional well-being D. Everything’s fine, or I’m already getting the support I need</p> <p>Please reply even if you don’t need support.</p>
School B (CFA)	<p>Hello {firstname}!</p> <p>I’m {agentname}, the support bot of your center. I’m here to help you resolve any questions you may have regarding academic guidance and to find information about the different studies and programs in which {schoolname} participates.</p> <p>I can also help you find some municipal services that might be useful to you, as well as assist you in choosing further studies after those you have completed at {schoolname}.</p> <p>Is there anything you would like to know?</p>
School C (IE)	<p>👋 Hello! We’re {schoolname}, and this is {agentname}, your support bot 🤖. With the Easter holidays coming to an end, this could be a good time to pause and reflect on how you’re feeling.</p> <p>Would you like to chat with the bot? It helps us understand you better so we can support you during the final part of the school year.</p> <p>It’s quick, confidential, and it might be really helpful!</p>
School D (INS)	<p>Hello {firstname}, I’m {agentname}, the support bot from {schoolname}. During this school year, I’ll be checking in with you to offer help whenever you might need it. I hope your school year has started well!</p> <p>Can I help you with something, or is everything okay? Please reply with 1, 2, or 3:</p> <ol style="list-style-type: none"> <li>1. I need help with my studies.</li> <li>2. I need help with my health or mood.</li> <li>3. Everything is fine, I don’t need any help.</li> </ol> <p>Please reply to this message even if you don’t need help. Best, {agentname}, your support bot</p>
School E (PC)	<p>Hello {firstname}!</p> <p>I’m {agentname}, the support bot of your center 🤖. Nice to meet you! I’m here to help</p>

you with whatever you need. From now on, I'll always be available so you can write to me whenever you want. Is there anything that's worrying you?

SOURCE: Prepared by the authors.

### **3. Results**

#### *3.1. Engagement and support requests*

Across the five participating schools, student engagement with the AI agents varied considerably (table 2). Campaign-level participation ranged from 3% to 56%, with most rounds drawing between 15% and 40% of students. When aggregated across campaigns, overall engagement per school ranged from roughly one in five to about one in two students. These figures indicate that a notable proportion of students responded to the invitation and interacted with the system at least briefly.

Support requests were relatively infrequent. On average, 3.6% of students initiated a request for human assistance, and several nudge rounds generated none. While the agents did identify a small number of students requiring professional follow-up, for the majority the interaction did not translate into formal requests for support.

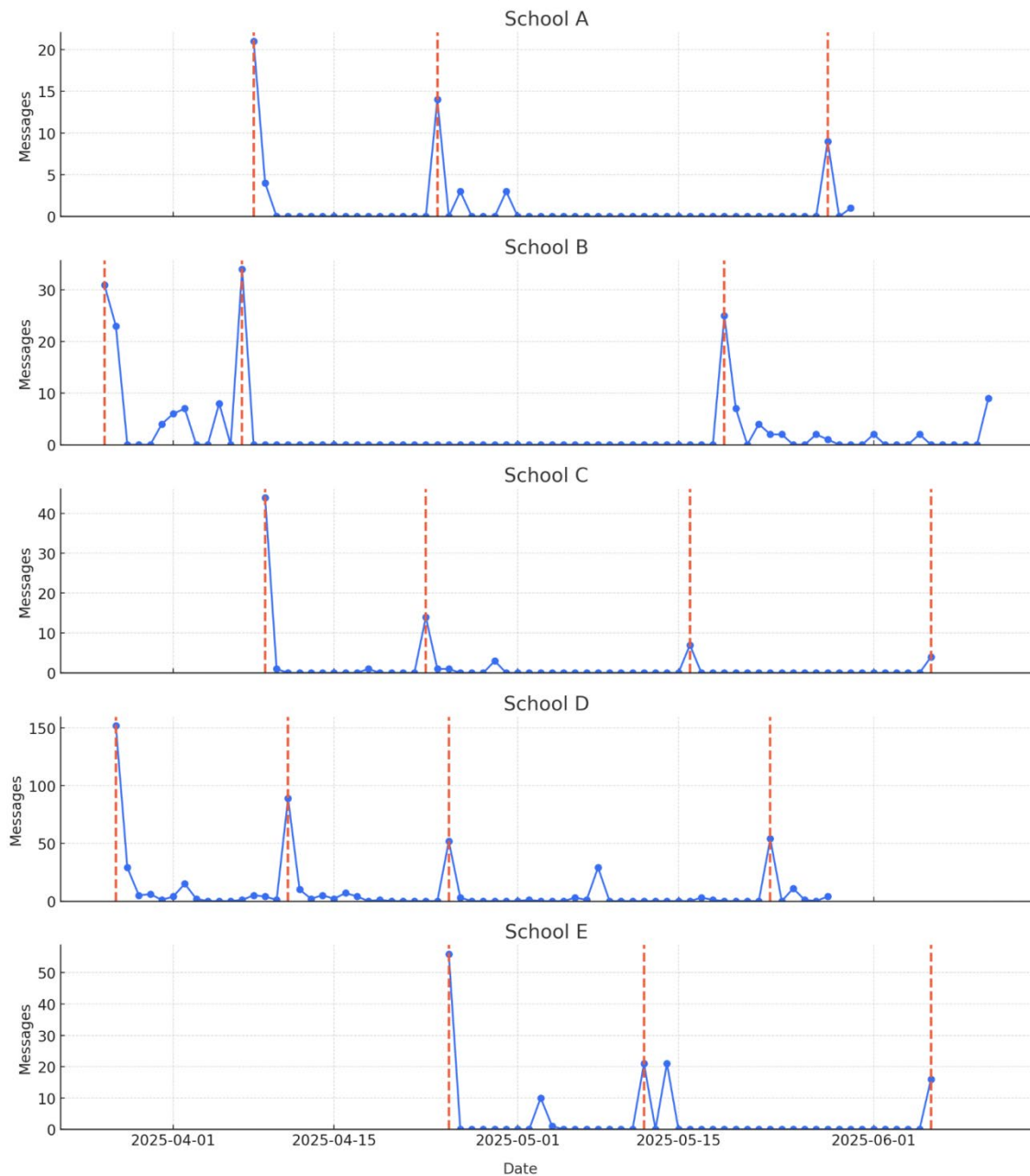
Analysis of daily activity patterns (figure 1) shows that engagement clustered strongly around the days when nudge messages were sent, with interaction dropping quickly afterward and remaining near zero until the next round. This indicates that structured outreach was the primary driver of use, while spontaneous engagement was rare.

TABLE 2  
*Engagement and support requests*

<b>Institution</b>	<b>Received the nudge <i>N</i></b>	<b>Engaged with the agent <i>N</i> (%)</b>	<b>Requested support <i>N</i> (%)</b>
School A (CFA)			
- First round (April 8th)	46	11 (23.9%)	2 (4.3%)
- Second round (April 24th)	46	7 (15.2%)	1 (2.2%)
- Third round (May 28th)	45	8 (17.8%)	0 (0.0%)
School B (CFA)			
- First round (March 26th)	81	25 (30.9%)	2 (2.5%)
- Second round (April 7th)	81	11 (13.6%)	2 (2.5%)
- Third round (May 19th)	75	20 (26.7%)	0 (0.0%)
School C (IE)			
- First round (April 9th)	70	27 (38.6%)	0 (0.0%)
- Second round (April 23rd)	60	11 (18.3%)	0 (0.0%)
- Third round (May 16th)	63	3 (4.8%)	0 (0.0%)
- Fourth round (June 6th)	62	2 (3.2%)	0 (0.0%)
School D (INS)			
- First round (March 27th)	109	61 (56.0%)	4 (3.7%)
- Second round (April 11th)	106	49 (46.2%)	0 (0.0%)
- Third round (April 25th)	102	31 (30.4%)	0 (0.0%)
- Fourth round (May 23rd)	97	36 (37.1%)	2 (2.1%)
School E (PC)			
- First round (April 25th)	57	28 (49.1%)	0 (0.0%)
- Second round (May 12th)	56	20 (35.7%)	0 (0.0%)
- Third round (June 6th)	55	13 (23.6%)	0 (0.0%)

SOURCE: Prepared by the authors.

FIGURE 1  
Daily student activity across the five participating schools during the spring 2025 pilot



Note: Blue dots represent the number of student messages on each day, while red dashed lines indicate the timing of nudge messages sent by schools.

SOURCE: Prepared by the authors.

### 3.2. Themes of conversation

Analysis of conversation themes shows that students most frequently used the agents to signal that “everything was OK” ( $n = 211$ ; 58.0% of conversations), which constituted the largest single category across all institutions. Among those who did express specific needs, the most common themes were related to life plans, career, and identity ( $n = 29$ ; 8.0%), followed by study planning ( $n = 21$ ; 5.8%), and teaching and learning arrangements ( $n = 18$ ; 4.9%). Issues

related to psychological well-being were raised in 15 cases (4.1%), while social well-being ( $n = 6$ ; 1.6%), subject matter ( $n = 6$ ; 1.6%), and physical health ( $n = 4$ ; 1.1%) were less frequent. Learning skills and special needs ( $n = 3$ ; 0.8%), documents and certificates ( $n = 2$ ; 0.1%), finance and housing ( $n = 1$ ; 0.03%), and tools and materials ( $n = 1$ ; 0.03%) were only occasionally mentioned.

The distribution of themes showed clear institutional differences (table 3). Schools B (CFA) and D (INS) were the most diverse in topics, ranging from life and study planning to psychosocial well-being. School D also had a high share of “other or unknown” responses, indicating less focused engagement. School A (CFA), School C (IE) and School E (PC) generated relatively few topical conversations, with students largely confirming they were fine.

TABLE 3  
*Conversation themes*

Theme	School A	School B	School C	School D	School E	Total
Life plans, career & identity	2	12	10	4	1	29
Study planning	1	15	1	2	2	21
Teaching and learning arrangements	0	14	0	1	3	18
Psychological well-being	3	1	1	9	1	15
Social well-being	0	0	0	6	0	6
Subject matter	2	0	0	2	2	6
Physical health	0	0	0	4	0	4
Learning skills & special needs	0	0	0	3	0	3
Documents & certificates	0	2	0	0	0	2
Finance & housing	0	1	0	0	0	1
Tools & materials	0	1	0	0	0	1
Other or unknown	0	7	0	35	5	47
Everything OK	14	3	31	109	54	211
<b>Total number of conversations</b>	<b>22</b>	<b>56</b>	<b>43</b>	<b>175</b>	<b>68</b>	<b>364</b>

SOURCE: Prepared by the authors.

### 3.3. *Student experience*

Student feedback on the AI agents was mixed (table 4). In terms of satisfaction, the largest group of respondents rated their experience as “very satisfactory” ( $n = 36$ ; 35.0%), followed by “good” ( $n = 28$ ; 27.2%), “could be improved” ( $n = 20$ ; 19.4%), and “not satisfactory” ( $n = 19$ ; 18.4%), indicating that while many students valued the experience, a notable minority remained dissatisfied. A slight majority (52.5%) was willing to continue use. Written feedback was sparse, with three students providing favorable comments (“Annie is the best!”, “Being able to talk about how you feel, with classmates and on a personal level, is really important for our emotions”, “Nothing to

improve”), one neutral (“For now, I’m satisfied”), and one critical note (“Giving reminders linked to the school from time to time would be more useful”).

Variation between institutions was evident: one adult education center (School B) generated the highest proportion of very satisfied students, while the public secondary education schools (Schools C and D) saw more critical responses. Overall, these findings suggest that while not universally embraced, the AI agents were perceived as useful by a substantial share of students.

TABLE 4  
*Student experience*

Institution	School A	School B	School C	School D	School E	Total
Satisfaction						
1 Not satisfactory	2	2	6	8	1	19
2 Could be improved	0	1	6	9	4	20
3 Good	0	6	6	8	8	28
4 Very satisfactory	2	24	5	3	2	36
Would like to continue use						
No	2	5	9	11	2	29
Yes	0	9	7	6	10	32
Written feedback						
Critical	0	0	0	0	1	1
Neutral	0	1	0	0	0	1
Positive	0	3	0	0	0	3

SOURCE: Prepared by the authors.

## 4. Conclusions and discussion

### 4.1. Overview of the findings

The pilot demonstrated that conversational AI agents successfully engaged a notable share of secondary students in the participating Catalan schools, though engagement varied across institutions and campaigns. In most cases, participation clustered strongly on the days when nudge messages were sent, with little spontaneous activity between rounds. This suggests that student interaction was closely tied to the structured outreach initiated by schools rather than self-directed use, in line with earlier findings in digital learning technologies, where structured prompts, nudges, or teacher-led cues are often required to sustain engagement (Lodge et al., 2018; Lim et al., 2023; Heikkinen et al., 2023).

Differences between schools were notable. In the institution where students reported the most positive experience, engagement rates were also among the highest, and

conversations covered the widest variety of themes. By contrast, other schools saw narrower patterns of use, with some dominated by short confirmations that “everything was OK”. This might suggest that outcomes were shaped by local context and implementation practices, in line with previous research describing how adoption of AI depends on workload, ownership, and trust rather than technical quality alone (Cukurova et al., 2023; Viberg et al., 2025).

Prior studies have suggested that AI can lower barriers to discussing sensitive issues (Ta et al., 2020; Bojd et al., 2024; Branley-Bell et al., 2023). In the present study, however, only a relatively small share of students escalated their concerns into formal support requests, especially compared with earlier rule-based systems (Anonymous, 2021). One possibility is that many students felt the short exchange with the agent was already sufficient and did not require further help. Another possibility is that the more personal style of LLM-based agents may unintentionally recreate barriers common in human help-seeking – compared to the structured, multiple-choice style of rule-based bots, which may feel more like filling in a form than admitting problems in conversation. Further research is needed to investigate the mechanisms underlying differences in help-seeking across interaction styles.

Finally, student feedback showed a mixed reception. Students with more positive experiences were also those in schools with higher engagement. This resonates with Anonymous (2025), who showed that willingness to engage with AI guidance depends on both satisfaction and openness to emotional expression. Not all students may be equally ready to use AI agents for support, highlighting the importance of tailoring implementation to different levels of receptivity.

#### 4.2 *Limitations and future work*

This study has several limitations. First, although the results showed differences between schools, our design does not allow to explain what caused these differences. They could be due to the characteristics of the student populations, differences in school culture, or the way each school introduced and supported the agents. Future studies should use designs that allow examining these factors separately to better understand the role of context.

Second, the thematic analysis used automated coding that assigned one main theme to each conversation. This method made it possible to compare results across schools, but it also simplified the data. By forcing each conversation into a single category, we could not capture the co-occurrence of different themes. This may hide important patterns, such as when academic and well-being concerns appear together.

Third, our analysis was nomothetic in nature, focusing on aggregated statistics across schools and nudge rounds. While this allowed us to identify overall patterns, it did not capture how individual students experienced or developed their interactions with the agents over time. Future research could benefit from an idiographic perspective, examining student experiences at the individual level. Such an approach would make it possible to follow how engagement unfolds across subsequent conversations and to

observe the combinations of academic and well-being concerns that may arise for each student.

#### 4.3. *Educational implications*

The findings suggest several educational implications. First, the role of school staff in designing and framing the use of AI is essential. Teachers and counselors were central in setting the boundaries of the agents in this pilot, and their involvement helped ensure that the systems aligned with local practices and professional responsibilities. At the same time, it is important that schools do not work in isolation. Sharing and learning from best practices across institutions could help refine implementation strategies, showing what kinds of nudges, framing, or follow-up practices lead to higher engagement and more meaningful use.

Second, the findings suggest that students differ in how natural they find interactions with conversational AI. While some engaged in extended dialogue, others responded only minimally. This underlines the importance of identifying student preferences and designing systems that respond to different levels of receptivity. AI support should offer layered options, ranging from highly structured to more open conversation, so that students can regulate the depth and pace of interaction. Providing agency in whether and how to engage is likely to enhance trust and acceptability.

Taken together, the findings indicate that conversational AI agents can extend existing support systems by lowering barriers to initial engagement and by highlighting cases that warrant human follow-up. Their role should not be seen as a substitute for professional services but rather as an additional channel that can complement them. In conclusion, conversational AI agents show potential in supporting student well-being. Realizing this potential, however, will require continued research and refinement of both technology and implementation practices.

## 5. References

- Af Ursin, P., Tikkanen, J., & Järvinen, T. (2023). Student disengagement in Finnish comprehensive schooling. In *Finland's Famous Education System: Unvarnished Insights into Finnish Schooling* (pp. 435-450). Springer Nature Singapore. [https://doi.org/10.1007/978-981-19-8241-5\\_27](https://doi.org/10.1007/978-981-19-8241-5_27)
- Aguirre Velasco, A., Cruz, I. S. S., Billings, J., Jimenez, M., & Rowe, S. (2020). What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. *BMC psychiatry*, 20(1), 293. <https://doi.org/10.1186/s12888-020-02659-0>
- An, P., Holstein, K., d'Anjou, B., Eggen, B., & Bakker, S. (2020). The TA framework: Designing real-time teaching augmentation for K-12 classrooms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-17). <https://doi.org/10.1145/3313831.3376277>

- Andreessen, M. (2023). The techno-optimist manifesto. Andreessen Horowitz. <https://a16z.com/the-techno-optimist-manifesto>
- Arslan, G. (2021). School belongingness, well-being, and mental health among adolescents: Exploring the role of loneliness. *Australian Journal of Psychology*, 73(1), 70-80. <https://doi.org/10.1080/00049530.2021.1904499>
- Berral-Ortiz, B., Ramos-Navas-Parejo, M., Lara-Lara, F., & Moreno-Palma, N. (2022, December). School dropouts in Spain: A systematic review. *Frontiers in Education*, 7, 1083774. Frontiers Media SA. <https://doi.org/10.3389/feduc.2022.1083774>
- Bojd, B., Garimella, A., & Yin, H. (2024). Overcoming the stigma barrier: Conversational information-seeking from AI chatbots vs. humans. In *Academy of Management Proceedings 2024*, 13882. Academy of Management Valhalla, NY. <https://doi.org/10.5465/AMPROC.2024.13882abstract>
- Branley-Bell, D., Brown, R., Coventry, L., & Sillence, E. (2023). Chatbots for embarrassing and stigmatizing conditions: Could chatbots encourage users to seek medical advice? *Frontiers in Communication*, 8, 1275127. <https://doi.org/10.3389/fcomm.2023.1275127>
- Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Harvard University Press. <https://doi.org/10.2307/j.ctvk12qnw>
- Cukurova, M., Miao, X., Brooker, R. (2023). Adoption of Artificial Intelligence in Schools: Unveiling Factors Influencing Teachers' Engagement. In: Wang, N., Rebolledo-Mendez, G., Matsuda, N., Santos, O.C., Dimitrova, V. (eds) *Artificial Intelligence in Education. AIED 2023. Lecture Notes in Computer Science*, vol 13916. Springer, Cham. [https://doi.org/10.1007/978-3-031-36272-9\\_13](https://doi.org/10.1007/978-3-031-36272-9_13)
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(1), 155-174. <https://doi.org/10.1111/bjet.13544>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681-694. <https://doi.org/10.1007/s11023-020-09548-1>
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry*, 10, 113. <https://doi.org/10.1186/1471-244X-10-113>
- Merikko, J. & Moral Banal, J. (2026). Supporting student well-being with AI agents: A pilot study in Catalan secondary schools. *Revista Catalana de Pedagogia*, 29, 166-185. <https://doi.org/10.2436/20.3007.01.239>

- Hadar-Shoval, D., Asraf, K., Mizrachi, Y., Haber, Y., & Elyoseph, Z. (2024). Assessing the alignment of large language models with human values for mental health integration: cross-sectional study using Schwartz's theory of basic values. *JMIR Mental Health, 11*, e55988. <https://doi.org/10.2196/55988>
- Hatch, S. G., Goodman, Z. T., Vowels, L., Hatch, H. D., Brown, A. L., Guttman, S., Le, Y., Bailey, B., Bailey, R. J., Esplin, C. R., Harris, S. M., Holt, P. Jr., McLaughlin, M., O'Connell, P., Rothman, K., Ritchie, L., Top, N. Jr., & Braithwaite, A. R. (2025). When ELIZA meets therapists: A Turing test for the heart and mind. *PLOS Mental Health, 2*(2), e0000145. <https://doi.org/10.1371/journal.pmen.0000145>
- Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A. D., Salzhauer, A. J., Griffin, T. Z., & Jacobson, N. C. (2025). Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI, 2*(4), AIOA2400802. <https://doi.org/10.1056/Aloa2400802>
- Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions—a systematic literature review. *Education and Information Technologies, 28*(3), 3059-3088. <https://doi.org/10.1007/s10639-022-11281-4>
- Jones, J. M. (2024, July 8). U.S. confidence in higher education now closely divided. *Gallup*. <https://news.gallup.com/poll/646880/confidence-higher-education-closely-divided.aspx>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences, 103*, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123*, 106878. <https://doi.org/10.1016/j.chb.2021.106878>
- Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., ... & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior, 139*, 107547. <https://doi.org/10.1016/j.chb.2022.107547>
- Lodge, J. M., Panadero, E., Broadbent, J., & de Barba, P. G. (2018). Supporting self-regulated learning with learning analytics. In *Learning analytics in the classroom* (pp. 45-55). Routledge. <https://doi.org/10.4324/9781351113038-4>
- López, N., Opertti, R., & Vargas Tamez, C. (2017). *Youth and Changing Realities: Rethinking Secondary Education in Latin America*. UNESCO Publishing. <https://doi.org/10.54675/OQYR2841>

- Luckin, R., Cukurova, M., Kent, C., & Du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, 3, 100076. <https://doi.org/10.1016/j.caeai.2022.100076>
- Madan, N., & Subramanian, L. (2025). In-Context Alignment at Scale: When More is Less. In *2nd Workshop on Models of Human Feedback for AI Alignment*. <https://icml.cc/virtual/2025/50225>
- Malfacini, K. (2025). The impacts of companion AI on human relationships: Risks, benefits, and design considerations. *AI & SOCIETY*, 40, 1-14. <https://doi.org/10.1007/s00146-025-02318-6>
- Maples, B., Cerit, M., Vishwanath, A., & Pea, R. (2024). Loneliness and suicide mitigation for students using GPT-3-enabled chatbots. *npj Mental Health Research*, 3(1), 4. <https://doi.org/10.1038/s44184-024-00004-4>
- Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martínez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham Shum, S., Gašević, D., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, 3, 100056. <https://doi.org/10.1016/j.caeai.2022.100056>
- Marrapese, A., Suleiman, B., Ullah, I., & Kim, J. (2024). A novel nuanced conversation evaluation framework for large language models in mental health. *arXiv preprint arXiv:2403.09705*. <https://arxiv.org/pdf/2403.09705>
- Merikko, J., Räsänen, P. (2026). Exploring the Role of AI Guidance in Internet-Based Acceptance and Commitment Therapy: Links to Human-Computer Trust and Help-Seeking Barriers. In: Degen, H., Ntoa, S. (eds) *HCI International 2025 – Late Breaking Papers. HCII 2025. Lecture Notes in Computer Science*, vol 16345. Springer, Cham. [https://doi.org/10.1007/978-3-032-13184-3\\_14](https://doi.org/10.1007/978-3-032-13184-3_14)
- Merikko, J., & Silvola, A. (2024). An AI agent facilitating student help-seeking: Producing data on student support needs. In *Joint Proceedings of LAK 2024 Workshops, co-located with 14th International Conference on Learning Analytics and Knowledge (LAK 2024), Kyoto, Japan, March 18-22, 2024*. Rheinisch-Westfaelische Technische Hochschule Aachen. <https://ceur-ws.org/Vol-3667/>
- Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, 53(4), 914-931. <https://doi.org/10.1111/bjet.13232>
- Nevala, A. M., Hawley, J., & Manoudi, A. (2011). *Reducing early school leaving in the EU. European Parliament Commissioned Report on Early School Leavers*. [https://www.europarl.europa.eu/RegData/etudes/etudes/join/2011/460048/I\\_POL-CULT\\_ET%282011%29460048%28SUM01%29\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/etudes/join/2011/460048/I_POL-CULT_ET%282011%29460048%28SUM01%29_EN.pdf)
- Merikko, J. & Moral Banal, J. (2026). Supporting student well-being with AI agents: A pilot study in Catalan secondary schools. *Revista Catalana de Pedagogia*, 29, 166-185. <https://doi.org/10.2436/20.3007.01.239>

- Pesonen, J.A. (2021). 'Are You OK?' Students' Trust in a Chatbot Providing Support Opportunities. In: Zaphiris, P., Ioannou, A. (eds) *Learning and Collaboration Technologies: Games and Virtual Environments for Learning. HCII 2021. Lecture Notes in Computer Science*, vol 12785. Springer, Cham.  
[https://doi.org/10.1007/978-3-030-77943-6\\_13](https://doi.org/10.1007/978-3-030-77943-6_13)
- Selwyn, N. (2010). Looking beyond learning: Notes towards the critical study of educational technology. *Journal of computer assisted learning*, 26(1), 65-73.  
<https://doi.org/10.1111/j.1365-2729.2009.00338.x>
- Selwyn, N. (2019). *Should robots replace teachers?: AI and the future of education*. John Wiley & Sons.
- Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askeel, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2024). Towards understanding sycophancy in language models. In *The Twelfth International Conference on Learning Representations ICLR 2024*.  
[https://proceedings.iclr.cc/paper\\_files/paper/2024/file/0105f7972202c1d4fb817da9f21a9663-Paper-Conference.pdf](https://proceedings.iclr.cc/paper_files/paper/2024/file/0105f7972202c1d4fb817da9f21a9663-Paper-Conference.pdf)
- Sharples, M. (2022). Automated essay writing: An AIED opinion. *International journal of artificial intelligence in education*, 32(4), 1119-1126.  
<https://doi.org/10.1007/s40593-022-00300-7>
- Simon, S., Sankaranarayanan, S., Tajik, E., Borchers, C., Shahrokhian, B., Balzan, F., Strauß, S., Viswanathan, S. A., Ataş, A. H., Čarapina, M., Liang, L., & Celik, B. (2025, July). Comparing a human's and a multi-agent system's thematic analysis: Assessing qualitative coding consistency. In *International Conference on Artificial Intelligence in Education* (pp. 60-73). Springer Nature Switzerland.  
[https://doi.org/10.1007/978-3-031-98420-4\\_5](https://doi.org/10.1007/978-3-031-98420-4_5)
- Soriano, V., Ramos, J. M., López-Ibor, M. I., Chiclana-Actis, C., Faraco, M., González-Cabrera, J., González-Fraile, E., Mestre-Bach, G., Pinargote, H., Corpas, M., Gallego, L., Corral, O., & Blasco-Fontecilla, H. (2025). Hospital admissions in adolescents with mental disorders in Spain over the last two decades: A mental health crisis? *European Child & Adolescent Psychiatry*, 34(3), 1125-1134.  
<https://doi.org/10.1007/s00787-024-02543-2>
- Souza, R., Lim, J. H., & Davis, A. (2024). Enhancing AI-Driven Psychological Consultation: Layered Prompts with Large Language Models. *arXiv preprint arXiv:2408.16276*. <https://arxiv.org/abs/2408.16276>
- Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., ... & Loggarakis, A. (2020). User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *Journal of medical Internet research*, 22(3), e16235. <https://doi.org/10.2196/16235>

- Tuomi, I. (2023). Beyond mastery: Toward a broader understanding of AI in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00343-4>
- Van Der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Bannert, M., ... & Molenaar, I. (2021). Do instrumentation tools capture self-regulated learning?. In LAK21: 11th international learning analytics and knowledge conference (pp. 438-448). <https://doi.org/10.1145/3448139.3448181>
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509. <https://doi.org/10.1177/0002764213479363>
- Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., ... & Kizilcec, R. F. (2025). What explains teachers' trust in AI in education across six countries?. *International Journal of Artificial Intelligence in Education*, 35(3), 1288-1316. <https://doi.org/10.1007/s40593-024-00433-x>
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>
- Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., ... Gui, T. (2025). The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, 68(2), 121101. <https://doi.org/10.1007/s11432-024-4222-0>
- Zao-Sanders, M. (2025, April). How people are really using Gen AI in 2025. *Harvard Business Review*. <https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025>
- 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), 28. <https://doi.org/10.1186/s40561-024-00316-7>