# How should we teach chatbot interaction to students? A pilot study on perceived affordances and chatbot interaction patterns in an authentic K-12 setting

**Conference Paper** · October 2024

DOI: 10.18420/delfi2024-ws-22



**1 author:**

Maria Klar
University of Duisburg-Essen
**8** PUBLICATIONS **5** CITATIONS

SEE PROFILE

# How should we teach chatbot interaction to students?

## A pilot study on perceived affordances and chatbot interaction patterns in an authentic K-12 setting

Maria Klar [1]

**Abstract:** Understanding how to use generative AI can greatly benefit the learning process. Despite available concepts for teaching "how to prompt", little empirical evidence exists on students' current micro-level chatbot use that would justify a need for instruction on how to prompt. This pilot study investigates students' chatbot use in an authentic setting. Findings reveal general interaction patterns, including a notable lack of conversational patterns, indicating an underutilization of this central chatbot capability. However, despite having no formal instruction, some students discovered specific chatbot affordances. While basic prompting skills are displayed or acquired during exploration, explicit training on effective chatbot interaction could enhance skillful chatbot use. This training should integrate cognitive and metacognitive strategies as well as technological knowledge, helping students leverage the technology's full potential.

**Keywords:** Learning, Chatbots, Interaction Patterns, Chatbot Affordances, Prompting Instruction

## 1 Introduction

Understanding how generative AI *works* and knowing how to *use* it are equally important [Br16; LM20; Sc23]. Students themselves express a desire to learn effective AI use for educational purposes. In a recent survey, nearly half of the respondents ranked learning how to use AI as their most important objective concerning AI education [Vo24]. Although existing concepts for teaching prompt strategies are available [Th23], there is little empirical evidence available on learners' current use of chatbots at the micro level. This challenge is compounded by the numerous scenarios in which chatbots based on large language models (LLMs) can be utilized. These chatbots offer many *indirect affordances* — functionalities perceived by users but not intentionally designed [ADS17]. For instance, although ChatGPT was not specifically designed to be a Socratic tutor, it can function as one [Al23].

This versatility likely drives students' need for guidance on when and how to use chatbots for learning. Given the emerging nature of empirical studies on learners' authentic chatbot usage, it is crucial to investigate how to provide effective instruction on chatbot interaction that truly meets their needs.

1 University of Duisburg-Essen, Chair of Media Didactics and Knowledge Management, Universitätsstr. 2, 45141 Essen, Germany, maria.klar@uni-due.de, https://orcid.org/0000-0002-5298-8458

## 2 Related Work

Due to the recent publication of LLM-based chatbots, emerging research tends to be rather general [WPL24]. There are few studies examining how students use ChatGPT in authentic, detailed settings. Existing surveys indicate that students primarily use these chatbots for searching for information and clarifying concepts, less often for feedback and exam preparation [vM23; Vo24].

In contrast to the larger proportion of students who wish for guidance on effective use for learning [Vo24], there are also studies showing that university students perceive chatbots as easy to use [Ng24; Ro23a]. As the chatbot interface generally does not require technical skill to use, this perceived ease of use is understandable. In contrast to search engines, chatbots do not require abstracting search terms or selecting from numerous results [KKM24]. However, chatbots provide more than just answers to questions in natural language. In educational contexts, for example, they can provide individual feedback and more personalized instruction [Ka23], and they allow for deeper processing of learning material [Ro23b]. When students, therefore, see chatbots as easy to use, they might underestimate chatbots' affordances and overestimate their abilities to fully use them [SD18]. Additionally, they might perceive chatbots as an "easy" medium overall, overlooking functionalities that require a certain level of mental effort [Sa84].

To leverage a broader range of chatbot affordances, learners need to be aware of these functionalities and capable of using them effectively. This requires self-regulation and metacognitive skills [Ab24]. Learners must recognize their information needs, take appropriate actions, and evaluate their learning activities. However, self-regulated learning and the use of metacognition that it requires often pose a challenge for learners [Az12; Bo05]. It is not guaranteed that learners will engage in effective learning activities, even when they are generally aware of learning strategies.

Therefore, while the basic functionalities of chatbots probably can be explored by students without much instruction, it is unclear whether students can engage effectively in the more complex interactions that chatbots afford. This study aims to examine the perceived affordances and actual interaction patterns of students using chatbots in an authentic setting.

## 3 Method

This pilot study involved eight students, aged 15 to 18 in a 10th-grade class at a comprehensive school in Germany. Seven of them were male and one was genderqueer. The class teacher described the group as having diverse academic skills. The study took place in project-oriented learning in an umbrella subject that covers history, politics, geography, and ethics. In that project, the students were tasked with self-directed research on a self-selected topic. The study was conducted at the end of the school year, in July 2023. The data collection took place at the beginning of the project so that the students had decided on their self-selected

topic or were in the process of deciding on it. In the lesson, the students received the task from the teacher to conduct exploratory research on their topic in order to collect information and find or refine a research question that would guide their research during the project. Thus, the whole project and the lesson that was observed in this study were an open learning setting that required significant self-regulation by the students. The topics ranged from Martin Luther King to McDonalds and modern colonialism in Africa. Previously, the students had received no instruction on using the chatbot for exploratory search. A chatbot like ChatGPT had been used in a previous German lesson on how to write an argumentative text.

Data were collected in two ways: 1) Screen recordings were taken during the lesson where students were encouraged to use a chatbot as a research tool. 2) Students were interviewed in pairs about their perceived affordances and other related questions in semi-structured interviews.

Before or during the exploratory search session, the students did not receive any instruction on using the chatbot, neither in general terms nor specifically for the task at hand. The students had access to a GDPR-compliant version of ChatGPT-3.5. The lesson lasted about 50 minutes and students were allowed to use web search, any other tools of their choice, and to take notes during their research phase. The semi-structured interviews, lasting 12 to 25 minutes, were conducted after students had at least 30 minutes of research time with the chatbot. They included questions such as: *"For what can you use AI chatbots like the one you just tried in learning scenarios?"*, *"What can you not use it for in these scenarios?"*, and *"How often do you use tools like ChatGPT for school purposes? What for? Why?"*

Both types of data, the screen recordings and the interview transcripts, were coded inductively, and a selection of the main categories and frequencies are reported in the following section.

## 4 Results

In this authentic setting, the eight students showed varying levels of engagement with the task overall, and thus varying levels of engagement with the chatbot. While one student was highly engaged (74 codes), switching back and forth between chatbot, web search and note-taking, one student barely engaged with the task and chatbot (6 codes).

Tab. 1 shows a selection of the main categories. Notably, only one follow-up question was posed and there are relatively few instances of clarification requests, both essential affordances of chatbots compared to search engines. Instances of web search are also relatively few, indicating little integration of web search with chatbot interaction.

While these frequencies of interactions provide an overview of general chatbot interactions, a case-by-case analysis shows how differently the students used the chatbots within this session. One student (S) posed only nine queries, but five of them were adaptation prompts (e.g., "*Tell me everything important about Martin Luther King*," "*Summarize in bullet*

| Category with example | Freq. overall | No. of students |
|---|---|---|
| Question (“What was the battle of the Teutoburg Forest?”) | 44 | 6 |
| Meta-Question (“What could I ask about the topic?”) | 15 | 4 |
| Off-Topic (“What is my name?”) | 14 | 4 |
| Adaptation Prompt (“Make it easier”, “Give me an overview?”) | 12 | 5 |
| Web search (“modern colonialism in Africa”) | 12 | 5 |
| single word/concept (“about Yakuza”) | 7 | 3 |
| Clarify a concept (“What are ’allies’?”) | 4 | 4 |
| Follow-up question (“How does such a dependence develop?”) | 1 | 1 |

Tab. 1: Selection of main categories with frequencies

*points*”). Another student (L) had the chatbot create a profile about Pablo Escobar and engaged in in-depth reading of this single chatbot answer. R was the only one to engage in some back-and-forth conversation with several related questions and a follow-up question. One student (F) primarily used meta-questions to find a good research question, asking for a *“simple”* then a *“short”* and then a *“good”* guiding question. They repeatedly used similar prompts, perhaps because they were dissatisfied with the results but they were unable to reformulate their query more specifically. Another student (B) was the only student who interleaved web search and chatbot prompts, also using adaptation prompts. A student (P) spent the lesson in search of a topic and asked very specific, unrelated questions, such as “*Where is the Nazi Gold?*” but also very general meta-questions such as “*What is a good topic for two people?*”.

So, in addition to the general patterns of additive, sometimes unrelated questions without back-and-forth conversation, the individual cases showed a variety of specific chatbot interactions as well as workflows between note-taking, chatbots, and search engines.

In the interviews, the students were asked about the chatbots’ affordances and limits in learning settings. Several students mentioned core affordances of generative AI chatbots: generating text, generating answers to questions, and engaging in a conversation. They were aware that the chatbot can create all kinds of texts: from poems (L, P), bullet points (G), and summaries (B), to scripts for video games (L). The chatbot could answer questions *“like a teacher”* (L) or answer those questions for which a Google search yields no answer (B, F). Apart from these general affordances, they named quite specific ways to use chatbots for the task at hand, i.e., exploratory search. S suggested several ways the chatbot can support here: *“So, you can use it as a topic support, so, to find new topics, or see the topic, so to speak, in all directions and thereby you have more possibilities to decide for a direction of the topic.”*

These relatively specific affordances indicate that the students might be able to discern the chatbot’s functionalities for a specific context, while more general affordances in contexts of learning or studying are not mentioned yet. For example, receiving support while solving

tasks, receiving feedback and error correction, summarizing, or simplifying texts, and receiving an alternative explanation are not mentioned or mentioned just once.

When asked what chatbots can *not* be used for in contexts of studying, there is less variety in the answers and several students referred to the same limits: the chatbot cannot give *"hard facts"* (R), especially not on current events, and it does not have an opinion or emotions. These limits are not directly related to learning or studying but are of a general nature. Only S made several points about how the AI cannot teach competencies like asking good questions, it cannot give feedback where knowledge about the student is necessary and it does not actively give process-based feedback.

When asked about their use of AI chatbots for school purposes, only three students stated they had already used it, which might, of course, be due to social desirability, as chatbots are often not permitted by teachers. Several students doubted the usefulness of chatbots for learning, saying they prefer the *"good old Google Search"* (P), or that working with AI almost *"feels like managing a second person"* (S).

## 5 Discussion

This pilot study took place in an authentic setting where students, at the beginning of a project in social sciences, chose topics of interest and conducted exploratory searches. Screen recordings of one lesson and semi-structured pair-wise interviews were used to investigate students' use of chatbots during these activities.

General interaction patterns emerged. Only few students used conversational elements; instead, their questions were additive, with minimal follow-up or clarification questions. The chatbot affordance of back-and-forth conversation was underutilized. Also, students did not treat chatbots like search engines; almost all queries were fully formed questions rather than search terms, indicating that an (ineffective) transfer from search engine strategies did not occur [KKM24]. Within these general patterns, however, there was substantial variety in chatbot interactions. For example, some students used adapting prompts such as *"Give an overview"* or *"Make it easier,"* indicating independent learning about these chatbot's capabilities without prior instruction.

That students are able to identify such task-specific affordances is mirrored in the interviews. More task-general affordances of chatbots in learning settings are not yet mentioned by the students. The fact that several students preferred "traditional methods" over chatbots might be due to a lack of skills. For instance, students expressed dislike for lengthy chatbot responses but did not ask for shorter ones. Also, misconceptions were present, such as the belief that chatbots *"search in databases."* This suggests that exposure to a range of tasks with chatbot support, accompanied by instruction and reflection, is necessary to learn about chatbot capabilities and limits.

Importantly, chatbot prompting skills should be instructed in combination with cognitive and metacognitive learning strategies [Ab24]. For example, asking the chatbot for easier or shorter answers requires learners to recognize when the content is too demanding. Moreover, knowledge of how to construct elaborate prompts alone is less effective if learners do not know what next step would make the most sense in their learning process. Such integrated instruction could combine prompting tips with monitoring skills [MZD17] and generative learning activities to deepen information processing [FM16]: Students could explain their understanding to the chatbot and request feedback, enhancing their learning through active engagement. However, students often revert to simple learning strategies despite being aware of more effective ones [CN07]. Instruction therefore needs to point out the benefits of these strategies, so that learners are willing to exert the mental effort needed to perform them [Fe19].

Integrated instruction should also include aspects of how generative AI works [Ca23; Eh23]. For instance, students in this pilot study worked on formulating good research questions for their exploratory search. While they knew the criteria for a good question, they did not specify these in their chatbot queries. Knowing that the AI's training data might contain varying notions of a "good question" could have helped them enhance their queries with more context and detail. Furthermore, knowing that generative AI, at least in the chatbot used in this study, did not "search in databases" and that hallucinations might occur especially for content with limited information in the training data, they might have used web search, rather than the chatbot, for very specific questions. Pointing out the differences between traditional web search and LLM-generated content is crucial [KKM24] and instruction should include strategies to efficiently fact-check chatbot responses or focus on chatbot queries that are less prone to hallucinations like summaries or feedback.

Taking the context of this pilot study as an example, effective instruction would 1) explain technological differences between web search and LLM-chatbots, 2) suggest metacognitive and self-regulation strategies for exploratory search with chatbots, and point out their benefits, such as using the chatbot as a sparring partner for ideation, clarifying concepts in simple language, or asking for feedback, 3) combine this with general prompting tips, e.g., on adaptation prompts and providing enough context, 4) point out chatbot limits like hallucinations and respective strategies, 5) iteratively engage in reflection during the learning process, e.g., by collecting and discussing samples of chatbot interactions the students felt were helpful or unhelpful.

In summary, this pilot study showed that students displayed both general and diverse interaction patterns and an awareness of chatbot affordances without prior instruction. Allowing students to explore task-specific uses and combining prompting guidance with instruction on cognitive and metacognitive strategies as well as technological knowledge can help them fully leverage the chatbot's potential.

## References


[Ab24] Abdelhalim, S. M.: Using ChatGPT to Promote Research Competency: English as a Foreign Language Undergraduates' Perceptions and Practices across Varied Metacognitive Awareness Levels. Journal of Computer Assisted Learning 40 (3), pp. 1261–1275, 2024, ISSN: 1365-2729, DOI: 10.1111/jcal.12948, visited on: 02/05/2024.

[ADS17] Antonenko, P. D.; Dawson, K.; Sahay, S.: A Framework for Aligning Needs, Abilities and Affordances to Inform Design and Practice of Educational Technologies. British Journal of Educational Technology 48 (4), pp. 916–927, 2017, ISSN: 1467-8535, DOI: 10.1111/bjet.12466, visited on: 06/01/2023.

[Al23] Al-Hossami, E.; Bunescu, R.; Teehan, R.; Powell, L.; Mahajan, K.; Dorodchi, M.: Socratic Questioning of Novice Debuggers: A Benchmark Dataset and Preliminary Evaluations. In (Kochmar, E.; Burstein, J.; Horbach, A.; Laarmann-Quante, R.; Madnani, N.; Tack, A.; Yaneva, V.; Yuan, Z.; Zesch, T., eds.): Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023). Association for Computational Linguistics, Toronto, Canada, pp. 709–726, 2023, DOI: 10.18653/v1/2023.bea-1.57, visited on: 03/05/2024.

[Az12] Azevedo, R.; Feyzi Behnagh, R.; Duffy, M.; Harley, J.; Trevors, G.: Metacognition and Self-Regulated Learning in Student-Centered Leaning Environments. In (Land, S.; Jonassen, D., eds.): Theoretical Foundations of Student-Center Learning Environments. 2nd ed., Routledge, New York, pp. 171–197, 2012, ISBN: 978-0-203-81379-9.

[Bo05] Bol, L.; Hacker, D. J.; O'Shea, P.; Allen, D.: The Influence of Overt Practice, Achievement Level, and Explanatory Style on Calibration Accuracy and Performance. The Journal of Experimental Education 73 (4), pp. 269–290, 2005, ISSN: 0022-0973, DOI: 10.3200/JEXE.73.4.269-290, visited on: 07/30/2022.

[Br16] Brinda, T.; Brüggen, N.; Diethelm, I.; Knaus, T.; Kommer, S.; Missomelius, P.; Leschke, R.; Tilemann, F.; Weich, A.: Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt, 2016.

[Ca23] Casal-Otero, L.; Catala, A.; Fernández-Morante, C.; Taboada, M.; Cebreiro, B.; Barro, S.: AI Literacy in K-12: A Systematic Literature Review. IJ STEM Ed 10 (1), 2023, ISSN: 2196-7822, DOI: 10.1186/s40594-023-00418-7, visited on: 05/27/2024.

[CN07] Cao, L.; Nietfeld, J. L.: College Students' Metacognitive Awareness of Difficulties in Learning the Class Content Does Not Automatically Lead to Adjustment of Study Strategies. Australian Journal of Educational & Developmental Psychology 7, pp. 31–46, 2007, ISSN: 1446-5442, visited on: 06/07/2023.

[Eh23] Ehlers, U.-D.; Lindner, M.; Sommer, S.; Rauch, E.: AICOMP - Future Skills in a World Increasingly Shaped By AI. Ubiquity Proceedings 3 (1), pp. 230–239, 2023, DOI: 10.5334/uproc.91.

[Fe19] Feldon, D. F.; Callan, G.; Juth, S.; Jeong, S.: Cognitive Load as Motivational Cost. Educ Psychol Rev 31 (2), pp. 319–337, 2019, ISSN: 1573-336X, DOI: 10.1007/s10648-019-09464-6, visited on: 09/08/2023.

[FM16] Fiorella, L.; Mayer, R. E.: Eight Ways to Promote Generative Learning. Educ Psychol Rev 28 (4), pp. 717–741, 2016, ISSN: 1573-336X, DOI: 10.1007/s10648-015-9348-9, visited on: 04/12/2023.

[Ka23] Kasneci, E.; Sessler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günnemann, S.; Hüllermeier, E.; Krusche, S.; Kutyniok, G.; Michaeli, T.; Nerdel, C.; Pfeffer, J.; Poquet, O.; Sailer, M.; Schmidt, A.; Seidel, T.; Stadler, M.; Weller, J.; Kuhn, J.; Kasneci, G.: ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. Learning and Individual Differences 103, 2023, ISSN: 1041-6080, DOI: 10.1016/j.lindif.2023.102274, visited on: 02/27/2024.

[KKM24] Kerres, M.; Klar, M.; Mulders, M.: Informationskompetenz Neu Denken: Von Google Zu ChatGPT. Erwachsenenbildung 70, pp. 52–57, 2024.

[LM20] Long, D.; Magerko, B.: What Is AI Literacy? Competencies and Design Considerations. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20, Association for Computing Machinery, New York, NY, USA, 2020, ISBN: 978-1-4503-6708-0, DOI: 10.1145/3313831.3376727, visited on: 04/27/2023.

[MZD17] Moser, S.; Zumbach, J.; Deibl, I.: The Effect of Metacognitive Training and Prompting on Learning Success in Simulation-Based Physics Learning. Sci. Ed. 101 (6), pp. 944–967, 2017, ISSN: 00368326, DOI: 10.1002/sce.21295, visited on: 04/29/2022.

[Ng24] Ngo, T. T. A.; Tran, T. T.; An, G. K.; Nguyen, P. T.: ChatGPT for Educational Purposes: Investigating the Impact of Knowledge Management Factors on Student Satisfaction and Continuous Usage. IEEE Trans. Learning Technol. 17, pp. 1367–1378, 2024, ISSN: 1939-1382, 2372-0050, DOI: 10.1109/TLT.2024.3383773, visited on: 04/22/2024.

[Ro23a] Romero-Rodríguez, J.-M.; Ramírez-Montoya, M.-S.; Buenestado-Fernández, M.; Lara-Lara, F.: Use of ChatGPT at University as a Tool for Complex Thinking: Students' Perceived Usefulness. N.Appr.Ed.R 12 (2), pp. 323–339, 2023, ISSN: 2254-7339, DOI: 10.7821/naer.2023.7.1458, visited on: 03/18/2024.

[Ro23b] Rospigliosi, P. 'a.: Artificial Intelligence in Teaching and Learning: What Questions Should We Ask of ChatGPT? Interactive Learning Environments 31 (1), pp. 1–3, 2023, ISSN: 1049-4820, DOI: 10.1080/10494820.2023.2180191, visited on: 02/23/2023.

[Sa84] Salomon, G.: Television Is "Easy" and Print Is "Tough": The Differential Investment of Mental Effort in Learning as a Function of Perceptions and Attributions. Journal of Educational Psychology 76 (4), pp. 647–658, 1984, ISSN: 1939-2176, DOI: 10.1037/0022-0663.76.4.647.

[Sc23] Schüller, K.; Rampelt, F.; Koch, H.; Schleiss, J.: Better Ready than Just Aware: Data and AI Literacy as an Enabler for Informed Decision Making in the Data Age. In: INFORMATIK 2023 - Designing Futures: Zukünfte Gestalten. Gesellschaft für Informatik e.V., Bonn, pp. 425–430, 2023, ISBN: 978-3-88579-731-9, DOI: 10.18420/inf2023_49.

[SD18] Sanchez, C.; Dunning, D.: Overconfidence among Beginners: Is a Little Learning a Dangerous Thing? Journal of Personality and Social Psychology 114 (1), pp. 10–28, 2018, ISSN: 1939-1315, DOI: 10.1037/pspa0000102.

[Th23] Theophilou, E.; Koyutürk, C.; Yavari, M.; Bursic, S.; Donabauer, G.; Telari, A.; Testa, A.; Boiano, R.; Hernandez-Leo, D.; Ruskov, M.; Taibi, D.; Gabbiadini, A.; Ognibene, D.: Learning to Prompt in the Classroom to Understand AI Limits: A Pilot Study. In (Basili, R.; Lembo, D.; Limongelli, C.; Orlandini, A., eds.): AIxIA 2023 – Advances in Artificial Intelligence. Springer Nature Switzerland, Cham, pp. 481–496, 2023, ISBN: 978-3-031-47546-7, DOI: 10.1007/978-3-031-47546-7_33.

[vM23] von Garrel, J.; Mayer, J.: Artificial Intelligence in Studies—Use of ChatGPT and AI-based Tools among Students in Germany. Humanit Soc Sci Commun 10 (1), pp. 1–9, 2023, ISSN: 2662-9992, DOI: 10.1057/s41599-023-02304-7, visited on: 11/10/2023.

[Vo24] Vodafone-Stiftung: Pioniere Des Wandels: Wie Schüler:Innen KI Im Unterricht Nutzen Möchten, 2024, visited on: 05/10/2024.

[WPL24] Wong, L.-H.; Park, H.; Looi, C.-K.: From Hype to Insight: Exploring ChatGPT's Early Footprint in Education via Altmetrics and Bibliometrics. Journal of Computer Assisted Learning, pp. 1–19, 2024, ISSN: 1365-2729, DOI: 10.1111/jcal.12962, visited on: 02/26/2024.