

# Using Chatbots for Feedback Exchanges in Educational Settings

---

Jennifer Schluer

**Abstract** *Rapid developments in the field of Artificial Intelligence (AI), particularly since the public release of ChatGPT-3 in November 2022 (and of more recent versions), offer several opportunities for enhancing learning and teaching processes. The present paper concentrates on one crucial factor contributing to learning success, which is feedback. After suggesting a definition of chatbot feedback, the chapter outlines various advantages and discusses prevailing limitations. Moreover, the chapter provides recommendations for utilizing chatbots for feedback purposes in educational settings. Altogether, it is argued that chatbots can serve as feedback assistants for teachers and learners and can play a supportive role in the development of learners' and teachers' feedback literacy.*

**Keywords** *chatbots; artificial intelligence; chatbot feedback; technology-generated feedback; feedback training*

## 1. Introduction

Recent years have seen a rapid increase in the use of chatbots for various commercial and non-commercial purposes (cf. Essel et al., 2022, p. 1). Especially since the public release of ChatGPT in November 2022, AI-powered technologies have caught the interest of the general public and been vividly discussed since then. They have also evoked an intense debate among educators and researchers regarding the opportunities and dangers coming along with the existence and utilization of these technologies (Qureshi, 2023, p. 1).

Previously, AI had been mainly used for language learning software, translations and text corrections within and beyond the educational realm. The generated feedback, however, was mostly limited to mechanics (e.g. spelling and punctuation) and word choice, while still working unreliably for coherence, style and paragraph structure. This notably pertained to Automated Writing Evaluation (AWE) programs, as discussed by Schluer (2022). Further common drawbacks were the brief and formulaic nature of the comments (Chen & Cheng, 2008, p. 101) as well as the preponderance of negative feedback (Dembsey, 2017, p. 89). It often also lacked a human touch, even though some programs incor-

porated a virtual avatar or provided feedback in conversational language (as reviewed by Schluer, 2022, p. 85). Nevertheless, the unidirectionality of the feedback (from the program to the user) might not only be demotivating for the users but also contradicts the current conceptualization of feedback as part of an ongoing learning dialogue (Schluer, 2022, pp. 33–34). These limitations notwithstanding, the AWE technologies offered several benefits for advancing writing skills and building students' linguistic confidence (Schluer, 2022, pp. 81–83).

With the advent of more mature AI technologies and ongoing developments, however, additional avenues for facilitating educational processes arise for learners and teachers alike. The aim of the current contribution therefore is to review and discuss the opportunities and challenges of utilizing chatbots for feedback processes in educational settings. It will begin with a definition of chatbot feedback and a description of the most recent developments at the time of writing this chapter. Afterwards, the affordances and limitations will be synthesized along with recommendations for practical use. Finally, the chapter will close with a discussion of future needs to further improve AI-powered feedback processes.

## 2. Definition and Examples

Chatbot feedback refers to the use of artificial intelligence (AI) in chatbot applications to interact with a human user and provide written or oral feedback to submitted data and prompts. These chatbots have been trained on a vast amount of data (Ray, 2023, p. 133) and utilize Natural Language Processing (NLP) as well as sentiment analysis (Khanna et al., 2015, cited by Adamopoulou & Moussiades, 2020, p. 1) to analyze the user's input and provide personalized feedback on assignments, quizzes, or other learning tasks. The chatbot feedback system aims to simulate human-like interaction while offering students timely and constructive feedback, answering their questions, and assisting them in their learning process (Adamopoulou & Moussiades, 2020, p. 1; Chang et al., 2022, p. 15; Yin et al., 2021, p. 155). Furthermore, chatbots can produce samples of feedback requests, feedback interactions and action plans, which may serve as a model or inspiration for learners to become more active participants in the feedback process. Likewise, teachers can prompt chatbots to generate more elaborate and detailed feedback than they could offer themselves for reasons of time (Dai et al., 2023a, p. 323; Essel et al., 2022, p. 4).

All in all, the pedagogical roles of chatbots can be of various but often interconnected and increasingly overlapping kinds. Wollny et al. (2021, p. 12), for instance, distinguished between the functions of learning, assisting, and mentoring. The learning role refers to “the support in learning or teaching activities”, such as knowledge acquisition (Wollny et al., 2021, p. 12). The assisting function is understood as the provision of services for managing everyday life (Wollny et al., 2021, p. 12). The mentoring role, in turn, more specifically denotes the support of learners' personal development, e.g. by fostering their self-regulated learning (Wollny et al., 2021, p. 12). Others suggested a more general differentiation between chatbots as learning assistants and chatbots as teaching assistants. The latter are alternatively labelled “classroom assistants” and offer functions such as generating knowledge and providing intelligent feedback (as summarized by Essel et al., 2022,

p. 2). For both learners and teachers, the affordances of chatbots' feedback functionalities will be reviewed in this chapter.

While many AI-powered chatbots exist, ChatGPT has initially attracted most attention. Beforehand, chatbots had mainly been used by companies, e.g. for "automated customer service support" (Qureshi, 2023, p. 2), "website support" (Fryer et al., 2020, p. 9), or feedback surveys about client or employee satisfaction (cf. Lechler et al., 2019, p. 10). Accordingly, the majority of research papers concentrated on these contexts of use (see e.g. Lechler et al., 2019). In November 2022, though, OpenAI made ChatGPT (version 3) accessible to the general public (at <https://chat.openai.com/>), which has evoked vivid discussions in the academic sphere since then. GPT stands for "generative, pre-trained transformer" and is an advanced AI-powered chatbot which "uses a combination of deep learning and statistical methods to generate high-quality text" (Frackiewicz, 2023, n.p.). The successor ChatGPT-4 was trained on an even more extensive database of 570GB (Qureshi, 2023, p. 2) and was launched in May 2023 (Frackiewicz, 2023). With this large database and computational mechanisms, it can mimic human conversations more realistically (Kohnke et al., 2023, 2) and provide extensive output when prompted appropriately (Qureshi, 2023, p. 6; Schmied, 2023, p. 8). This "tool for natural language generation (NLG)" (Frackiewicz, 2023, n.p.) can produce more natural-sounding, detailed and context-specific responses for a vast array of topics than other tools before (Dai et al., 2023a, p. 323). By contrast, AI-powered digital assistants, such as Alexa or Siri, were mainly able to perform simple tasks upon request, as for instance switching on the light or the radio or reporting the weather forecast for a particular region (Kohnke et al., 2023, 2). As opposed to this unidirectionality, ChatGPT permits more elaborate "back-and-forth conversations" with its users (Kohnke et al., 2023, p. 2), which makes it more compliant with current feedback conceptualizations.

According to contemporary perspectives, feedback is regarded as a dialogic and interactive process aimed at enhancing learning processes and outcomes (Schluer, 2022, pp. 17, 33). Therein, feedback ideally operates in a cyclical or spiral-like manner towards continuous improvement (see Carless, 2019, and the illustration by Schluer, 2022, p. 44). In this process, learners and teachers adopt complementary responsibilities. More precisely, learner feedback literacy pertains to students' understanding, attitudes, and abilities to seek, provide, comprehend, and utilize feedback, as well as manage emotions (Carless & Boud, 2018, pp. 1316–1317). Conversely, teacher feedback literacy encompasses educators' capabilities to design learning environments that are conducive to students' feedback literacy development. This involves planning and coordinating feedback activities to facilitate proactive and multidirectional feedback exchanges, as well as fostering reflective and effective feedback utilization (Boud & Dawson, 2023, p. 158; cf. the review by Xie et al., 2022, pp. 28–29). The supportive role of chatbots in these processes will be considered in this chapter.

In the following sections, the discussion of advantages, limitations and possible implementations will not exclusively, but primarily center on ChatGPT to exemplify recent developments at the time of writing the chapter.

### 3. Affordances

ChatGPT (and similar chatbot technologies) can be used beneficially for manifold educational purposes. Several advantages thus arise, both for teachers and for learners, which will be outlined in the following sections.

#### 3.1. Chatbots as Feedback Assistants for Teachers

Especially in classes with numerous students, the provision of individual support becomes challenging for educators. In that regard, chatbot adoption can enable personalized assistance for each student by catering to their unique knowledge levels and learning paces (Essel et al., 2022; Kuhail et al., 2023, p. 974; Mageira et al., 2022, pp. 2–3; Zhang et al., 2023, p. 14). This can result in a substantial reduction of teachers' workload (Escalante et al., 2023, p. 13) as well as better student support. For instance, (in- and pre-service) teachers might use AI to generate ideas and tasks for teaching, to create lesson plans and syllabi (Schmied & Ivanova, 2023, p. v), as well as personalized assessments, including open-ended or multiple-choice questions, quizzes, and exams that are tailored to the student's proficiency level (cf. Amin, 2023, p. 3; Bommasani et al., 2021, cited by Cotton et al., 2023, p. 2; Wahyudi, 2023, p. 1). Moreover, and most relevant for this chapter, is the possibility to obtain AI feedback on learner texts (or other assignments). For instance, instructors could enter a(n anonymized) student text (e.g. essay or exam) into ChatGPT, together with a specification of the assessment criteria according to which the submission should be analyzed. In their study about ChatGPT's capacity to provide feedback to students on a data science project at an Australian university, Dai et al. (2023a) fed the following prompt into the chatbot:

"Please give feedback on the following text in terms of a clear description of the goals of the project, appropriateness of the topic to data science, clear description of the business benefits, novelty/creativity and overall clarity of the report. <INSERT THE TEXT OF A REPORT>." (p. 324)

They pasted each student's report into ChatGPT and obtained generated feedback (Dai et al., 2023a, p. 324).

This process can be facilitated even further through continuous improvements of the bot. For instance, on August 03, 2023<sup>1</sup>, OpenAI announced on their Twitter account that it will soon be possible to upload multiple files for analysis. However, these and other functions are only available for Plus users who subscribe to the premium version and thus have access to the more powerful ChatGPT-4. Moreover, they stated that it also allows to "generate insights across multiple files", which brings along several benefits for feedback processes as well. Assessors could thus compare students' performance within the classroom or over time for each learner. Coupled with learning analytics dashboards and digital feedback portfolios, this could provide a fruitful avenue for future feedback practices. Furthermore, on the classroom level, teachers can analyze conversation logs

1 <https://twitter.com/OpenAI/status/1687159114047291392>

and statistics generated during student interactions, helping them identify areas of difficulty and weaknesses in students' understanding (Mageira et al., 2022, pp. 2–3; cf. Essel et al., 2022).

In their study, Dai et al. (2023b) also compared the readability and nature of the feedback generated by ChatGPT with the feedback provided by a human instructor. They found that the bot-generated feedback was more detailed, more consistent and more process-oriented, more prose-like and more positive as well as “significantly more readable than instructor feedback ( $p < 0.001$ , examined by paired t-test)” (Dai et al., 2023b, p. 4). While the readability scores ranged between 3.75 to 4.0 for the ChatGPT output, more than three quarters of the instructor feedback was below the score of 3.75 (Dai et al., 2023a, pp. 324–325). Nevertheless, the generated feedback showed high agreement with the human rating along the five dimensions of the marking rubric (Dai et al., 2023b, p. 2). In terms of content and structure, the chatbot feedback commonly started with a brief summary of the student work, followed by an evaluation as well as “an explanation of how this feedback was given” (Dai et al., 2023b, p. 4) and suggestions for learning strategies that could enhance students' performance further (Dai et al., 2023b, p. 2). By contrast, the feedback by the human instructor “consisted mostly of simple expressions about the evaluation of student performance” (Dai et al., 2023b, p. 4).

In fact, the brevity of written feedback produced by human assessors is a phenomenon that has been widely observed in the literature and in teaching practice. The main reason for this is the high amount of time that humans need in order to produce elaborate written feedback. Several scholars have therefore investigated alternative ways to accelerate the process of feedback provision and to offer more comprehensible and detailed explanations at the same time. One of these suggestions is to use the oral modality or audiovisual feedback, e.g. in the form of screencast feedback (Schluer, 2020). Another frequently cited option is the combination with automated feedback, e.g. by using AWE software (Kim, 2018, cited by Schluer, 2022, pp. 207, 233). Beforehand, however, the automated comments mainly seemed to be useful for simple corrections of spelling and grammar, while showing several limitations with regard to feedback on the content (cf. the review by Schluer, 2022, pp. 78–91). Large language models, as they are used in more recent technologies (e.g. in ChatGPT), can help to overcome this limitation (e.g. Escalante et al., 2023, p. 12).

Especially when teachers have to deal with large numbers of students in their classes, AI-enhanced feedback can be a useful supplement, while they still need to be aware of its limitations (see section 4 below). Even before the advent of ChatGPT-3, other chatbots were used to help teachers manage a high amount of student queries. For example, in the research reported by Essel et al. (2022), a chatbot was incorporated into a messenger app to provide automatic responses to students in Ghanaian higher education. Outsourcing replies to simple student queries to AI-powered bots might lead to a reduction of teachers' workload (Essel et al., 2022, p. 2). Furthermore, it has been argued that summative assessments produced by ChatGPT can free up teachers' time so that they can concentrate more on high-level tasks, such as formative feedback provision and student support (Cotton et al., 2023, p. 3; cf. Qureshi, 2023, p. 1). Yet, with continuous advancements, chatbots can help teachers to assess student work formatively as well (cf. Kohnke et al., 2023, p. 2).

Teachers could also ask ChatGPT to produce assessment rubrics and to apply them to the submitted texts (see Gönen, 2023, cited in chapter 1 in this volume). Such an AI-supported application of assessment rubrics to submitted texts might create greater consistency and objectivity in feedback provision and the grading of student work (see the study by Dai et al., 2023b, p. 5; cf. Cotton et al., 2023, pp. 2–3; Escalante et al., 2023, p. 12; Qureshi, 2023, p. 1). Going beyond corrections, ChatGPT has the capacity to generate suggestions for further learning as well as to produce “personalized learning materials” (Qureshi, 2023, p. 3). Furthermore, the latest AI systems continuously adapt to the learner’s progress and can thus finetune the feedback (Amin, 2023, p. 8). Such timely, individualized and elaborate feedback can help to create a “more engaging learning experience for the students” (Essel et al., 2022, p. 2). The learners’ perspective will therefore be inspected next.

### 3.2. Chatbots as Learning Assistants

As learning assistants, chatbots can fulfill multiple functions. Zhang et al. (2023) reviewed the advantages that were pinpointed in previous studies from a pedagogical perspective and summarized them by using the acronym “RAISE”, which stands for “Repetitiveness, Authenticity, Interactivity, Student-centredness, and Enjoyment” (p. 13). They will be discussed in the current contribution as well, albeit in different order.

To start with, one major advantage of chatbots as learning assistants is that they can provide students with rich linguistic input, which is particularly beneficial for language learning. AI-driven chatbots might thus serve as convenient and accessible language partners, especially when foreign language instructors are not readily available (Haristiani, 2019, p. 4; cf. Fryer et al., 2017; 2020; Ruan et al., 2021, p. 434). They offer learners an opportunity to practice language skills endlessly, enabling them to explore a wide variety of language input that they might not encounter with traditional classroom partners (Fryer et al., 2020, p. 16). However, it needs to be borne in mind that the database on which the AI system has been trained is limited in scope linguistically, culturally, and temporally (Ray, 2023, p. 147; see section 4.3). Especially when English is used as a lingua franca by people with different linguistic backgrounds, human language partners are likely to produce more authentic input.

Additionally, though, chatbots are capable of repeating the same material as many times as needed (Haristiani, 2019, p. 4; cf. Fryer et al., 2020, p. 16). Hence, students can progress at their own pace and can continuously become more self-directed in their learning (Essel et al., 2022, p. 4). Citing the study by Cai et al. (2020), Zhang et al. (2023) remarked “that the chatbots allowed students to practise repeatedly and make constant self-correction based on the chatbot feedback, helping them gradually enhance their accuracy and fluency of speaking” (p. 10). Likewise, ChatGPT can offer tailored assistance to guide students to improve their skills and help them in problem-solving (Qureshi, 2023, p. 6), such as by suggesting learning techniques and resources that they could engage with (cf. Cotton et al., 2023, pp. 2–3; Dai et al., 2023b, p. 5). Hence, they can take control of their learning and progress at their own pace without being dependent on teachers’ availability (Essel et al., 2022, p. 4).

Indeed, after the presentation of learning materials on rank one, performance evaluation and feedback provision were the second most frequent types of pedagogical application found in Zhang et al.'s (2023) review of chatbot studies (p. 6). To exemplify, learners can formulate specific feedback requests and enter them into the ChatGPT interface, together with their work in progress. For this, they obtain immediate feedback at any time and place (Essel et al., 2022, p. 14; Haristiani, 2019, p. 4; Mageira et al., 2022, p. 3) “[t]hrough a broad range of devices” with internet access (Fryer et al., 2017, p. 463). The immediate feedback helps to eradicate errors at an early stage already (Wahyudi, 2023, p. 3). Moreover, the speedy responses might lead to greater engagement, especially if learners consider them as relevant. As such, the availability of AI chatbots can “promote learning outside of the traditional classroom setting” (Alshahrani, 2023, p. 2035).

There also exist chatbots or extensions that do not only allow for written exchanges, but also for audio interactions (Fryer et al., 2017, p. 463). Moreover, ChatGPT can provide feedback in multiple languages, making it useful for language learners who want to improve their skills in languages other than their native language. These multilingual interactions (Kohnke et al., 2023, p. 4) might be particularly profitable for foreign language learners at the beginner level (see final chapter 17 by Schluer in this volume).

Related to that, the AI system can assess learners' proficiency and provide them with tailored tips for improvement (Frackiewicz, 2023; cf. Farrokhnia et al., 2023, p. 5). In addition, ChatGPT remembers past conversations (Farrokhnia et al., 2023, p. 3), which enhances its potentials for formative feedback. In the future, a more comprehensive consideration of individual learning histories also appears likely to further improve the personalized support (cf. Qureshi, 2023, p. 6). So far, ChatGPT has already demonstrated the ability to provide valuable process-focused feedback, which is considered effective in shaping students' task strategies and developing learning skills (Dai et al., 2023b, p. 5). It thus moves beyond a simple checking of grammar and linguistic correctness, as many AWE tools do. This personalized guidance can also boost learners' motivation (cf. the review by Kohnke et al., 2023, p. 2), especially when paired with positive feedback (see Dai et al., 2023b, p. 4). Crucially, ChatGPT provides explanations, not only for negative feedback but also for positive feedback. This way, users learn about the reasons for the (in-)correctness of the text that they have entered. This can be valuable for raising learners' language awareness and for improving their future performance. Further affective and (meta-)cognitive benefits will be elaborated next.

### 3.3. Affective Benefits for Learners

Chatbots can contribute to the learning experience not only (meta-)cognitively, but also affectively, for example regarding motivation, interest, and self-efficacy (Zhang et al., 2023, p. 12). Many students are scared or too shy to seek help and ask questions to their teachers, especially when they feel that their request might bother other students or the teacher or slow down the overall progress in class. This is particularly true of classroom contexts in which learners perceive a high power distance (cf. Kohnke et al., 2023, p. 2). Students might also be afraid of making mistakes and of the resultant negative feedback (Oktaria & Soemantri, 2021; Verleger & Pembridge, 2018; both cited in Essel et al., 2022, p. 4). It could make them appear incompetent (Fryer et al., 2020, p. 16), especially when

“communicating in a foreign language” (Ruan et al., 2021, p. 435). Accordingly, they do not dare to seek assistance for reasons of shame and lacking self-confidence. Chatbots, in turn, could reduce language anxiety by providing a non-authoritative (Li et al., 2023, p. 38) and non-judgmental learning environment (Essel et al., 2022; Ruan et al., 2021) in which the learners can ask questions, request feedback (Essel et al., 2022, p. 3) and practice their language skills (Amin, 2023, p. 8; Brinegar, 2023, p. 224; Fryer et al., 2020, p. 16; Hoang et al., 2023, pp. 151–152; Kohnke et al., 2023, p. 2). They provide feedback without any biases or preconceptions, which makes them a non-judgmental source of feedback that can help users improve their writing skills without feeling embarrassed or judged (cf. the review by Zhang et al., 2023, p. 14). It has consequently been argued that chatbots offer users comfortable and efficient assistance by responding directly to their problems in an attractive and friendly manner (Adamopoulou & Moussiades, 2020, p. 1). Through engaging answers, they can even make the learning experience more enjoyable (Adamopoulou & Moussiades, 2020, p. 1; Ruan et al., 2021, p. 434; cf. Tseng et al., 2020, cited by Zhang et al., 2023, p. 12).

Overall, the following affective benefits of chatbots for feedback provision stand out (cf. Lechler et al., 2019): Feedback can be sought at any time and about almost any topic in a risk-free environment. The instantaneously provided feedback can be motivating and confidence-building (Frackiewicz, 2023; see also Zhang et al., 2023, p. 12; Wahyudi, 2023, p. 3), particularly since it may give learners a feeling of acknowledgment, appreciation and care. Empathy can be conveyed through emotionalized language and the recognition of previous efforts. It can stimulate an achievement orientation and foster a sense of growth. The positivity in the formulation of feedback can lead to a faster adoption of suggestions even though it gives the learners a sense of choice.<sup>2</sup> This, in turn, can promote learner agency and autonomy. As such, it can increase students’ willingness to learn and facilitate competence development. One major catalyst of this continued improvement probably is the conversational nature of the chat interactions, which can help them stay motivated (Frackiewicz, 2023) and will be treated next.

### 3.4. Human-Like Conversation

As Kohnke et al. (2023) emphasized, “[i]ntelligent chatbots have caught the attention of language educators because they can interact with learners in their target language instantly and realistically (Fryer et al., 2020; Lee et al., 2020)” (p. 1). They can simulate human conversation (Frackiewicz, 2023) and might even promote social interaction abilities (Ruan et al., 2021, p. 434). Chatbots have therefore alternatively been termed “conversational agents” (cf. Zhang et al., 2023, p. 4). As ChatGPT has been trained on a large data corpus, the responses sound more natural and context-specific than those from preceding chatbot types (Dai et al., 2023b, p. 1). Apart from written interactions, voice conversations are also possible with chatbots to a certain extent (cf. the review by Essel et al., 2022, p. 2; Kohnke et al., 2023, p. 1).

---

2 However, in ChatGPT, there is also the possibility for custom instructions in which users might specify their need to get brief answers in a more neutral tone (see section 5.3 in this chapter).

The two-way interactions (Kohnke et al., 2023, p. 2) are further enhanced by improvements to ChatGPT-4. To help overcome writer's block (Duval et al., 2020, cited in Cotton et al., 2023, p. 2) and engage the users, ChatGPT introduced prompt examples at the beginning of each chat (OpenAI via Twitter on August 03, 2023). Furthermore, with the August 2023 update, ChatGPT additionally suggests possible ways to continue the conversation (OpenAI via Twitter on August 03, 2023).

ChatGPT can hence solve several problems that had been reported about other chatbots in the previous literature, such as the limited personalized support, its "inauthentic" appeal (Fryer et al., 2017, p. 463), and the reduced range of possible areas of application. For example, Adamopoulou and Moussiades (2020) noted that "human–chatbot communication has noticeable differences in the content and quality in comparison to the human–human discussion" (p. 1), whereas Dai et al. (2023b) argued that this disadvantage has been resolved to a certain extent, as "specific training allows ChatGPT to generate more natural-sounding and context-specific responses" (p. 1). Moreover, the user interface is "simple and intuitive" (Essel et al., 2022, p. 2), making it easy to use (*ibid.*, p. 12) and lowering the threshold for interaction even further.

Due to the chatbot's conversational nature, it is frequently applied in language learning (Amin, 2023, p. 8; Haristian, 2019; Zhang et al., 2023, p. 9), e.g. learning EFL/ ESL (Fryer et al., 2017; 2020) or Japanese (Haristian, 2019, p. 3). In that respect, chatbots can be used for a wide range of assignments, depending on their specific features. For example, ChatGPT can be utilized for written assignments to produce feedback on grammar, spelling, writing style and text structure. Mondly and Duolingo more specifically support the learning of grammar and vocabulary. English Liulishuo, in turn, can be utilized for oral practices, particularly by Chinese-speaking students practicing speaking in English (Ruan et al., 2021, p. 436). With its quick development, chatbots can now be employed in many other subjects as well, such as computer science (Abbasi et al., 2019; Qureshi, 2023), nursing education (Chang et al., 2022), engineering (Kohnke et al., 2023), etc.

### 3.5. Idea Generation and Reflection

Apart from submitting a draft to the system, learners can benefit from generative AI already before writing a paper. In addition to corrections, it could be deployed to generate ideas and provide information about a certain topic. Learners can thus brainstorm ideas and collect different perspectives, either before starting to write or while crafting a paper.

Whenever students ask for feedback on a piece of work, the system could stimulate further reflections (Schmied, 2023, p. 12). Learners, however, should not directly copy the suggested ideas (as this would constitute academic fraud, see section 4 below), but regard them as a starting point for critical reflection and searches in approved resources and reliable databases (Schmied, 2023, p. 12). Moreover, the stereotypical output of the AI system does not necessarily comply with the standards that are set by a particular discipline or teacher, nor does it represent the uniqueness of an academic writer, which means that the generated output requires modification and refinement (Schmied, 2023, p. 12).

To give an example, Cotton et al. (2023) composed their paper entitled "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT" almost completely by making use of ChatGPT. Mainly, they relied on the prompted ChatGPT output while only

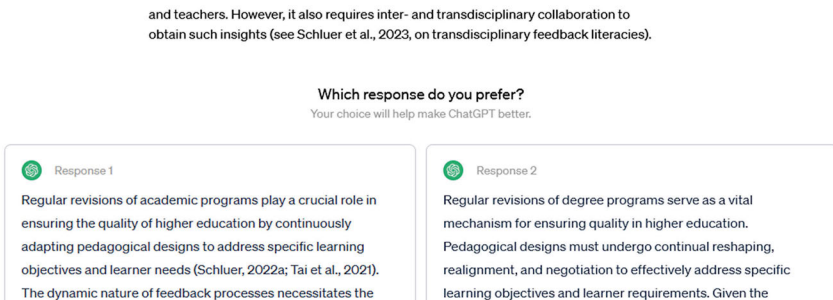
adding the sub-headings and references manually (p. 6). This article has rapidly attracted widespread attention, with almost 50,000 reads within a couple of months. In the paper, they list several advantages and limitations of ChatGPT as well as potential instructional techniques, as suggested by ChatGPT itself. Some of these points are also cited in the present contribution and backed up by additional evidence.

### 3.6. Feedback to the Chatbot

So far, we have seen that chatbot feedback can be useful for teachers and learners alike. Teachers can ask ChatGPT to give feedback on students’ texts, and learners can utilize chatbots for self-feedback. In addition, the AI system relies on the users’ feedback to further improve its self-learning mechanism. As Essel et al. (2022) explain, “AI relates to systems or machines that mimic human intelligence and self-alter based on accumulated data (Angelov et al., 2021)” (p. 3). The AI-powered chatbot can hence “continuously learn from prior interactions [and thus] improve over time” (Kohnke et al., 2023, p. 2).

On the start page, OpenAI explicitly encourages the users to provide feedback about the accuracy of the responses (“This system is optimized for dialogue. Let us know if a particular response was good or unhelpful”). For this, they may use the thumbs-up or thumbs-down symbols that are placed below each generated response (Kohnke et al., 2023, p. 9). In a pop-up window that appears, they can provide additional feedback by stating the problem and suggesting potential solutions in an open-response field and/or by ticking one or several of the given options (“This is harmful/ unsafe”, “This isn’t true”, “This isn’t helpful”). Furthermore, for regenerated answers, they can vote with a thumbs up, thumbs down, or equal sign. Sometimes, ChatGPT directly produces two solutions, asking the user to select the one they find more suitable (see Figure 1). Additionally, users can visit OpenAI’s discord server to give more detailed feedback (<https://discord.gg/openai>).

Figure 1: Response Preferences in ChatGPT



Another way to provide feedback to the system is to reply to the generated output in the chat itself, e.g. by specifying in how far the response was useful or incorrect. Often, it already helps to formulate the prompt in a different and more precise manner so that ChatGPT can resort to relevant data (see the importance of adequate prompts in section

5). Moreover, users could ask the chatbot to review its own output and provide feedback on it (Gharani, 2023).

Finally, there is still another contribution that ChatGPT can make to the feedback process, which is fostering users' feedback literacy and learner agency. The details will be shared in section 5, while a brief preview will be given in the subsequent subsection.

### 3.7. Chatbots as Feedback Models

Above, we have discussed that AI can be used for language modeling, providing learners with examples of correct language and helping them develop a more accurate understanding of grammar, vocabulary, and syntax. Moreover, ChatGPT has the capacity to model various text genres (Kohnke et al., 2023, p. 4), such as business letters or academic essays. This capability can also be exploited for developing learners' and (pre-service) teachers' feedback literacy.

A feedback process ideally starts with learners actively seeking guidance from their teachers, as this presupposes self-reflection about their own work in progress. One feature of ChatGPT is to generate questions about a text or topic (Kohnke et al., 2023, p. 8), including questions that could be asked to initiate a feedback dialogue. Hence, chatbots are not only capable of responding to feedback, but also to model the initiation of a feedback exchange (see example in section 5). This, in turn, can stimulate further questions from the students (Essel et al., 2022, p. 4), lower the threshold of daring to ask questions, and may ultimately foster learners' agency in the feedback process.

The same is true of follow-up conversations and actions: Once learners have received feedback from their teacher or peer, they can utilize ChatGPT to seek guidance in developing an action plan and to clarify open questions that the previously provided feedback did not answer. As shown in the study by Dai et al. (2023a), the chatbot-generated feedback often appears to be more readable and comprehensible as well as process-oriented towards further improvement. Moreover, students could use ChatGPT again to obtain potential follow-up questions that they could ask to their teacher or peer.

To seize these affordances of ChatGPT, the process should be explained and introduced by the teacher beforehand. Accordingly, educators themselves need to be aware of the chances and challenges of ChatGPT. They can utilize a feedback conversation generated by ChatGPT in order to analyze its structural components and linguistic features together with the learners (see samples in section 5). In the classroom, they should also brainstorm alternative phrases and questions that could replace the ones suggested by the generated model. Furthermore, they can discuss modifications of the model for different assignment types and feedback modes (e.g. written, oral, or video feedback).

Teachers may ask ChatGPT to create texts and feedback in different proficiency and complexity levels (Farrokhnia et al., 2023, p. 5; Schmied, 2023, p. 10). Moreover, the tone and length of the feedback can be adjusted when prompted accordingly (see Farrokhnia et al., 2023, p. 5, for examples). Chatbots can thus be used to generate reviews that exemplify positive, negative and balanced feedback (Schmied, 2023, p. 11). These could be used profitably for feedback training in teacher education or in classrooms. Pre-service teachers can try to identify characteristics of balanced feedback and reflect on the impact of overly positive and negative feedback, respectively. Similarly, by closely analyz-

ing the linguistic features, especially in comparison to overly positive and negative examples, learners' language awareness can be raised (cf. Schmied, 2023, p. 11). This might also leverage their own feedback provision, e.g. when being asked to engage in peer feedback processes for the first time.

Finally, it would be important to critically discuss the ChatGPT output with the learners when the AI system is used for self-feedback on a submitted draft. They need to become aware of the typical nature of the prompted feedback and the limitations of the AI output. Unless prompted differently, the generated feedback is often formulated in a positive and inoffensive manner (e.g. "In Italian, we prefer to say X and not Y"; Schmied, 2023, p. 12). Linguistically and culturally, this can be challenging for learners who have difficulties in understanding implicit and hedged suggestions. Moreover, they need to realize that the generated suggestions can be wrong, thus requiring their own critical thinking and search skills to verify the correctness of the AI output.

### 3.8. Learning Gain

While it has been widely argued that ChatGPT (and similar chatbots) can be beneficially deployed to assist teachers and learners in the feedback process, the actual learning gain is disputed (Farrokhnia et al., 2023, p. 11; cf. the review by Zhang et al., 2023), especially from a long-term perspective. Moreover, the experimental evidence is still limited to a few disciplines and assignment types. For example, Pardos and Bhandari (2023) "found that 70% of the hints offered by ChatGPT led to positive learning gains for students" with regard to "elementary and intermediate Algebra topics" (as cited by Qureshi, 2023, p. 3). Essel et al. (2022) focused on a multimedia programming course and concluded that "the experimental cohort who engaged with the chatbot performed better than students in the control cohort who interacted with the course instructor" (p. 14). Conversely, Basic et al. (2023, cited by Qureshi, 2023, p. 3) conducted a small-scale study and inferred "that ChatGPT does not necessarily improve essay quality." These mixed results lead us to a discussion of the challenges and limitations of chatbot feedback.

## 4. Challenges and Limitations

On the start page of ChatGPT, OpenAI already draws the users' attention to the limitations of the chatbot (at least pertaining to ChatGPT 3.5 as of August 4, 2023):

- May occasionally generate incorrect information
- May occasionally produce harmful instructions or biased content
- Limited knowledge of world and events after 2021

These and additional limitations will be discussed in the present section with particular regard to feedback processes.

#### 4.1. Difficulties in Interpreting Input and Providing Accurate Output

One major limitation of ChatGPT (and other chatbots) is that their capability to interpret the input and generate adequate output is limited (e.g. Farrokhnia et al., 2023). Since ChatGPT has been trained on a limited (albeit huge) database, AI chatbots could have problems in accurately recognizing and interpreting natural language, particularly in languages other than English (Mageira et al., 2022, p. 3). For instance, chatbots may have difficulties in understanding users' questions and prompts (Qureshi, 2023, pp. 3, 6) so that the resultant feedback might be of limited use. Notably, chatbots may also struggle to answer complex, open-ended, or unexpected questions (Bailey et al., 2021, cited in Zhang et al., 2023, p. 13), causing frustration among students or even erroneous learning (cf. Mageira et al., 2022, p. 3). A main reason for this is that a chatbot usually has less specific contextual background knowledge about a users' needs and learning trajectory than a teacher or fellow student who is well familiar with the learner. Moreover, the AI system does not have the profound knowledge of a human subject matter expert and may thus lack a thorough understanding of underlying concepts (Qureshi, 2023, p. 3; cf. Grassini, 2023, p. 6). Also, it "lacks the human ability to assess the credibility of the data it was trained on" (Farrokhnia et al., 2023, p. 7, with reference to Lecler et al., 2023). This can lead to inaccurate feedback and recommendations for specific learning contexts. For instance, Dai et al. (2023b) remarked that "ChatGPT could not offer a reliable assessment of student performance compared to the instructor" (p. 5). To increase the reliability, it would be necessary to "train ChatGPT by feeding examples including students' assignments of different quality and associated golden feedback that accurately evaluates their performance" (Dai et al., 2023b, p. 5). In the end, "chatbots are only as good as the datasets of text and statistical patterns that they are trained on", as Aithal and Silver (2023) noted (n.p.).

Another problem is that the ChatGPT output quite often looks promising at the textual surface, but demonstrates difficulties at the deeper, conceptual level (cf. the reviews by Farrokhnia et al., 2023, pp. 6, 8, and by Grassini, 2023, p. 6). Indeed, ChatGPT has attracted attention due to its capacity to produce "credibly sounding texts" (Schmied & Ivanova, 2023, p. v). While the generated texts might look convincing at first sight (especially due to its linguistic accuracy, clear structure and the use of fitting cohesive devices), they "can give false confidence to users" since the information provided may not be fully correct or could even contain invented links and references (Schmied & Ivanova, 2023, p. v). For example, Qureshi (2023) discovered inaccuracies and inconsistencies in the code that was generated for a programming assignment, which affected overall performance. Furthermore, Cotton et al. (2023) spotted fictional references in the ChatGPT output and replaced them by genuine ones to compose their paper (p. 7). In addition, the AI does not always provide stylistically appropriate answers. Sometimes, it is already helpful to give very specific commands to ChatGPT, e.g. about the genre of the submitted text. Schmied and Ivanova (2023) therefore argued that users not only need the skills to write adequate prompts but also critical text-analytical skills to scrutinize the generated output (p. v). Similarly, Dai et al. (2023b) reasoned that "prompt engineering" should play a greater role in the future to ensure higher effectiveness (p. 5). Moreover, the AI system also needs more training and a more reliable mechanism for controlling the quality of

the generated output. Nevertheless, learners should check the information that ChatGPT gives in order to prevent problems in their assignments.

#### 4.2. Immediate Responses Instead of Guidance

While it has been argued that ChatGPT can be a useful facilitator of learning and feedback processes, it may not fulfill this guiding role in a pedagogically adequate manner. It tends to give corrections rather quickly instead of guiding the users to come up with the correct solution on their own, e.g. through pinpointing the erroneous passage and asking questions about it. This is because ChatGPT's primary purpose is to generate responses to user inputs based on the patterns it has learned from the massive amount of data training. Thus, it may not always be capable of guiding users through the process of identifying and correcting errors themselves. As a result, an over-reliance on AI might diminish learners' ability to self-correct and engage in "critical thinking and independent problem-solving" (Ray, 2023, p. 140). However, as compared to the usually brief corrections by human assessors, it formulates the suggestions in an action-oriented manner, and users often need to access relevant resources on their own or ask follow-up questions to the system. Nevertheless, "it may not necessarily improve critical thinking and problem-solving skills" (Qureshi, 2023, p. 1) as a teacher could do. This is primarily due to the dearth of detailed back-and-forth interactions and the systems' "[i]nability to think in depth" (Essel et al., 2022, p. 13). Sometimes the chatbot also generates almost the same response when asked to reformulate the reply, or it repeats the same response for a different question (Essel et al., 2022, p. 13). Efforts are made to improve ChatGPT's capacity, e.g. by automatically suggesting prompts for follow-up interactions (OpenAI via Twitter on August 03, 2023). Probably, though, chatbots will not be able to raise further questions as human language partners would do based on their individual and interpersonal knowledge.

Another difficulty arises from the often definite and "authoritative nature of the responses", which "can lead users to take them at face value" (Kohnke et al., 2023, p. 9), even though the output might be erroneous. These ready-made solutions could decrease learners' ability to think critically and autonomously as well as undermine their unique identity as a writer (Schmied, 2023, p. 12). Quite often, the output sounds stereotypical of a particular genre without creative deviations that might be necessary to solve a new task and could help advance a field further (cf. Qureshi, 2023, p. 3). In addition, the replies are often wordy and repetitive (Adamopoulou & Moussiades, 2020, p. 13; Kohnke et al., 2023, p. 9; Ray, 2023, p. 148), thus lacking the conciseness that might be needed in a specific context. When prompted to list its own limitations regarding feedback provision, ChatGPT acknowledges that it can only offer "limited feedback on creative writing: While ChatGPT can provide feedback on grammar and style, it may not be as effective in providing feedback on creative writing, such as poetry or fiction" (OpenAI, 2023b). At present, chatbots are therefore not yet suitable for all kinds of assessment tasks and learning objectives (Lo, 2023; Schmied & Ivanova, 2023).

Finally, ChatGPT has not yet been optimized for feedback purposes with regard to the way in which the comments are provided (Escalante et al., 2023, p. 14). In contrast to AWE programs, it does not annotate an existing text through in-text comments or col-

oring (as e.g. Grammarly does) but produces feedback in prose form. While such lengthy comments provide language learners with additional input, they might impose too much cognitive load for students at lower proficiency levels (cf. Li et al., 2023, p. 38).

### 4.3. Ethical Concerns

There are also several ethical concerns, especially in the academic sphere, about ChatGPT and other AI-powered chatbots (e.g. Amin, 2023). The generated output is based on existing texts, but usually lacks proper citation. If given, the sources are often incorrect or non-existing (Kohnke et al., 2023; Schmied & Ivanova, 2023, p. v). Hence, the question arises whether the utilization of ChatGPT constitutes academic fraud or contributes to academic dishonesty. In particular, copied passages from chatbots could be classified as plagiarism (cf. Cotton et al., 2023, pp. 1, 3; Kohnke et al., 2023, p. 8; Qureshi, 2023, p. 1; Schmied & Ivanova, 2023, p. v). On the other hand, the output from ChatGPT usually is not a direct copy from the original texts on which it has been trained. It is thus a borderline case that has not previously been captured by definitions of plagiarism, but which seems to be located in a grey zone of academic misconduct. Recent improvements, e.g. through Bing.ai search (Microsoft Copilot) or Perplexity.ai, could, however, help to reduce the amount of improperly cited or uncited works. At any rate, if direct copies or paraphrases are utilized without an acknowledgment of the source, this would count as plagiarism (Kohnke et al., 2023, p. 8).

Another problem from an ethical perspective is the linguistic, cultural, ideological, and temporal bias of the database on which ChatGPT has been trained (Ray, 2023, p. 147). The majority of the data corpus consists of texts written in English, which were then translated into other languages (Kohnke et al., 2023, p. 9). In addition, the corpus mostly represents written language rather than spoken language (Kohnke et al., 2023, p. 9). These texts also tend to mirror a specific sociocultural stance and hence cannot be regarded as culturally neutral even though the very use of AI may imply such neutrality (Kohnke et al., 2023, p. 9). Additionally, it is also restricted in time since only a limited portion of all available texts have been fed into the database (Ray, 2023, p. 147). In particular, this refers to old texts as opposed very recent ones that are not yet part of the data corpus. Hence, even though ChatGPT has been trained on a massive database, it only provides a limited perspective of the existing multitude of texts and resources. Accordingly, chatbot responses may be influenced by a cultural bias present in the source texts and in the working of the algorithms, potentially impacting students from diverse cultural backgrounds in language education (Kohnke et al., 2023, p. 9). Moreover, due to its self-learning mechanism, the chatbot might reinforce these existing biases and stereotypes even further (cf. Amin, 2023, pp. 7, 11). Due to recurring associations, some topics might appear as trending, whereas less common perspectives could diminish, thus undermining innovation and creativity (cf. the novelty bias cited by Ray, 2023, pp. 147–148). For these and other reasons, “[t]eachers worry that students will become too reliant on the chatbot to generate ideas, questions, and conversation topics rather than thinking creatively and originally” (Brinegar, 2023, p. 228). For users, it would be important to consider these inherent biases (Qureshi, 2023, p. 6) and the resultant cultural appropriation as well as possible hegemony of particular writing styles and

worldviews (cf. Ray, 2023, pp. 147–148). The effectiveness of chatbots in education thus heavily depends on the quality and quantity of information and knowledge included, necessitating an effective data collection mechanism (Chang et al., 2022, p. 17; Kuhail et al., 2023). Crucially, users need to be aware of these limitations and develop strategies in handling them confidently (Aithal & Silver, 2023).

Furthermore, there are potential risks related to privacy and data security when using chatbots, which raises additional ethical concerns (Adamopoulou & Moussiades, 2020, p. 13; Grassini, 2023, p. 7; Ray, 2023, p. 141). They pertain to the compilation of data for the AI system, notably the collection of user data, as well as an unequal access to AI technologies resulting from differences in the free and premium versions that require subscription (cf. Amin, 2023, p. 7).

#### 4.4. Novelty Effects and Other Affective Factors

As with any unprecedented technology, the use of chatbots may cause novelty effects, which makes them an attractive feedback companion upon initial exploration, but might make users lose interest over time (Fryer et al., 2020, p. 16; Li et al., 2023, p. 38; Yin et al., 2021, quoted by Zhang et al., 2023, p. 13). Learners' motivation and engagement may decrease further (Kuhail et al., 2023, p. 1005) as soon as they become aware of the limitations of the chatbot. For example, they might realize that “chatbots are often unable to answer complex or unexpected questions [...] and produce inappropriate dialogue flows”, which can lead to frustration (Mageira et al., 2022, p. 3) or even faulty knowledge acquisition.

Overall, chatbot communication may feel inauthentic and unreliable, as it often lacks the empathetic qualities of human-human interaction (Adamopoulou & Moussiades, 2020, p. 1; Li et al., 2023, p. 38; cf. the review by Zhang et al., 2023, p. 13). Some chatbots are perceived as predictable and redundant, with a limited ability to memorize previous conversations, leading to repetitive interactions and user frustration (Mageira et al., 2022, p. 3). For instance, ChatGPT responses often follow the same standard pattern (introduction, main part, conclusion) and utilize similar formulations. This can result in boredom and also have an impact on the way the feedback is perceived and used (cf. Adamopoulou & Moussiades, 2020, p. 1; Dai et al., 2023b; Fryer et al., 2017, p. 463). However, it is also possible to customize the output to a certain extent, as described in section 5.3 below.

Due to the still-prevailing lack of emotional intelligence (see the review by Grassini, 2023, p. 6), chatbots are not capable of understanding or empathizing with a user's emotional state, which, however, is important for feedback uptake. In fact, the long-term learning gain with chatbots has not been extensively studied, and comparisons with traditional non-intelligent systems are scarce (Ruan et al., 2021, p. 435).

#### 4.5. Costs and Server Capacity

Qualitatively better chatbots (or premium versions, including ChatGPT) come with additional costs (cf. Kohnke et al., 2023, p. 2). Lower-performing language models may also suffer more often from server overload or slower loading times. The question of access and availability can also be regarded as another ethical issue in addition to the other

points that were listed above. The “most capable model”, as stated on the ChatGPT-4 website in 2023, is currently only available to paying users. Such a distinction between paid and free versions certainly puts individuals and countries with a low economic status at disadvantage.

Another area of concern relates to the energy consumption resulting from the enormous computing resources that are needed to create and sustain an AI system (Ray, 2023, p. 141). These, in turn, can have negative effects on the environment (Ray, 2023, p. 141).

#### 4.6. Skills Required

Implementing chatbots for specific teaching and learning needs requires a high level of experience and skills, which may be a hurdle for educators (Wollny et al., 2021, p. 12) and learners (Woo et al., 2023). For instance, in the study by Woo et al. (2023), the “[s]tudents reported high cognitive load” when using the chatbot, especially to formulate adequate prompts (preprint p. 3). Also, the development of critical thinking and problem-solving skills presents a challenge, given that chatbots fulfill an increasingly sophisticated range of functions. Therefore, the following section will give advice about the implementation of chatbot feedback in educational contexts.

### 5. Suggestions for Implementation

The present section will outline important preconditions as well as suggestions for using chatbots for self-feedback, assessment, and feedback training.

#### 5.1. Required Equipment and Competencies

Chatbots can be used on tablets, laptops, desktop PCs and mobile phones. The free version of ChatGPT can be accessed at <https://chat.openai.com/>, whereas the premium version comes with additional costs. Other AI chatbots are Google Gemini (<https://gemini.google.com/app>), Bing AI (Microsoft Copilot), Perplexity AI (<https://www.perplexity.ai>) and several others.

Crucially, a pedagogically valuable as well ethical use of chatbots for feedback purposes requires an array of competences from the users, i.e. teachers and learners (Kohnke et al., 2023, p. 10). In their article, Kohnke et al. (2023) discuss several dimensions of digital competence that appear to be needed for using ChatGPT. This involves awareness of the features and of the limitations of ChatGPT (see sections 3 and 4; cf. Javier & Moorhouse, 2023, p. 1) and staying up-to-date with regard to the newest changes (Kohnke et al., 2023, p. 10). An understanding of the underlying mechanisms of the chatbot is also beneficial for becoming critically aware of its biases and the risks involved in its usage (Javier & Moorhouse, 2023, p. 1; Kohnke et al., 2023, p. 10).

Moreover, educational institutions should set up clear “guidelines for using such tools, modify their teaching and assessment practices, and consider how to best prepare students for a world where AI-driven digital tools are a normal part of daily life” (Kohnke et al., 2023, p. 10). Hence, AI tools cannot or should not be ignored, but a safe and ethical

use of the tools should be practiced (Kohnke et al., 2023, p. 10). Learners' attention should be drawn to legitimate resources and tools for working on their assignments (Qureshi, 2023, pp. 6–7) as well as to the concept and consequences of plagiarism (Cotton et al., 2023, p. 3; Qureshi, 2023, pp. 7–8). In addition, they could discuss strategies that help them verify the suggestions and identify incorrect information of chatbot-generated output (Qureshi, 2023, p. 1). In that respect, teachers could demonstrate and practice the utilization of chatbots and other resources together with the students while closely monitoring learners' usage of these tools (Qureshi, 2023, p. 7). Overall, an iterative and scaffolded approach to utilizing AI chatbots is recommended for sensitizing students to their responsible use (cf. Javier & Moorhouse, 2023).

To pre-empt an overuse of AI tools, educators should develop tasks that cannot easily be solved by a bot (Cotton et al., 2023, p. 3) and which alert students to existing limitations and dangers. This includes assignments that require students' own creative and "critical thinking, problem-solving, and communication skills" (Cotton et al., 2023, p. 5), e.g. through specific real-world projects that are solved in a team. Furthermore, continuous guidance through in-process feedback conversations based on regular draft submissions helps to ensure that students craft their works on their own and have understood the underlying notions. In that respect, regular in-process discussions and follow-up interviews can prove helpful (cf. Qureshi, 2023, p. 7). Hence, a combination of different assessment techniques and feedback channels is recommended (Cotton et al., 2023, p. 3; Qureshi, 2023, p. 7). In that respect, peer approaches appear promising in which students explore AI tools collaboratively within the protected space of the classroom. In the end, students will not always be able to rely on chatbots, especially when engaging in face-to-face conversations that require immediate answers and social skills.

For this, feedback skills and teamwork skills play a central role. Learners should thus be made aware of the importance of feedback as well as of their agentive role in the feedback process (see section 2). Indeed, the utilization of ChatGPT could promote students' agency at different phases of the feedback process and in different situations and modalities, as will be discussed further below. Moreover, it can aid teachers' feedback and learners' self-correction process, as will be shown next.

## 5.2. Formulation of Prompts

Apart from a critical analysis of the generated output, the usefulness of ChatGPT depends on the input that the users provide to the system (Schmied, 2023, p. 11). Notably, an adequate formulation of prompts plays a vital role (Javier & Moorhouse, 2023, pp. 2–4; Qureshi, 2023, p. 6; Schmied, 2023, p. 8; Woo et al., 2023, preprint p. 8). These prompts should be clear and specific and provide sufficient contextual information (Ray, 2023, p. 126). Consequently, the successful use of ChatGPT for feedback purposes hinges "on the quality of the questions" (Kohnke et al., 2023, p. 8), together with an awareness of relevant assessment criteria (desired output). Both of them are important components of feedback literacy, which means that practicing the provision of prompts can help students to develop this capacity. For example, during feedback training (see also section 5.5 below), students could create an assessment rubric together with the teacher, apply it, and evaluate the resultant chatbot output. However, depending on the specificity

and complexity of the assignment, learners might notice the limitations of the system. Accordingly, they might try out different prompts, ask ChatGPT to help them improve the prompts they have entered, or even prompt ChatGPT to produce suitable prompts (cf. Gharani, 2023). One common strategy in that regard is “iterative prompting” (Ray, 2023, p. 126). This means that users continuously refine the prompt or divide it into sub-questions to obtain the desired guidance (see also Woo, 2023, and Woo et al., 2023, about the iterative process of chatbot interactions and prompt engineering). These specific feedback requests might refer to grammar, spelling, pronunciation (Frackiewicz, 2023), source use, paragraph structure, and style. Escalante et al. (2023), for instance, prepared a set of prompts pertaining to different assessment criteria, such as “Using simple language, examine the use of transitional phrases in the paragraph. Are they appropriate for academic writing? How could they be improved? Start your feedback with [...]” (p. 16).

Likewise, teachers can ask ChatGPT to generate an assessment rubric for a specific assignment. The results could be compared to the rubric that has been co-created with the students, leading to potential modifications or evidencing the system’s limitations. The finalized rubric can then be fed into ChatGPT again and applied to student work (Cotton et al., 2023, p. 3). As Dai et al. (2023b) showed, there are benefits and drawbacks to such a procedure. Similarly, learners are likely to experience affordances and challenges when utilizing ChatGPT for the purpose of self-feedback, to which we will turn next.

### 5.3. Implementation of ChatGPT for Self-Feedback

To utilize ChatGPT for self-feedback, learners should first reflect on their work’s strengths and areas of difficulty themselves. For this, teachers could provide learners with reflective questions (cf. chapter 2 on course re-design by Schluer in this volume), such as

- What do you think are the strongest aspects of your assignment? OR: What are the areas you feel confident about in your assignment?
- What areas of your assignment do you think need to be improved? OR: What are the areas you are unsure of in your assignment?
- I would particularly like feedback on (list up to three specific areas): ...

These are self-reflective questions that students were asked to attach to their draft submissions in some of my classes (as inspired by Winstone & Carless, 2020, pp. 108, 110; see also Schluer, 2022, pp. 37, 42–43).

In a subsequent step, they should go to the ChatGPT platform (or another generative AI tool), sign up and log in, respectively (see also the information about chatbot feedback on the “Digital Feedback Map” by Schluer, 2023). Next, they need to create a prompt that instructs ChatGPT on the specific assignment and the type of feedback that they desire (see above). For example, learners could use a prompt like: “Please review the attached essay and provide constructive feedback on its structure, clarity, and overall coherence.” For this, students need to upload or copy-paste their writing assignment into the ChatGPT platform. They need to ensure that the document is well-formatted and legible for the AI to process it accurately.

Afterwards, they should closely analyze the ChatGPT output. For example, the users might notice that the responses are too elaborate or too general and unrelated to their needs. To overcome this, there are at least four options. Since ChatGPT has been designed for back-and-forth conversations, (1) users can interact with it to ask clarifying questions or request more detailed feedback (Kohnke et al., 2023; see also Abdelhalim, 2024, p. 9). To give an example:

**User:** “Could you focus on the essay’s introduction and provide suggestions for improvement?”

**ChatGPT:** “Certainly! The introduction could be made more engaging by ...”

Learners should continue to review the feedback provided by ChatGPT and consider its suggestions. They may also (2) try out the feature for regenerating the response in order to compare it to the previous one. If the feedback still differs from what they expect, they could (3) rephrase the prompt in a clearer and more specific way (Kohnke et al., 2023, p. 8). Indeed, ChatGPT often simply needs sufficient information to produce responses that are tailored to the user’s needs. To avoid repeating the same background information in every single prompt, users can benefit from (4) the feature called “custom instructions” that can be found in the user settings since summer 2023 (OpenAI announcement July 20, 2023, at <https://openai.com/blog/custom-instructions-for-chatgpt>; cf. Gharani, 2023). Two text fields will appear in which users can characterize their own background (“What would you like ChatGPT to know about you to provide better responses?”) and expectations (“How would you like ChatGPT to respond?”). To exemplify, they could write that they are university students and expect responses to be written in a succinct academic style. They might also specify whether they prefer feedback written in a direct style which focuses on constructive critique rather than praise. In that respect, they could even describe a persona (e.g. historical figure) whose rhetorical style should be imitated in the chatbot replies (Kostka & Toncelli, 2023, p. 2).

Users can experiment with the level of detail they provide in these fields to evaluate their impact on the responses that ChatGPT provides. Otherwise, it will produce generic responses that follow a standard pattern which might not comply with the user’s needs. If users are not sure what their own (or expected) writing style is, they could also feed in good examples of their desired output (e.g. previous texts they have produced) and ask ChatGPT to define it and use it in its answers (Gharani, 2023).

If this customization does not help, unsatisfactory replies might evidence its true limitations. Certainly, the AI-generated feedback is not perfect (Dai et al., 2023b; Fryer et al., 2017, p. 463), which is why learners should utilize their own critical judgment to assess the quality of the recommendations (Haristiani, 2019). For example, Li et al. (2023) underscore the importance of comparing the ChatGPT corrections with users’ own drafts and of giving reasons for why they will (not) use certain suggestions (p. 45). Users also should not hesitate to incorporate their own ideas and perspectives (Cotton et al., 2023, p. 7) that go beyond or could even contradict the generated feedback. To verify the AI suggestions and back up their own ideas, they should consult reliable external sources (e.g. online dictionaries, grammar books, literature databases). To some extent, they can also ask the AI bot to provide further explanations (Li et al., 2023, pp. 45–46). Quite of-

ten, however, ChatGPT even reminds the users of its own limitations. For instance, when asked to explain in how far ChatGPT might be useful for self-assessment (on February 07, 2023), ChatGPT replied that

“[...] it’s important to note that the accuracy and effectiveness of language models for self-assessment will depend on the quality and comprehensiveness of the training data, as well as the specific task and context. Further research is needed to fully understand the capabilities and limitations of language models in this area.” (OpenAI, 2023c)

Hence, it is essential to keep in mind that AI-generated feedback should not be a substitute for the instructor’s expertise and guidance (Mageira et al., 2022). Rather, it should be employed in conjunction with instructor assessment to achieve a better learning gain. However, to some extent, the provided feedback can be beneficial, for example for minor revisions as well as for developing an action plan. Often, ChatGPT suggests several strategies for what learners could do in order to improve further. By checking the suggestions and consulting additional resources, they can fine-tune the action plan and implement it step-by-step as well as discuss it with their teacher if needed.

#### 5.4. Combination of ChatGPT with Peer Feedback and Teacher Guidance

As indicated above, chatbot feedback should be used in combination with other information sources, including the reading of relevant literature as well as human feedback from peers or teachers (see e.g. Escalante et al., 2023, p. 13; Li et al., 2023, pp. 45–48; cf. the review by Wahyudi, 2023, p. 2). Several scholars therefore argue for “a mixed approach” (Escalante et al., 2023, p. 13) that “strik[es] a balance between AI-assisted innovation and human expertise” (Ray, 2023, p. 121).

Overall, there are many possible ways in which technology-generated and human feedback can be combined in the classroom. For instance, students could engage in peer feedback activities first and then compare or enrich their assessment with the help of AI. Through such comparisons, they might become aware of the benefits and limitations of chatbots. To encourage (meta-)cognitive engagement and critical thinking, the utilization of a “feedback tracker sheet” (in analogy to Gießler’s [2024] change tracker sheet for AWE feedback) may prove valuable. Therein, learners can collect the feedback comments they have received from both sources (peers and chatbots), note down the explanations they have obtained (if available), indicate whether or in how far they will consider them in their further learning or in their draft revisions as well as provide reasons for (not) using them. At this point, they can engage in follow-up dialogues with the chatbot (Li et al., 2023, p. 45) and the peers or the teacher to clarify or deepen their understanding. This way, students might become aware of the chatbot’s limitations and the importance of “double-checking the correctness and accuracy of provided resources and citations” as well as of “approach[ing] them [the given responses] with a critical reading mindset” (students cited in Abdelhalim, 2024, p. 9). Accordingly, teachers should provide students with sufficient “opportunities [...] to discuss AI-generated feedback [...] and ask follow up questions” (Escalante et al., 2023, p. 13).

In fact, pedagogical guidance is crucial at any stage of the AI-supported feedback process. Teachers should therefore offer manifold scaffolded activities to foster critical thinking (Abdelhalim, 2024, p. 10) but also to manage emotions and develop interpersonal skills in human interactions. Not only should the utilization of chatbot feedback be addressed in the classroom, but their potential to develop students' feedback literacies can be profitable in additional ways, as will be discussed below.

## 5.5. Utilization of ChatGPT for Feedback Training

Apart from prompting ChatGPT to correct sentences or provide feedback on a draft, the capacities of AI bots can be exploited for feedback training and the development of learners' and (pre-service) teachers' feedback literacy. As delineated above (in section 2), feedback literacy involves the ability to request feedback, provide feedback and utilize the received feedback as well as to engage in follow-up feedback dialogues (Carless & Boud, 2018, pp. 1316–1317). While the reception and use of feedback has been explored above, the initiation, provision and follow-up conversation will be focused on in the present section.

Here is an example of a feedback conversation that was generated by ChatGPT (OpenAI, 2023a, August 04, 2023, 10:15 AM CEST):

**Prompt:** Can you give me an example of a feedback conversation in the context of learning English as a foreign language? The example should demonstrate the active role that learners play throughout the feedback process, starting with a request for feedback.

**ChatGPT:** Of course, I'd be happy to provide an example feedback conversation in the context of learning English as a foreign language:

**Learner (Maria):** Hi there, I hope you're doing well. I've been working hard on improving my English speaking skills, and I recently recorded myself having a conversation about a random topic. Would you be willing to listen to it and give me some feedback?

**Peer (John):** Hi Maria! I'd be glad to help. That's a great initiative! You can send me the recording, and I'll listen to it carefully.

**Learner (Maria):** Thank you so much, John! I really appreciate it. Here's the recording: [attaches recording]. I'm particularly focusing on my pronunciation and fluency.

**Peer (John):** I've listened to your recording, Maria. Overall, you're doing a fantastic job! Your pronunciation has improved a lot since the last time I heard you. Your speech is clear and easy to understand. One thing I noticed is that your fluency is quite good, but there were a few moments where you hesitated or repeated some words. To work on this, you might want to practice more free-flowing conversations or try some tongue twisters to improve your speech rhythm.

**Learner (Maria):** Thank you, John! I'm glad to hear that my pronunciation is better. I'll definitely take your advice to heart regarding fluency. Do you think you could point out those specific parts where I hesitated or repeated? That would be really helpful.

**Peer (John):** Of course, Maria! Let me give you some timestamps and examples. Around the 2-minute mark, you hesitated slightly when transitioning to a new idea. Instead of saying "um," you could have used filler phrases like "well," or "you know." Also, at around 4 minutes, you repeated the word "interesting" a couple of times.

Maybe you can explore synonyms like “fascinating” or “captivating” to add variety to your speech.

**Learner (Maria):** Thank you for the detailed feedback, John! I'll go back to those parts and practice more. Your suggestions make a lot of sense, and I'm excited to work on these areas.

**Peer (John):** You're welcome, Maria! I'm glad I could help. Remember, learning a new language takes time and practice, so don't be too hard on yourself. Keep up the good work, and I'm here anytime you need more feedback or just want to practice a conversation.

**Learner (Maria):** I really appreciate your support, John. Having a peer like you to give me feedback and encourage me is invaluable. I'll definitely reach out again soon. Thanks once again!

**Peer (John):** No problem at all, Maria! I'm here to help. Looking forward to our next feedback session. Keep up the great work, and take care!

Directly below this example, ChatGPT gives the following explanations that illuminate central characteristics of feedback dialogues:

“In this example, Maria takes an active role in seeking feedback, provides context about her learning goals, and engages in a meaningful dialogue with her peer, John. The conversation demonstrates a constructive feedback process where Maria's strengths are highlighted, areas for improvement are pointed out, and specific suggestions are given. Maria actively seeks clarification on feedback and expresses gratitude, creating a positive and supportive learning environment.” (OpenAI, 2023a)

The output was generated within seconds, together with the explanation of what is going on. It illustrates a potential peer feedback exchange between learners of English.

In addition, it is possible to “adjust the complexity of the dialogue to make it more suitable for beginners or advanced learners” (Kohnke et al., 2023, p. 4) or for specific contexts. Here is another example of a sample feedback exchange prompted in ChatGPT 3.5 on August 04, 2023, 10:22 AM CEST (OpenAI, 2023a):

**Prompt:** Thank you. Can you give me an example of a written feedback exchange between a learner of English as a foreign language and her teacher? The example should come from a university context. The learner intends to become a teacher of English herself. She is at advanced level already and has composed a research proposal for examining students' perceptions of different accents in English.

**ChatGPT:** Certainly! Here's an example of a written feedback exchange between a university student (Learner) and her English as a foreign language teacher (Teacher) in the context of a research proposal:

**Subject: Request for Feedback on Research Proposal**

**Learner (Maria):** Dear Professor [Teacher's Last Name],

I hope this message finds you well. I have been working on my research proposal and I would greatly appreciate your feedback before I finalize and submit it. As an aspiring English teacher, I am particularly interested in examining students' perceptions of

different accents in English and how it might influence their language learning experience.

I have attached the draft of my research proposal to this email. I am particularly concerned about the clarity of my research questions and the overall structure of the proposal. Any guidance you can provide would be immensely valuable to me. Thank you in advance for your time and support.

Best regards, Maria

**Teacher (Professor Smith):** Dear Maria,

I hope you're doing well. I'm glad to see your enthusiasm for your research proposal. It's an interesting and relevant topic, especially considering your future goals. I've reviewed your draft, and I'm impressed with the depth of your understanding and your well-defined research questions. Your passion for becoming an English teacher shines through in your writing.

In terms of structure, your introduction sets the stage effectively by highlighting the importance of accents in language learning. However, I'd suggest adding a brief literature review to provide context for your research questions and demonstrate your awareness of existing studies in this area.

Your methodology section is well thought out, and your choice of qualitative interviews to gather students' perceptions is appropriate. Make sure to detail the process of participant selection and ethical considerations more comprehensively.

Regarding your research questions, they are clear and focused. To enhance them further, consider embedding the rationale for each question within its description. This will help your readers understand the significance of each question in relation to your overall study.

Overall, you're on the right track. Make the suggested revisions, and I believe your proposal will be even stronger. Feel free to reach out if you have any questions or need further assistance. Keep up the good work!

Best regards, Professor Smith

**Learner (Maria):** Dear Professor Smith,

Thank you so much for your prompt and thorough feedback. I truly appreciate your insights and guidance. I will certainly revise the proposal to incorporate the changes you've suggested. Your idea about including a literature review makes a lot of sense, and I will make sure to provide a more detailed explanation of the participant selection process and ethical considerations.

I'm grateful for your encouragement and support as I work towards refining my research proposal. I will keep you updated on my progress, and I look forward to your continued mentorship.

Best regards, Maria

In this written feedback exchange, the learner, Maria, seeks guidance and feedback from her teacher, Professor Smith, on her research proposal. The teacher provides detailed feedback, highlighting both the strengths of the proposal and areas for improvement. The exchange demonstrates a respectful and collaborative learning environment in a university context, where the learner actively engages with the feedback and expresses gratitude for the teacher's guidance.

The second example represents a teacher-student feedback dialogue and could prove useful for the training of teachers' and learners' feedback literacy alike. Based on these

and further feedback samples, pre-service teachers might be asked to identify central characteristics of feedback exchanges and suggest alternative variants. However, they also need to become aware of the potential cultural bias that is inherent in these examples: The professor and the feedback-giving peer are male and have stereotypical English (British or US) names (Professor Smith and John, respectively). Moreover, a repeated use of the same standard formulations and structures may also leave a negative impact on the students, once they recognize that the feedback comes from a template and is not specific enough. Likewise, students need to be taught that an over-reliance on chatbot-generated feedback requests and replies might diminish their own critical thinking skills and could be identified by their teachers. Similarly, chatbot-generated writing can be detected to some extent by teachers or specialized software. To pre-empt such cheating and over-reliance on AI programs, educators can resort to the strategies that are sketched below.

## 5.6. Further Advice for Teachers

Beyond the precautionary measures mentioned above, there are some further techniques that teachers could take in order to detect an illegitimate use of chatbots. As the generated output often adopts similar phrases and structures (Cotton et al., 2023, p. 5), teachers might notice learners' use of AI, especially when it contrasts with students' previous linguistic performance. This becomes even more evident when there are obvious differences in language use within the same assignment (Cotton et al., 2023, p. 4). Moreover, the non-citation of sources or the types of sources cited provide important clues. So far, several chatbots are not yet able to reliably incorporate references (Cotton et al., 2023, p. 4). Even more so, the range of available sources is limited and might not comply with the literature list that was used in class or that was recommended by the instructor. As Cotton et al. (2023) note, "human writing tends to be more contextually aware and responsive to the needs of the audience, while writing generated by AI may be more generic and less tailored to a specific context" (p. 5). There could even be factual errors in the writing when the content was produced by a machine (Cotton et al., 2023, p. 5). While some plagiarism detection tools might be able to identify AI language to a certain extent (Cotton et al., 2023, p. 5; Qureshi, 2023, p. 7), they do not work reliably throughout. Human expertise and strategies to verify the correctness of the contents are therefore essential.

## 6. Discussion

Chatbots fulfill an increasing range of functions as learning, teaching and feedback assistants in educational contexts and beyond (cf. e.g. Essel et al., 2022, p. 2). Due to their prevailing limitations and the importance of critical and creative thinking as well as social interaction skills, AI-powered chatbots should at present merely be considered as a supplement to feedback exchanges between humans (cf. Escalante et al., 2023, p. 13; Qureshi, 2023, p. 3; Ray, 2023, p. 121). It is to be seen in how far they will become an integral part of teaching and learning in the future. To some extent, AI systems can already be a useful guide for the training of feedback literacies as well as for self-assessment, corrections, and grading.

More precisely, AI-powered engines like ChatGPT can be very helpful in generating specific feedback requests, checking grammar, providing recommendations for improvement, and giving positive, constructive and action-oriented feedback to learners. Especially the given explanations for both negative and positive feedback may help users understand the reasons behind their errors or successes, which can improve their language awareness and self-directed learning. At the same time, critical language awareness and analytical skills are needed to check the generated suggestions, formulations and structures. In that respect, the consultation of additional resources for learning and teaching is indispensable, particularly since several studies have observed a high amount of false information and fictional references.

Through an advanced training of chatbots, though, it might become possible to produce real-time, personalized, context-specific and natural-sounding responses more reliably. This can also allow “educators [to] provide personalized feedback of consistently high quality for a larger scale of class in less time” (Dai et al., 2023b, p. 5). Likewise, ChatGPT’s ability to provide personalized, immediate, non-judgmental, and multilingual feedback can make it a powerful asset for language learners seeking to improve their communication skills. Their motivation and engagement in enhancing their skills continuously can increase further if the use of chatbots occurs as guided practice in teaching settings. Through pedagogical integration, it will also become easier to cater for individual needs and learning preferences, foster active learning, decision-making, and problem-solving skills. The combination of personalized assistance and feedback by humans and AI bots can eventually contribute to improved learning achievement and self-efficacy (Essel et al., 2022, pp. 4, 12) and boost students’ confidence in their language skills (cf. Chang et al., 2022; Fryer et al., 2020, p. 16, referring to Fryer & Carpenter, 2006).

With advancements in AI, it might even become easier for educational institutions to create their own chatbots for specific needs (Essel et al., 2022, p. 2), as some universities have done already. Moreover, there are numerous apps and extensions that draw on AI. The language learning app Duolingo, for instance, makes use of GPT-4 (Frackiewicz, 2023). It is likely that technological improvements will further enhance feedback processes in the future by moving beyond simple corrective feedback. One critical and highly debated area for future research is the emotional sensitivity of AI-generated feedback, including non-verbal ways of emotional reactions (cf. Fryer et al., 2020, p. 15). On the one hand, this could have affective benefits; on the other hand, it would blur the boundaries between humans and technologies even further, making them less distinguishable from each other. This could be perceived as dangerous, especially with respect to the development of interpersonal skills.

Other attempts at improvement relate to the incorporation of sources to the bot output. For example, the chatbot Perplexity.ai (<https://www.perplexity.ai/>) lists sources together with the bot answers, even when not explicitly asked for. Several of them indeed lead to existing websites, whereas some others seem non-existent or at least inaccessible. However, these sources are not embedded within the running text, but given in a separate section (i.e. in-text citations are missing). Aside from that, several extensions are being developed, including voice communication (e.g. Talk to ChatGPT) and video sum-

maries (e.g. YouTube Summary with ChatGPT), which could enhance the (multimodal) feedback process further.

Overall, it is important for users to continuously examine these developments critically in order to exploit their potentials effectively while being aware of their limitations. The aim of the present chapter was to make a contribution to this highly dynamic and promising field. Naturally, some of the limitations might have become obsolete in the years following this publication, which is in line with the idea of dynamic feedback literacy development (Schluer, 2022, p. 238). Both teachers and learners need manifold and agile skills for using and critically inspecting technological developments (Schluer, 2022, pp. 238–249).

Generally, these developments cannot be fully ignored to prepare students for real-life workplace contexts which progressively draw on new technologies for an increasing number of tasks (Schmied, 2023, pp. 12–13). In sum, then, guided practice within educational contexts would be a recommended. Educators should offer an open-minded but critical stance (Schmied & Ivanova, 2023, p. vi) whilst exploring and discussing the affordances and dangers of AI, as well as use a meaningful combination of manual and automated feedback methods (Cotton et al., 2023, p. 3).

## 7. Conclusion

This chapter has reviewed several challenges and affordances of chatbots with an emphasis on enhancing feedback processes in educational settings. This includes using chatbots as facilitators of self-feedback among learners, of feedback processes by teachers, and as inspirations for initiating and continuing feedback conversations in the classroom.

With chatbot systems undergoing constant transformation, their impact on the learning and teaching experience needs to be inspected continuously and critically. Together with their students, teachers should openly discuss these developments to avoid an illegitimate use of the AI-powered tools (see also closing chapter 17 by Schluer in this volume). Moreover, it is crucial to combine chatbot-generated feedback with other feedback methods, such as teacher feedback or peer feedback in various written, oral and multimodal ways in face-to-face, online and hybrid learning constellations. To choose an adequate feedback method, the primary impetus should always be the intended learning objective (as reviewed by Schluer, 2022, pp. 53–54). For each feedback mode and tool, then, specific training will be required for teachers and learners alike. In that respect, there is still lots of underexplored territory from an empirical and pedagogical perspective. It is hoped that the present contribution will give inspirations for future research and teaching practice concerning a pedagogically sound use of chatbot feedback.

## References

- Abbasi, S., Kazi, H., & Hussaini, N. N. (2019). Effect of chatbot systems on student's learning outcomes. *SYLWAN*, 163(10), 49–63.

- Abdelhalim, S. M. (2024). 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*, 1–15. <https://doi.org/10.1111/jcal.12948>
- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, Article 100006, 1–18. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Aithal, V., & Silver, J. (2023). *Enhancing learners' critical thinking skills with AI-assisted technology*. Cambridge. <https://www.cambridge.org/elt/blog/2023/03/30/enhancing-learners-critical-thinking-skills-with-ai-assisted-technology/>
- Alshahrani, A. (2023). The impact of ChatGPT on blended learning: Current trends and future research directions. *International Journal of Data and Network Science*, 7(4), 2029–2040. <https://doi.org/10.5267/j.ijdns.2023.6.010>
- Amin, M. Y. M. (2023). AI and Chat GPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, 4(4), 1–15. <https://doi.org/10.33422/ijhep.v4i4.554>
- Boud, D., & Dawson, P. (2023). What feedback literate teachers do: An empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, 48(2), 158–171. <https://doi.org/10.1080/02602938.2021.1910928>
- Brinegar, M. (2023). Chatbots as a supplementary language learning tool: Advantages, concerns, and implementation. *International Journal of Education and Social*, 6(6), Article 875, 223–230. <https://doi.org/10.37500/IJESSR.2023.6615>
- Carless, D. (2019). Feedback loops and the longer-term: Towards feedback spirals. *Assessment & Evaluation in Higher Education*, 44(5), 705–714. <https://doi.org/10.1080/02602938.2018.1531108>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Chang, C.-Y., Kuo, S.-Y., & Hwang, G.-H. (2022). Chatbot-facilitated nursing education. *Educational Technology & Society*, 25(1), 15–27.
- Chen, C.-F. E., & Cheng, W.-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112. <http://lt.msu.edu/vol12num2/chencheng/>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 33, 1–12. <https://doi.org/10.1080/14703297.2023.2190148>
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023a). Can large language models provide feedback to students? A case study on ChatGPT. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 323–325). IEEE. <https://doi.org/10.1109/ICALT58122.2023.00100>
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023b). *Can large language models provide feedback to students? A case study on ChatGPT*. Preprint. EdArXiv. <https://osf.io/hcgzj> <https://doi.org/10.35542/osf.io/hcgzj>

- Dembsey, J. M. (2017). Closing the Grammarly® gaps: A study of claims and feedback from an online grammar program. *The Writing Center Journal*, 36(1), 63-96, 98-100. <https://www.jstor.org/stable/44252638>
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1), Article 57, 1-20. <https://doi.org/10.1186/s41239-023-00425-2>
- Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19, Article 57, 1-19. <https://doi.org/10.1186/s41239-022-00362-6>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1-15. <https://doi.org/10.1080/14703297.2023.2195846>
- Frackiewicz, M. (2023). *The Chat GPT-4 approach to language learning: Teaching and learning with AI*. <https://ts2.space/en/the-chat-gpt-4-approach-to-language-learning-teaching-and-learning-with-ai/>
- Fryer, L., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior*, 75, 461-468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning*, 24(2), 8-22.
- Gharani, L. (2023). *Don't use ChatGPT until you watch this video*. <https://www.youtube.com/watch?v=wBANcMA98ls>
- Gießler, R. (2024). EFL writers' cognitive engagement with AWE feedback. *Language Awareness*, 33(2), 428-445. <https://doi.org/10.1080/09658416.2023.2269088>
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), Article 692, 1-13. <https://doi.org/10.3390/educsci13070692>
- Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *Journal of Physics: Conference Series*, 1387, 1-6. <https://doi.org/10.1088/1742-6596/1387/1/012020>
- Hoang, N. T., Ngoc Han, D., & Le, D. H. (2023). Exploring Chatbot AI in improving vocational students' English pronunciation. *AsiaCall Online Journal*, 14(2), 140-155. <https://doi.org/10.54855/acoj.231429>
- Javier, D. R. C., & Moorhouse, B. L. (2023). Developing secondary school English language learners' productive and critical use of ChatGPT. *TESOL Journal*, Article e755, 1-9. <https://doi.org/10.1002/tesj.755>
- Kohnke, L., Moorhouse, B. L., & Di Zou (2023). ChatGPT for language teaching and learning. *RELC Journal*, 10(3), 1-14. <https://doi.org/10.1177/00336882231162868>
- Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *Teaching English as a Second or Foreign Language Journal (TESL-EJ)*, 27(3), 1-19. <https://doi.org/10.55593/ej.27107int>

- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lechler, R., Stoeckli, E., Rietsche, R., & Uebernickel, F. (2019). Looking beneath the tip of the iceberg: The two-sided nature of chatbots and their roles for digital feedback exchange. In Association for Information Systems (Ed.), *Proceedings of the 27th European Conference on Information Systems (ECIS)* (pp. 1–17). AIS Electronic Library (AISeL).
- Li, J., Ren, X., Jiang, X., & Chen, C.-H. (2023). Exploring the use of ChatGPT in Chinese language classrooms. *International Journal of Chinese Language Teaching*, 36(55), 35–55. <https://doi.org/10.46451/ijclt.20230303>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(Article 410), 1–15. <https://doi.org/10.3390/educsci13040410>
- Magiera, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. *Applied Sciences*, 12, 1–16. <https://doi.org/10.3390/app12073239>
- OpenAI (2023a). *ChatGPT [Large language model]: Active Feedback Conversation*. Retrieved August 04, 2023, from <https://chat.openai.com/chat>
- OpenAI (2023b). *ChatGPT [Large language model]: AI Feedback for Writing*. Retrieved March 18, 2023, from <https://chat.openai.com/chat>
- OpenAI (2023c). *ChatGPT [Large language model]: ChatGPT Education Research*. Retrieved February 07, 2023, from <https://chat.openai.com/chat>
- Qureshi, B. (2023). Exploring the use of ChatGPT as a tool for learning and assessment in undergraduate computer science curriculum: Opportunities and challenges. [https://www.researchgate.net/publication/370227070\\_Exploring\\_the\\_Use\\_of\\_ChatGPT\\_as\\_a\\_Tool\\_for\\_Learning\\_and\\_Assessment\\_in\\_Undergraduate\\_Computer\\_Science\\_Curriculum\\_Opportunities\\_and\\_Challenges](https://www.researchgate.net/publication/370227070_Exploring_the_Use_of_ChatGPT_as_a_Tool_for_Learning_and_Assessment_in_Undergraduate_Computer_Science_Curriculum_Opportunities_and_Challenges)
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Ruan, S., Jiang, L., Xu, Q., Liu, Z., Davis, G. M., Brunskill, E., & Landay, J. A. (2021). Englishbot: An AI-powered conversational system for second language learning. In *26th International Conference on Intelligent User Interfaces* (pp. 434–444). ACM. <https://doi.org/10.1145/3397481.3450648>
- Schluer, J. (2020). Individual learner support in digital ELT courses: Insights from teacher education. *International Journal of TESOL Studies*, 2(3), 41–63. <https://doi.org/10.46451/ijts.2020.09.17> (Special Issue: ELT in the Time of the Coronavirus 2020 (Part 2)).
- Schluer, J. (2022). *Digital feedback methods*. Narr Francke Attempto.
- Schluer, J. (2023). *Digital feedback map: Overview of digital feedback methods*. <https://tinyurl.com/DigitalFeedbackOverview/>
- Schmied, J. (2023). Confidence and trust in online academic discourse: Integrating new technologies into teaching and learning. In J. Schmied & M. Ivanova (Eds.), *Comparing confidence and trust online and offline* (1st ed., pp. 1–14). *Research in English and Applied Linguistics (REAL Studies): Vol. 19*. Cuvillier.

- Schmied, J., & Ivanova, M. (2023). Preface. In J. Schmied & M. Ivanova (Eds.), *Comparing confidence and trust online and offline* (1st ed., pp. v–viii). *Research in English and Applied Linguistics (REAL Studies): Vol. 19*. Cuvillier.
- Wahyudi, Z. A. (2023). The role of chatbots in English language education: A comprehensive review. *Pustakailmu.Id*, 3(6), 1–19.
- Winstone, N. E., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Society for Research into Higher Education (SRHE). Routledge.
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 1–18. <https://doi.org/10.3389/frai.2021.654924>
- Woo, D. J. (2023). *Evaluating a learning design for EFL writing using ChatGPT* [Preprint]. <https://www.researchgate.net/publication/373265937>
- Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2023). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process [Preprint]. <https://www.researchgate.net/publication/376813082>
- Xie, X., Nimehchisalem, V., & Rafik-Galea, S. (2022). Paradigm shifts in peer feedback within learning-oriented language assessment. *ASEAN Journal of Applied Linguistics*, 1(1), 20–34.
- Yin, J., Goh, T.-T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, 59(1), 154–177. <https://doi.org/10.1177/0735633120952067>
- Zhang, R., Di Zou, & Cheng, G. (2023). A review of chatbot-assisted learning: Pedagogical approaches, implementations, factors leading to effectiveness, theories, and future directions. *Interactive Learning Environments*, 110(9), 1–29. <https://doi.org/10.1080/10494820.2023.2202704>

## Acknowledgments

Special thanks go to Shanqing Gao and Erhan Altay for assisting me in the literature searches for this chapter. This support was made possible by the *Stiftung Innovation in der Hochschullehre* as part of the project “Didaktische Orientierung für digitales Feedback (Pedagogical guidance for using digital feedback): Digital Feedback Map (DFM)”, funded from 09/2022 to 11/2023 (funding ID: FRFMM-181/2022).

