International Journal of Educational Technology in Higher Education

RESEARCH

Open Access

Designing a mobile chatbot-based learning journaling system for intrinsic motivation and engagement

Sven Scheu[1*], Simone N. Loeffler[2] and Alexander Maedche[1]

*Correspondence:
Sven Scheu
sven.scheu@kit.edu
[1]Human-Centered Systems Lab, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]Learning Lab, House of Competence, Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract
Journaling enables students to reflect on their learning processes and thereby strengthen their self-regulation, a key competency for meeting academic goals. Previous work has shown that students benefit from digital support when creating learning journals, such as through mobile learning journaling systems. Yet, a major issue with such systems is the rapid decline in user motivation and engagement, often occurring after only a brief period of use. To address this challenge, we present a mobile chatbot-based learning journaling system that guides students through structured reflective journal entries and supports writing through an LLM-based journaling assistant. The novel system combines (1) an example-based built-in course that teaches reflective journaling through modeled responses with (2) an interactive journaling assistant that scaffolds students' entries by generating follow-up questions and rewrite suggestions. In a randomized field experiment with 179 students using the system over 22 days, we examined the impact of both design principles on intrinsic motivation and behavioral engagement. While our results indicate that the built-in course can increase intrinsic motivation, we find no evidence that the LLM-based journaling assistant improves intrinsic motivation. Regarding engagement, both design principles provide benefits in different ways: the course shows a rather constant positive influence, whereas the LLM-based assistant appears to form a feedback loop with continued use, increasing engagement over time.

Keywords Self-regulated learning, Journaling, Chatbot, Motivation, Engagement

Introduction

Learning journals are a promising approach for engaging students in critical reflective activities regarding self-regulated learning (SRL) and increasing the awareness of their learning processes (Hiemstra, 2001; Hubbs & Brand, 2005; Luft et al., 2026; Nückles et al., 2020). SRL describes learners as active participants in their learning processes that can shape and develop their cognitive and behavioral actions in a successful way (Boekaerts, 1999; Efklides, 2011; Schunk & Greene, 2018). Proficient self-regulated learners employ cognitive strategies to improve their success in learning (Zimmerman & Pons, 1986) and utilize metacognition to refine their learning processes continuously

(Schunk & Greene, 2018). In practice, students often struggle with self-regulating their learning, and only a fraction of learners are competent self-regulators (Barnard-Brak et al., 2010; Ning & Downing, 2014; Valenzuela et al., 2020). Consequently, many students fail to convert their intellectual capacities into academic achievement (Bjork et al., 2013; Boekaerts & Cascallar, 2006).

Self-reflection is regarded as a prerequisite for students to effectively evaluate their strengths and weaknesses and adapt their learning behavior accordingly. During SRL, self-monitoring and self-reflection processes are crucial for the selection of appropriate SRL strategies; They depend on individual characteristics, the subject manner, the current individual situation, and external circumstances (Broadbent et al., 2020; Schmitz & Wiese, 2006). Learning journals are considered useful tools for engaging in SRL, foster self-reflective activities, and allow the writers to increase metacognitive activities in their learning processes (Wallin & Adawi, 2018). However, to take full advantage of these benefits, students need guidance accompanying the creation of their learning journals (Dörrenbächer & Perels, 2016; Fabriz et al., 2014; Lu & Wang, 2022; Pesonen et al., 2020). Previous work has shown that students benefit from digital support when creating learning journals, such as through mobile learning journaling systems (e.g., Nepal et al., 2024). Yet, a major issue with such systems is the rapid decline in user motivation and engagement, often occurring after only a brief period of use (e.g., Wong et al., 2026).

In this work, we explore how mobile learning journaling systems should be designed to better assist students in creating learning journals, especially by strengthening their intrinsic motivation and maintaining engagement over time. Prior studies have explored the potential of structured learning diaries for prompting the application of SRL strategies revealing that this approach alone cannot replace SRL trainings (see, e.g., Ewijk et al. 2015; Fabriz et al., 2014; Perels et al. 2007; Schmitz & Wiese, 2006). We build on the ideas of scaffolding and structured learning diaries by designing a mobile chatbot-based journaling system that guides students through their journaling process. We argue that chatbots are a natural advancement of the previous structured approaches because of their adaptability, their role as a form of social support, and their ability to guide users through their individual reflection processes (Brandtzæg et al., 2021; Feine et al., 2019; Kocielnik et al., 2018; Lee et al., 2021).

A major advantage of mobile learning journaling systems lies in their seamless integration in daily life and their scalability. However, a well-known critical issue regarding their voluntary use is the fading of engagement and subsequent dropout of the participants after a few usage days (Baumeister & Vohs, 2007; Baumel et al., 2019). Thus, to address this challenge, we articulated and implemented two design principles: (a) An example-based built-in course and (b) a journaling assistant leveraging large language models (LLMs) to support students in maintaining their intrinsic motivation and engagement to keep their learning journals. In the present study, we evaluate the impact of the proposed design principles and their instantiation on subjective and behavior-related motivational constructs aiming to answer the following superordinate research question: *What is the effect of the proposed design principles implemented in a mobile chatbot-based learning journaling system on students' intrinsic motivation and engagement in keeping a learning journal?.*

To investigate this question, we implement a fully functional system and perform a randomized field experiment to explore the effects of the two proposed design principles on intrinsic motivation and engagement as well as on SRL.

Background and related work

Learning journals

Journal writing emerged as a promising method to support self-reflection processes, identifying behavioral patterns, cognitive organization, and processing of experiences, as well as emotion regulation and stress management (Alt & Raichel, 2020; Greenleaf Brown et al., 2022; Nepal et al., 2024; Nückles et al., 2020; Smyth et al., 2018). A learning journal is a tool for students to record their "thoughts, reflections, feelings, personal opinions, and even hopes or fears during an educational experience" (Hiemstra, 2001, p. 20) and by doing so, the learning journal helps students to reflect on their learning (Hubbs & Brand, 2005; Miller, 2017). In the long run, journaling supports students in becoming more aware of their learning practices (Broadbent et al., 2020) and enables them to identify their learning gaps and comprehension difficulties (Luft et al., 2026; Nückles et al., 2020). However, it should be noted that a learning journal is not a simple recollection of past events like a *learning diary*. Instead, the defining element of *learning journals* is that they include reflection as part of the creation process (Fabriz et al., 2014; Park, 2003). In addition, the creation of a learning journal is not a single reflective action but the cumulation of many reflective entries about the entire learning process of the writer over multiple learning sessions (Moon, 2019; Nückles et al., 2009). The journal's creation can help students become more cognitively aware of their actions during learning (Lindroth, 2015) by supporting the learners to connect their "thoughts, feelings, and actions" (Hubbs & Brand, 2005, p. 62) retrospectively, without being pressured to act on practical issues immediately (Morrison, 1996).

Journaling methods have been used frequently and successfully with nursing students to promote reflection, clinical judgment, and emotional competence (Greenleaf Brown et al., 2022). Moreover, in recent years, reflective journaling has been applied in combination with wearable devices to record physiological parameters, in some cases also with integrated LLMs, aiming to promote mental and physical health (Nepal et al., 2024; Ren et al., 2025; Zhao et al., 2026). However, there have been few new developments regarding reflective writing journals promoting SRL in higher education that have been systematically evaluated in adequate longitudinal studies during SRL processes in daily learning routines. An exception is the Freiburg Self-Regulated Journal Writing Approach; in that context, various instructional methods to support SRL by optimizing cognitive load through journal writing were developed and tested in the laboratory and in the field (Nückles et al., 2020).

Like any learning strategy, the skill to write a learning journal is not inherent to students but requires training and feedback to develop (Bain et al., 2002; Hume, 2009). For example, students who are unfamiliar with the creation of a learning journal might be unsure about what to write in their learning journal (Dincel & Savur, 2019; Jarvis & Baloyi, 2020; Kasprabowo et al., 2021).

Previous studies have focused on a spectrum of strategies to support the creation of learning journals. The most open strategy is allowing users to create their learning journal freely and provide a complementary SRL training (see, e.g., Broadbent et al., 2020;

Dörrenbächer & Perels, 2016). A more structured approach is to utilize explicit prompts that serve as a foundation for a journal entry (see, e.g., Berthold et al. 2007; Hübner et al. 2010; Nückles et al. 2009). The most restricted approach is the usage of standardized diaries based on questionnaires on SRL, in which users self-assess various aspects of their SRL mostly by providing ratings on Likert scales (see, e.g., Ewijk et al., 2015; Fabriz et al.,2014; Perels et al., 2007; Schmitz & Wiese, 2006).

In their review article based on 16 experimental and 4 correlative studies, Nückles et al. (2020) show that from the perspective of cognitive load theory, journal writing is promising, since it can serve as an adequate scaffold for promoting SRL by offering the following advantages: Writing gives learners the opportunity to externalize their own thoughts, reread them, and develop them further, with the written text acting as memory aid or feedback. Externalizing thoughts reduces the cognitive processing load, allowing more cognitive activity to be devoted to germane processing, such as metacognitive reflection (see also Luft et al., 2026). However, to benefit adequately from these advantages, journaling needs to be supported by instructions, as unguided learners do not engage sufficiently in germane processes, tending to keep the mental effort during journaling to a minimum. In the SRL context, prompting proved to be the most important support method; the best learning outcomes were achieved with prompting of all main SRL sub-components (Nückles et al., 2020).

In our view, chatbots are the natural extension of prompt-based journaling scaffolds and standardized learning diaries because they can provide assessments and serve as an interactive platform for prompts (Schick et al., 2022). These benefits have long been used in mental health to provide chatbots that supply users with assessment, treatment, and journaling capabilities (Abd-Alrazaq et al., 2020; Kawasaki et al., 2020). For example, Kawasaki et al. (2020) and Lee et al. (2021) built a chatbot enabling users to create their journals by chatting with a chatbot and showed that users react similarly to the guidance of a chatbot and humans. This guidance can even be effective if it is static or mostly static and does not rely on natural language understanding to steer a dynamic conversation but presents a set of prompts (Kocielnik et al., 2018; Wolfbauer et al., 2022, 2020).

Engagement, self-determination, and intrinsic motivation

Learners' engagement is essential for the learning process and learning success. Academic task engagement has been conceptualized as a multidimensional construct involving affective, behavioral, and cognitive dimensions (Ben-Eliyahu et al., 2018). In the present work, we focus on the behavioral dimension of engagement during journal writing. Behavioral engagement refers to active involvement and persistence in task completion. In the context of writing tasks, students' behavioral engagement has been commonly operationalized as time investment and writing productivity (Bråten et al., 2022; Fleckenstein et al., 2024; Namkung & Kim, 2024). Bråten et al. (2022) confirmed significant relationships between engagement, motivation, cognition and performance showing that in student post-reading written reports behavioral engagement (writing time and length of the written responses) predicted comprehension performance, and that behavioral engagement mediated the effects of cognitive prerequisites and intrinsic reading motivation on comprehension performance.

Moreover, engagement is often a concern for mobile systems that target a change in behavior or acquisition of new behavior, like starting to journal regularly, because engagement naturally declines over time (Baumel et al., 2019). This is especially an issue for long-term engagement, or "the degree of involvement a user chooses to have with a system over time" (Bickmore et al., 2010, p. 649) (Baumel et al., 2019; Lipschitz et al., 2023).

In a recent study, Wong et al. (2026) applied learning analytics to identify user profiles from log data collected through the voluntary use of a mobile learning app offering multiple-choice questions during a nine-week undergraduate university course and to examine differences in exam grades. Disengaged users ($n = 54$) started using the app early (18.5 days before the exam), but stopped prematurely (8.5 days before the exam) and answered few questions. Utilitarian users ($n = 57$) only answered multiple-choice questions two days before the exam for immediate preparation. Active users ($n = 82$) worked the most with the questions, started the app early (17.5 days before the exam) and continued until one day before the exam. 117 students did not try the app at all (non-users). Whereas active and utilitarian users achieved significantly better exam grades than the non-users, there were no differences between the disengaged users and the non-users.

One approach to promote engagement is to motivate the user in some form, as motivation is the fundamental driver of action (Achtziger & Gollwitzer, 2018; Deci & Ryan, 2000). Consequently low motivation may result in the remaining motivation not being able to overcome the mental barrier that is a result of the effort required to engage in a task (Baumel et al., 2019). This motivation can be either extrinsic or intrinsic to an action, and any action might have both extrinsic and intrinsic motivational components (Ryan & Deci, 2000). Examples of extrinsic motivated action include tasks that are coupled to a reward like money or food, while in intrinsically motivated action the activity itself is rewarding to the individual (Ryan & Deci, 2000).

According to self-determination theory the most important drivers for intrinsic motivation are that the activity satisfies the human needs for autonomy, competence, and relatedness (Deci & Ryan, 2000). In this context, autonomy refers to the perception that the outcome of any event depends on the person's actions instead of being outside of his or her control (Deci & Ryan, 1985), whereas competence is related to how effective a person perceives oneself at performing a non-trivial task (Deci & Ryan, 1985). Relatedness is the "desire to feel connected to others–to love and care, and to be loved and cared for" (Deci & Ryan, 1985, p. 231).

Chatbots promoting SRL and motivation

In almost all previous work on promoting SRL through chatbots, the completion of tasks generated specifically for the study was supported by a few task-oriented cognitive or metacognitive strategies conveyed through very specific instructions (Guan et al., 2025). Moreover, educational chatbots rarely took motivational aspects into account. When motivational constructs were considered, motivation was usually assessed only once, typically at the beginning of the study, based on self-reports (Guan et al., 2025; Huang et al., 2025). An exception is the study of Liu and Reinders (2025) that compares two chatbot versions (a pre-scripted chatbot and a chatbot powered with ChatGPT) designed to promote SRL skills through reflection depicting the entire SRL cycle. The chatbot serving as a learning coach showed a more positive impact on motivational development. The following reasons were suggested: meeting of basic psychological needs as the generative AI-model was instructed to converse in a friendly and constructive way, novelty effect, adaptability and flexibility. Although chatbots generally have the potential to boost motivation, empirical results to date have been rare and heterogeneous (Huang et al., 2025).

In the context of higher education, the active role of students is critical, as the environment is characterized by considerable degrees of freedom and flexibility in the organization of learning processes. Consequently, there is a lack of chatbots that provide comprehensive SRL support to help students achieve their individual learning goals in their daily study routine considering dynamics and behavioral measures of motivation and engagement (cf., Guan et al., 2025; Huang et al., 2025).

AI-based support in this context should be understood as a form of scaffolding or co-regulation rather than as a replacement for student self-regulation. The system can prompt, summarize, and reformulate, but the student remains responsible for selecting goals, evaluating suggestions, and deciding what to write. This distinction is important because external support can also lead to cognitive offloading, i.e., the use of external tools or actions to reduce cognitive demand (Risko & Gilbert, 2016). Such offloading can be beneficial when it frees resources for higher-order reflection, but it can also become maladaptive when users overly rely on external support (Fröscher et al., 2022; Gilbert, 2015).

A chatbot-based learning journaling system

In order to contribute by bridging the research gap described above (cf. Guan et al., 2025; Huang et al., 2025), we designed and implemented a mobile chatbot-based system that supports students to achieve their personal learning goals by creating an individual learning journal. Therefore, we prompted active reflective processes on SRL and explicitly foster motivation and engagement. In the following, we describe the underlying design principles and hypothesize their effects. We also detail their implementation and integration into a fully functional mobile app.

Design

While existing examples of chatbots show promise using prompting to assist the creation of learning journals, current approaches appear to have potentially negative motivational effects on learning and journaling itself; e.g., a loss of motivation after a short time or time-consuming activities with the learning systems (Fabriz et al., 2014; Luft et al., 2026; Wolfbauer et al., 2022). This presents an issue, as motivation is crucial for maintaining and creating a learning journal (García & Pintrich, 1991). Consequently, a sufficient degree of motivation is inherently required to engage in any self-regulation process (Zimmerman, 2008; Zimmerman & Moylan, 2009); and self-evaluation during SRL influences the motivation to learn and engage in further self-regulation processes(Zimmerman, 2008; Zimmerman & Moylan, 2009). However, to the best of our knowledge, no previous work addresses the question of how to design a chat-based learning journaling system to increase the motivation and engagement of students.

Example-based course (design principle 1)

The first step of SRL skill acquisition is the observation and imitation of a proficient model (Schunk & Zimmerman, 1997; Zimmerman, 2000). In addition to providing practical guidance on how to implement SRL, a model like a peer or an instructor is also a primary source of motivation to continue engaging with the self-regulation cycle (Zimmerman, 2000, 2013). This is important because motivation and SRL are "reciprocally interactive" (Zimmerman, 2008), meaning a certain level of motivation is required for SRL, but successful SRL also improves motivation. Before this reciprocal relationship becomes self-sustaining, an initial external motivational source is usually required (Zimmerman, 2008). However, since a mobile chatbot-based system is inherently individualistic, external motivational resources and human models are unavailable. We argue that addressing these shortcomings requires both an alternative model and a motivational source to help users initiate the SRL process.

The role of providing guidance on how to self-regulate one's learning traditionally falls to teacher and peers that can provide instructions on how to self-regulate (Carr, 1996; Hofer et al., 1998). Following this approach, a chatbot could also provide learning material on how to self-regulate and how to write a learning journal. E.g., in our chatbot-solution, we prompted metacognitive strategies (self-reflection, self-monitoring, self-regulation, self-evaluation, and, in some cases, strategy application) regarding all SRL main constructs (motivation, requirement level, goal setting, planning, time management, situational interest, resource-management, learning emotions, cognitive learning strategies, peer-learning, and satisfaction with learning outcomes) to support the participants in reaching their personal learning goals (cf., Nückles et al., 2020; Zimmerman, 2002). But, since this does not necessarily imply the presence of a model, we decided to complement the chatbot-solution with a model-based approach borrowed from general education by making meaningful application examples available to the students and to allow them to use these examples as a model for their journaling (Atkinson et al., 2000). Combining these approaches leads us to our first design principle: include a course in the journaling system with a course structure of seven days and providing examples of how the prompts could be answered. The examples fill the role of a model providing users with a basis for imitation to satisfy the observational level of regulation. Moreover, the course offered metacognitive background knowledge about the reflective prompts and their value, which has been shown to improve motivation to engage in SRL (Schunk & Greene, 2018; Schunk & Rice, 1987) and thereby supporting perceived competence (Deci & Ryan, 2000).

Journaling assistant (design principle 2)

Learning requires students to perform a task that has some challenging aspects (Bjork, 1994). At the same time, the task should not be experienced as too challenging, since self-determination theory tells us that perceived competence would then lead students to avoid the task altogether. The appropriate level of challenge depends on the learner; while some students might find the task trivial, other students might be unsure about what to write (Dincel & Savur, 2019; Jarvis & Baloyi, 2020; Kasprabowo et al., 2021). In such cases where students feel unprepared, providing scaffolding during the creation of journal entries could help them in this initial phase, increasing their perceived competence and, in turn, engagement by providing an adequate challenge (Deci & Ryan1985; Järvelä 1995). Especially adaptive scaffolding has been shown to support students in their learning (Azevedo et al., 2005, 2004). This dynamic matching of difficulty with the current competence of the user should also help in sustaining motivation (Bandura & Schunk, 1981; Bjork & Bjork, 2011; Bjork et al., 2013). New large language model based

AI tools could provide scaffolding in a way that takes into account the current competence of the user while affording users with the autonomy to apply these tools as they see fit. However, such support must remain bounded. If the system begins to formulate, structure, or organize reflection on behalf of the student, it risks encouraging cognitive offloading rather than reflective self-regulation (Risko & Gilbert, 2016). We thus propose our second design principle: the inclusion of an AI-based assistant to support the creation of a learning journal, while encouraging reflective action by the student.

Hypotheses

We assume that providing the example-based course and the journaling assistant should increase intrinsic motivation and engagement. Building on this conceptualization and the proposed design, we articulate the following hypotheses:

H1 Providing users an example-based course leads to greater intrinsic motivation for creating a learning journal compared to users without the course.

H2 Providing users a journaling assistant leads to greater intrinsic motivation for creating a learning journal compared to users without the assistant.

H3 Providing users a example-based course leads to more engagement with the learning journal compared to users without the course.

H4 Providing a journaling assistant leads to more engagement with the learning journal compared to users without the assistant.

Implementation

The foundational application used to evaluate our hypotheses is part of the Amsl project (Scheu et al., 2023). This study forms another iteration in the overarching design science research (DSR) project (Peffers et al., 2007; Vaishnavi & Kuechler, 2008). To instantiate the proposed design principles, we reused the core frontend and backend of the application including the overall design aesthetic but removed all functionality that was not needed for the newly designed journaling process. The resulting mobile chatbot-based learning journaling system is depicted in Fig. 1.

The design of the journaling process is comparable to publicly available commercial mobile applications for journaling with the additional features instantiating the design principles described above. The home screen, depicted in the leftmost screenshot,

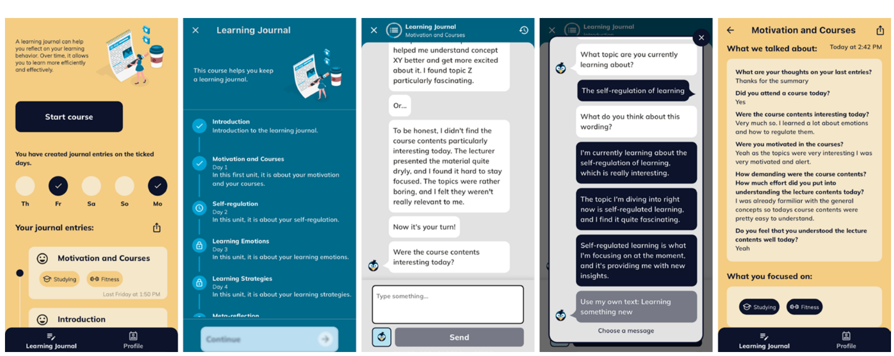

Fig. 1 Screenshots

prominently contains a button to start new journal entries (or the introduction, if it has not been completed yet), a list of past journal entries, and a quick overview of the participant's journaling activity over the last five days. From the home screen, users can view their past journal entries (cf. the right-most screenshot) or start a new journaling session based on a template. To keep the journal creation predictable and comparable, the chatbot was rule-based, with a primarily linear conversation flow following said template. All non-journaling-related conversation elements, like greetings, were handled using predefined reply buttons, while the journaling prompts allowed for full free text entry. To get the users into a reflective mindset and to promote a sense of relatedness in users with the chatbot, each journaling session started with a short greeting and a question about their current mood and what they focused on during the day. To assess mood, we used a 5-point scale from *very bad* to *excellent*. The available focuses were pre-set (e.g., *studying* or *fitness*), but users could easily add new ones in case they did not like the pre-set ones.

All application versions, including the baseline version, comprised seven topic-based journal templates, a general template, and optional daily summaries. The topic-based journaling templates were centered around specific topics related to SRL, for example, *learning strategies* or *motivation*. The general template was meant as a fallback in case the users did not want to create a journal entry for a specific topic and contained two prompts: (a) "*What are your thoughts about your day in general?*" and (b) "*What are your thoughts about your learning progress today?*".

Design instantiation

The design principles we derived were instantiated as follows.

Example-based course The example-based course was designed as a direct adaptation of the journaling templates. Each template, except for the general one, was converted into exactly one session in the course by extending it with information about why the prompt is relevant context and examples of possible responses to the prompts. For example, the mood prompt is contextualized in the course with "*Many factors influence your motivation. For example, your mood*". The example responses were designed to nudge the users to elaborate on their responses instead of providing short answers. In particular, they were formulated in complete sentences and slightly longer than expected for users to write naturally. For example, the model response for the prompt "*Were you motivated in the courses?*" was "*Yes, I was motivated to attend the courses despite the less interesting content. My motivation comes more from the desire to get good grades. Even if the topics aren't always super exciting, I try to listen actively and take notes to understand the material*". The course was also designed to teach users the daily habit of creating a journal entry. To do so, we allowed users only to do a single course session per day.

To keep the course and non-course versions as compatible as possible, the course contained the same prompts, in the same order, as the corresponding template, with the explanations and examples weaved in between the prompts. Each session in the course was time-gated to allow only a single entry per day to prevent users from rushing through the course and to encourage the habituation of journal entry creation.

Journaling assistant The assistant was implemented as a writing aid that asks follow-up questions to encourage deeper reflection. We implemented this in a way that is fully user-initiated: After the journal presented a prompt to the user and they wrote a preliminary response, they could ask the assistant to help them write a response. The assistant indicated its understanding of the preliminary response by summarizing it. Subsequently, the assistant asked a clarifying question in order to let the user elaborate on the original response. For example, if the user answered the prompt "*Were the course contents interesting today?*" with "*Yes*", the assistant could then ask "*Why did you think the course contents were interesting today?*". The assistant would then generate up to three new responses based on the original prompt, the original answer, the inquiry, and the response to the inquiry. If the response to the previous example inquiry was "*Because it was very math-heavy, and I like math*", the assistant might suggest "*The courses were very interesting today because they were very math-centric*" as one of the possible responses to the original prompt. Finally, the users could select one of the suggested responses or keep their own. In addition, the whole process could be repeated until the user was satisfied with the response, and users could always edit the final answer before submitting it.

The journaling assistant utilized OpenAI's (2023) *gpt–3.5-turbo-1106* model using zero-shot, instruction-based prompting with schema-constrained function calling (cf., Reynolds & McDonell, 2021). More specifically, the model received the current journal prompt(s), the user's unfinished response, and explicit instructions to (1) summarize the response briefly, and (2) generate a single follow-up question. No exemplar input–output pairs were included in the prompt; instead, the output format was constrained through a predefined function schema. In the second step, the model received the original prompt-answer pair and the follow-up question-answer-pair. From this it was tasked with generating two to three alternative formulations of a possible journal entry in the first person.

Other components

In addition, at the start of every journal entry, the user was asked whether the chatbot should summarize their last journal entries. This feature was added to motivate the usage of the application beyond what is possible on paper. Furthermore, rereading a learning journal can be productive work in itself (Moon, 2019). By reflecting on previous journal entries (meta-reflection), students can look for patterns and changes in their thinking over time (Stevens & Cooper, 2009). At the same time, engaging in rereading large portions of the journal is bound to take a long time and ties up the cognitive resources of the learner (Barrouillet et al., 2007; Nückles et al., 2020). However, current generative AI models can easily generate automated summaries and offer a lightweight alternative to re-reading the full journal entries while supporting users to get a better overview and reflect on their previous journal entries (Siriwardhana et al., 2023). In our implementation, we added a summary of the last five journal entries before every new journal entry. The summary included both a general summary as well as highlighted 5 topics the user focused on in their recent journal entries.

This feature also utilized OpenAI's (2023) *gpt–3.5-turbo-1106* model. To steer the format and style of the generated summaries, the summary feature used instruction-based prompting with schema-constrained function calling. In addition, the schema descriptions contained short examples of acceptable summaries and focus topic summaries, which provided exemplar-based guidance for the model outputs.

To help students remember to fill in their learning journal, we provided notifications because an important factor in adopting a new behavior is applying it consistently (Lally

& Gardner, 2013). In the context of using learning journals to investigate and support learning processes in higher education, this means that students must consistently create journal entries. A common tool to help attain this consistency is to provide users with reminders (Oinas-Kukkonen & Harjumaa, 2009). In the case of our system, we implemented notifications that reminded users that they had not created a journal entry for the day at 9 P.M. The notification was skipped if a user had already completed a journal entry for that day. Based on the work on time logging by Tabuenca et al. (2015) most users should have already completed their entry by this time, and based on the work by Schlarb and Kulessa (2012) the time should still be early enough that those that didn't create an entry yet had enough time to complete their journal entry before they went to bed.

Method

We conducted a three-week randomized field experiment to investigate the effectiveness of our mobile chatbot-based learning journaling system with the implemented design principles introduced above. We aimed to test our hypotheses using a randomized field experiment with a 2×2 full factorial design.

Study procedure

The participants for this study were students from many different study programs at a technical university in Germany who had native or near-native German-speaking proficiency and could utilize a personal smartphone to install the application. To recruit these participants, we utilized a panel of students at the university. The panel consists of students at our university who voluntarily sign up to be invited to participate in monetarily or tangibly compensated studies. We aimed at 200 total individuals split into four groups, with an equal distribution of males and females.

The invitations included general information on the procedure, incentives, and a link to join the study. Before participants could join the study, they were re-informed about participation requirements and that participation was entirely voluntary. In addition, they were provided information about the minimum requirements to receive compensation and data protection procedures. They were then informed of the entire study procedure and terms. All participants had to agree to these terms before they were able to participate.

To ensure an equal gender representation, we sent out two identical surveys and closed them each as soon as 100 complete responses were received. As participants who started their survey before it was closed were still allowed to finish it before the announced signup period was up, we ended up with slightly more participants. Invitations to participate in the study were sent to 751 males and 449 females. Overall, we received 210 complete responses to the pre-survey (108 males & 102 females). After the pre-survey, the participants were assigned to one of the four groups, which each group presenting one feature combination of our 2×2 design. The assignment was done using stratified randomization, based on gender (male/female) ($\chi^2 = 6.24, df = 9, p = 0.72$), age ($F = 0.76, p = 0.52$), and the four scales of the LIST-K (Klingsieck, 2018) (cognition, $F = 0.35, p = 0.79$; metacognition, $F = 0.34, p = 0.80$; strategies for internal resources, $F = 0.32, p = 0.81$; strategies for external resources, $F = 0.21, p = 0.89$). As a result, 53 of the 210 participants were assigned to the *baseline* (B) group that did not receive any of the support features, 53 to the *assistant supported* (A) group, 52 to the *course supported* (C) group, and 52 to the *course and assistant supported* (CA) group.

After the pre-survey, the participants were invited to install the application on their private smartphones and complete the ~10-minute in-app onboarding within the next two days. The participants who completed the onboarding within the set timeframe could use the application for the next 22 days. After the usage period, participants had to fill out a post-survey. For their participation in the study, all participants who fulfilled the requirements were compensated with €25. The requirements to receive the incentive were that the participants had to fill out the pre- and post-experiment surveys, install the application on their smartphones, and finish the application onboarding process within the two-day cutoff window. The application usage after the initial onboarding was completely voluntary and not tied to any compensation. All participants were informed about the study procedure and the requirements to receive the compensation in the study invitation and again before starting the pre-survey. Participants who could no longer receive the compensation (e.g., because they did not install the application in the said timeframe) were excluded from the study from that point forward and are not included in our evaluation. To test the stability of the effects, a follow-up survey was conducted 12 weeks after the post-survey. Participants were only invited to the follow-up survey if they completed the post-survey and were incentivized with an additional €4.

An overview of the full study procedure can be seen in Fig. 2.

Measures

We used validated scales from the literature to measure the treatments' effects on participants' intrinsic motivation for creating a learning journal with the application and their usage of SRL strategies. SRL was assessed in all three surveys (pre-, post-, follow-up) using the LIST-K questionnaire (Klingsieck, 2018). It is the short version of the LIST questionnaire (*Lernstrategien im Studium*, German for *learning strategies in academic studying*) inventory, a German adaptation of the commonly used MSLQ questionnaire (Pintrich et al., 1991; Schiefle & Wild, 1994). This scale measures SRL by assessing the use of SRL strategies for cognition, metacognition, and internal and external resource utilization. Intrinsic motivation was evaluated in the post-survey using the 22-item version of the Intrinsic Motivation Inventory (IMI) (Center for Self-Determination Theory, n.d.). The IMI measures self-reported intrinsic motivation (as enjoyment) and the subdimensions of perceived autonomy (as perceived choice and pressure), perceived competence, and effort.

We collected the following log data to complement the self-reports with behavioral data: we tracked the creation time of journal entries, the responses to all prompts, and which assistance features were used during the journaling process. Based on this data, we operationalized the engagement with the learning journal through the amount of

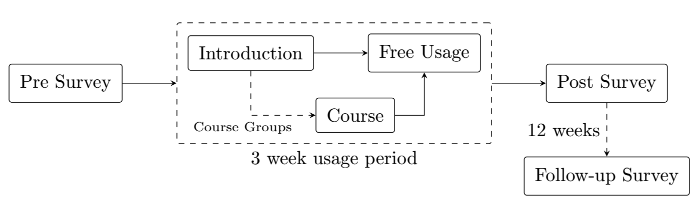

Fig. 2 Study overview

writing the participants did per prompt (measured in characters) in their learning journal (cf., Bråten et al., 2022). Number of characters was chosen as a measure for behavioral engagement instead of the total number of journal entries per participant because we suspected it should be less susceptible to external motivational factors like the daily notifications. In addition, we captured when each journal entry was created and which assistant features were used during its creation.

Results

In total, 200 participants installed the application on their smartphones and completed the onboarding session within the set timeframe. The post-survey questionnaire was completed by 179, and the follow-up questionnaire by 120 participants. The following analyses only include participants who completed the steps required for study inclusion, namely the pre-survey, the onboarding, and the post-survey.

Overall, the participants created 1904 journal entries using the application during the three-week study period. On average, each participant created 10.64 journal entries ($SD = 6.46$). While creating their journal entries, the participants with access to the assistant (groups A and CA) utilized it for every 0.23 of their journal entries. The course was completed by 75.58% of users with the course (groups C and CA).

In addition to the two major design principles, our application included some additional features. We evaluated the impact of the notifications and summaries by applying an exploratory approach combining quantitative and qualitative analyses. For the qualitative analyses, we included open questions in the follow-up survey. A single coder initially labeled the responses to these questions using open coding. After agreeing on category codes, two coders (including the original coder) re-coded the answers using categorical coding. Finally, we consolidated disagreements and settled on the final coding system. The mean Cohen's Kappa for the independent coding done by the two coders was 0.81 ($SE = 0.05$).

Intrinsic motivation

To test the influence of our design on the intrinsic motivation to create a learning journal (H1 & H2), we used self-reported data for intrinsic motivation that were measured through the *enjoyment* scale of the IMI. To ensure that we evaluated the actual effect of using the system, we did this analysis by excluding the participants who did not sufficiently use the system. But in order to avoid a potential selection bias, we also repeated the analysis with all participants included. We defined insufficient usage as having less than 4 days with journal entries, which coincides with the 15% quantile. This corresponded to the exclusion of 22 participants, leaving 157 *active participants* in the analyses presented here.

To test the effect of our design on intrinsic motivation, we employed a type-II ANOVA. This revealed a small significant effect ($\eta^2 = 0.03$) of the course ($F(1, 153) = 4.81, p < 0.05$), thus confirming H1. Regarding H2, no significant effects could be found for the assistant ($F(1, 153) = 0.18, p = 0.67$). There were also no significant interaction effects between the features ($F(1, 153) = 0.09, p = 0.77$). The results over all participants were similar for the course ($\eta^2 = 0.03, F(1, 175) = 5.96, p < 0.05$), assistant ($F(1, 175) = 0.48, p = 0.49$) and interaction ($F(1, 175) = 0.40, p = 0.53$).

Perceived competence and autonomy

We employed an ANOVA to analyze the difference in perceived competence between active participants based on the features the participants had access to. The treatment-based comparison indicated a small ($\eta^2 = 0.04$) but significant effect of the course on perceived competence ($F(1, 154) = 5.77, p < 0.05$). There was no significant difference in perceived competence between users who had access to the assistant and those who had not($F(1, 154) = 0.00, p = 0.95$). As with enjoyment, the results for competence over all participants were similar for the course ($\eta^2 = 0.05, F(1, 176) = 9.26, p < 0.01$) and the assistant ($F(1, 176) = 0.37, p = 0.54$).

Because the assistant was used by only 55.91% of the participants after the onboarding session, we investigated whether actual usage of the assistant would lead to a difference in perceived competence. To investigate this, we modeled perceived competence dependent on the number of days the assistant was used and the number of days a journal entry was created. This revealed no significant impact of the number of days the assistant was used ($t = 1.95, p = 0.054$) on the perceived competence, and no significant effect on the number of days a journal entry was created at ($t = 0.18, p = 0.86$).

The ANOVAs for choice were not significant for both the course ($F(1, 154) = 0.05, p = 0.83$) and the assistant ($F(1, 154) = 0.70, p = 0.40$). Here again the results were similar for the course ($F(1, 176) = 0.01, p = 0.92$) and the assistant ($F(1, 176) = 0.68, p = 0.41$), when all participants were included.

We also could not find any significant impact of the course ($F(1, 154) = 0.13, p = 0.72$) or the assistant ($F(1, 154) = 0.38, p = 0.54$) on pressure. When including all participants, we again observed similar results for the course ($F(1, 176) = 0.18, p = 0.67$) and the assistant ($F(1, 176) = 0.09, p = 0.77$).

The number of days the assistant was used had an impact on choice ($t = 2.05, p < 0.05$) but not on pressure ($t = -1.36, p = 0.18$).

Effects of notifications

As external triggers can influence action, we explored the behavioral impact of the notifications on the participants (Papies & Aarts, 2016). We investigated whether the participants relied on the notifications as an external trigger to create their journal entries or whether they did so without a reminder. Our initial assumption was that some users would create their journal entries without relying on the notifications, and some users would only create their entries after they received a notification reminding them to do so. For this analysis, we categorized users as generally *early* or *late* journal entry writers. Early users, in this case, were all who created their journal entries more than 50 % of the time before they received the notification that reminded them to create their journal entries; other users were late users. The contingency tables for early and late users with the course and assistant features can be seen in Table 1a and b respectively.

Table 1 Early/Late users contingency tables based on feature access

(a) By course access				(b) By assistant access			
	Early	Late	Total		Early	Late	Total
No course	26	50	76	No assistant	32	41	73
Course	52	29	81	Assistant	46	38	84
Total	78	79	157	Total	78	79	157

Fisher exact tests revealed that users with the course were significantly more often early ($OR = 3.45, p < 0.001$). The assistant users' results were not significant ($OR = 1.55, p = 0.11$).

To supplement these findings qualitatively, we asked the participants in the follow-up survey how they reacted to the notifications. As expected, most (84) users indicated that the notifications caused them to open the app and create their journal entries. In contrast, only a comparatively small group (23) indicated that the notifications generally had no effect and only very few participants (9) felt stressed by the notifications.

Behavioral engagement

To analyze the engagement with our system, we ran a multiple regression analysis modeling the number of characters written based on feature access and the day of the study. We included the day of the study because prior work on engagement with applications usually shows decreasing engagement over time (see, e.g., Baumel et al., 2019).

For the analysis, we focused on text length as an indicator of behavioral engagement. We used the number of characters participants wrote in the 7286 responses to journaling prompts of the chat-based learning journal. We focus on text length as an indicator of behavioral engagement, not cognitive engagement or reflective depth, because the latter cannot be inferred from log data alone (Ben-Eliyahu et al., 2018). In writing tasks, behavioral engagement has often been operationalized through observable productivity measures such as writing time and response length (Bråten et al., 2022; Fleckenstein et al., 2024; Namkung & Kim, 2024). All other messages, including small talk and greetings, were excluded from this analysis.

As this data contained large outliers in the form of single word responses and very long responses ($M = 56.82$, $SD = 75.24$, $\max = 1567$), we removed the top 1% of message lengths and removed messages that only consisted of a single word. This left us with 5181 unique messages ($M = 71.87$, $SD = 63.68$, $\max = 329$). The resulting regression was significant ($R^2_{\text{adj}} = 0.06, F(4, 1616) = 24.84, p < 0.001$). As seen in Table 2, the regression shows that both features had a significant impact on the amount written by the users, thus confirming H3 and H4. In general, the mean (A: 71.25, C: 79.26, CA: 83.04) and median (A: 51, C: 59, CA: 63) message lengths in all treatment groups were consistently longer than in the baseline group ($\text{mean} = 51.21$, $\text{median} = 37$). To rule out a selection bias introduced by removing specific entries, we re-ran the regression with all messages included, leading to the same conclusion ($R^2_{\text{adj}} = 0.03, F(4, 1691) = 15.82, p < 0.001$). As a robustness check and to account for the hierarchical nature of the data, we also estimated a mixed-effects model with written messages nested within participants; this arrived at the same conclusions.

Table 2 Regression results for text length based on feature access

Variable	*B*	*SE*	*t*	95.00% CI	
				LL	*UL*
Intercept	55.62***	2.92	19.06	49.89	61.34
Assistant	18.32***	3.07	5.96	12.29	24.35
Course	22.95***	3.42	6.71	16.25	29.66
Course * Assistant	− 15.21***	4.59	− 3.31	− 24.21	− 6.21
Day	− 0.77***	0.19	− 4.00	− 1.15	− 0.39

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$

We also analyzed potential influencing factors affecting the response length to gather further insights into how the features influence engagement. Firstly, we investigated whether users with the assistant wrote more, even if they did not use the assistant beyond the one required usage that was part of the introduction session. The result of this analysis revealed no significant difference $(F(1, 1158) = 1.33, p = 0.25)$ between the means of users with the assistant that only used the assistant a single time ($M = 56.31$, $SD = 43.03$) and the other users that had no access to the assistant ($M = 59.22$, $SD = 40.54$). Secondly, we tested the correlation of message length with the day the journal entry was created, the number of course days the user completed (*course days*), and the total number of days the participant used the assistant (*assistant days*). To keep the number of *course days* and the number of *assistant days* comparable, we clamped the maximum number of *assistant days* to the maximum number of *course days* (eight, including the introduction) for this analysis. The resulting regression seen in Table 3 was significant ($R^2_{\text{adj}} = 0.07, F(3, 723) = 19.32, p < 0.001$).

Following these, the amount of characters written slowly declines each day a journal entry is created, though the effect is relatively small compared to the average message length (71.87). In addition, while the previous analysis indicates that the course positively influences the message length overall, this analysis revealed that the course's impact was relatively constant, as the number of course days completed did not affect the message length. In the assistant case, we previously saw a minor overall impact of the treatment, with this analysis revealing a positive correlation between the days the assistant is used and the number of characters written per message ($B = 6.20, p < 0.001$), indicating that the messages the users send got longer the more they interacted with the assistant.

Self-regulated learning and self-reflection

In addition to testing our hypotheses, we validated the impact of our system with regards to improving SRL and self-reflection. To do so, we analyzed the change in each SRL sub-scale over time between the surveys (pre, post, follow-up). The comparisons only included participants who filled out both surveys. As seen in Table 4, significant differences exist between the pre-survey and the follow-up survey regarding the participants' usage of cognitive and metacognitive self-regulation strategies (including self-reflection) as well as resource-related strategies.

Our qualitative analysis of the respective question in the follow-up survey revealed that the summaries were considered helpful for the following reasons. Of the 97 participants that provided a reason for use 32 specifically indicated that the summaries helped them reflect on their previous summary entries, even though we never told the participants that was the intended purpose of the summaries (neither during the study nor during the survey). For example, P158 explicitly answered the question with "because it helped me reflect on my [journal] entries". Other than reflection, another reason for

Table 3 Regression results for message length based on treatment and usage

Variable	*B*	*SE*	*t*	95 % CI	
				LL	*UL*
Intercept	80.49***	16.59	4.85	47.92	113.07
Course days	− 0.34	2.14	− 0.16	− 4.55	3.87
Assistant days	6.20***	0.98	6.30	4.27	8.13
Day	− 1.44***	0.32	− 4.51	− 2.07	− 0.81

**p* < 0.05. ***p* < 0.01. ****p* < 0.001

Table 4 SRL changes over time

Survey	Scale	Total		Baseline		Assistant		Course		*F*	*p*	*n*
		M	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*			
1	C	3.32	(0.49)	3.31	(0.45)	3.30	(0.53)	3.34	(0.47)	–	–	157
2	C	3.38	(0.52)	3.38	(0.54)	3.32	(0.52)	3.41	(0.50)	3.737	0.055	157
3	C	3.58	(0.50)	3.40	(0.38)	3.60	(0.52)	3.65	(0.49)	22.677	**0.000**	107
1	MC	3.47	(0.68)	3.50	(0.74)	3.44	(0.66)	3.50	(0.65)	–	–	157
2	MC	3.52	(0.65)	3.55	(0.66)	3.50	(0.67)	3.57	(0.63)	1.013	0.316	157
3	MC	3.77	(0.63)	3.80	(0.65)	3.76	(0.60)	3.76	(0.60)	24.030	**0.000**	107
1	IR	3.24	(0.61)	3.15	(0.60)	3.30	(0.61)	3.21	(0.65)	–	–	157
2	IR	3.23	(0.60)	3.23	(0.65)	3.26	(0.57)	3.19	(0.62)	0.095	0.758	157
3	IR	3.30	(0.63)	3.26	(0.62)	3.33	(0.62)	3.25	(0.64)	0.195	0.659	107
1	ER	3.35	(0.51)	3.28	(0.50)	3.34	(0.55)	3.40	(0.49)	–	–	157
2	ER	3.37	(0.56)	3.36	(0.57)	3.38	(0.56)	3.38	(0.56)	0.320	0.572	157
3	ER	3.44	(0.58)	3.31	(0.58)	3.39	(0.55)	3.53	(0.59)	4.256	**0.042**	107

1 pre-survey, 2 post-survey, 3 follow-up survey

C cognition, *MC* metacognition, *IR* strategies for internal resources, *ER* strategies for external resources

F & p are the ANOVA results between the first survey and a given survey (combined for all groups)

Statistically significant p-values ($p < 0.05$) are shown in bold

using the summaries was how well the participants believed they remembered the content of their previous journal entries. For example, P116 indicated that he/she used the summaries "to better remember [his/her] answers". In contrast, P187 stated he/she "did not use the summaries because [he/she] could remember well what [his/her] last [journal] entries look like".

Discussion

Research so far has focused primarily on learning journaling as a teaching tool or to assess learning results. However, it has long been argued that there needs to be more research on how to support learning journaling itself (Jarvis, 2001). We argue that there is a research gap regarding the design of mobile chatbot-based learning journaling systems supporting students' motivation and engagement in generating their learning journal entries over extended time periods (Guan et al., 2025; Huang et al., 2025). Previous research on SRL included intrinsic motivation as a dependent variable but did not focus on the motivation to engage in SRL as a whole (see, e.g., Broadbent et al., 2020; Fabriz et al., 2014). Therefore, we conducted a study on the influence of the proposed design principles on students' intrinsic motivation and engagement to keep a learning journal. The present study indicates that our design fosters behavioral engagement and partly intrinsic motivation when creating a learning journal.

Contrary to previous research (see, e.g., Ewijk et al., 2015; Fabriz et al., 2014) that utilized structured learning journals, we showed that our reflective journal writing system with an integrated prompting concept positively affected SRL in all groups, including the baseline group. More specifically, the significant differences between the pre and the follow-up survey regarding the use of cognitive and metacognitive strategies (including reflective activities) as well as resource management strategies revealed increasing SRL strategy application in our student sample. In addition, the reasons for using the summaries given by the participants align with the idea that the summaries help to reduce cognitive load during journal writing. These results are in line with the findings of Nückles et al. (2020).

In terms of promoting motivation and engagement, we hypothesized that our design principles would positively influence the behavioral engagement of users compared to the baseline group (H3 & H4). We also hypothesized that the course and assistant would positively affect intrinsic motivation compared to the users that did not have access to these features (H1 & H2). Our hypotheses testing found a small but significant increase in behavioral engagement, operationalized as the amount of user-authored text per prompt, for the course participants and for the students using the assistant. Intrinsic motivation was increased in the course participants. In addition, a significant positive effect on perceived competence was revealed for the participants of the course group. Moreover, our exploratory analysis showed that while there were no significant differences in the number of days users created a journal entry, course users tended to rely less on notifications to create their daily journal entries. At the same time, we have reason to believe that the benefits of the course might only be temporary: In a post-hoc analysis, we compared the number of journal days for each participant in the course groups (C & CA) with the number of journal days for participants without the course (groups B & A). This revealed that while there is no significant increase in users with the course that had eight journal days ($OR = 1.41, p = 0.16$), significantly fewer participants with the course had ten days or more ($OR = 0.38, p < 0.001$). As the minimum number of journal days required to finish the course was seven, many participants completed the course, tried out the application for one more day, and then quickly stopped using it by having created eight journal entries in total (in accordance with the participants that were not provided with the course feature). This result echoes the finding of Lee et al. (2021) that the user attributes a large part of the application's value to one particular feature, and the lack of this feature causes them to stop using the application (like in our case, the application was the same for all groups, after the intervention). Possibly, following the theory by Bickmore et al. (2010), the reduction of perceived benefit led the participants to stop the usage of the application (Lee et al., 2021). Our analysis of behavioral engagement prints a similar picture: While the course had the largest impact on engagement, the effect was only related to having access to the course and not correlated with the number of days the students participated. As a result, the static increase in behavioral engagement provided by the course can be expected to fade after the course ends. From a practical perspective, this finding suggests that one-off static guidance may be sufficient to stimulate early journaling activity, but not to sustain stronger behavioral engagement over longer periods without additional reinforcement. For the design of reflective learning tools this implies that this scaffolding may need to be complemented by recurring or adaptive support if the goal is to maintain engagement over time. This could include follow-up prompts, phase-specific guidance, or timely interventions that respond to learners' changing needs.

Because of its nature as a fully optional feature, the effectiveness of the assistant was subject to the willingness of the participants to use it. Deriving our hypotheses from the literature, we envisioned the assistant and course would have a positive effect on engagement with the learning journal (H3 & H4). In our exploratory analyses, we also considered the temporal development. We saw some evidence of increasing intrinsic motivation (in terms of choice) in the students actively using the assistant and in their behavioral engagement. In addition, engagement increased over time with assistant usage but was unrelated to the amount of completed course days and decreased with journal entries created (cf. Table 3).

Taking a social cognitive view on the interaction between course or assistant usage and engagement suggests that users see these interventions as a model and adapt their behavior accordingly: According to Schunk and Zimmerman (1997), Schunk (1999), Usher and Schunk (2018), the development of new SRL skills is initially driven mainly by observing and imitating. The different modeling behaviors the course and assistant provided could explain how the engagement changes with time: Because the examples in the course stayed roughly the same length throughout the course, the model did not change, and users had no new outside influences to adapt their behavior. In contrast, the model the assistant provided was able to change. Future research could investigate if there is some positive feedback loop during the interaction with the assistant, that increases engagement over time. Our hypothesis is that as the users learned to write longer initial messages over time, the assistant would respond with longer messages, resulting in a feedback loop (in our testing, the assistant always responded with a longer message than the user initially provided). This would also explain the additional benefit of the combined treatment: The course examples provided an initial model that the assistant could replace over time.

Limitations and future work

To explore the effects of our chatbot-based learning journaling system, we conducted an experimental field study with a three-week usage period. This has several limitations. While the study was conducted over multiple weeks, it is still difficult to derive conclusions regarding the long-term effects of the proposed design principles. In particular, the declining rate of journal entry creation after course completion suggests that longer multi-phase interventions are needed to fully understand how to best utilize the proposed design principles. In addition, while the application usage (except for the onboarding session) was not compensated, and this fact was communicated to the participants before they started the study, the participants might still have felt required to use the application because of the study setting and the fact that they were part of a panel for study participation. Future work could extend on this and the previous point by providing a version of the app after the study ends to monitor users' long-term behavior.

Another potential issue is the environment and timing of the study: Because the surveys were provided at different times of the semester, there might be confounding effects due to seasonal and semester effects influencing the well-being of students (Lukmanji et al., 2020; Pitt et al., 2018). Because of the interconnectedness between SRL and well-being, seasonal and semester effects might influence the reported SRL (Boekaerts, 2011).

In addition, while the free-will approach to using the assistant and summary features should improve external validity, it comes at the risk of internal validity. On the one hand, the assistant was not used enough to impact most participants. This could either result from the assistant's lack of perceived value or because the assistant was not as tightly integrated as the summaries or the course. On the other hand, from the perspective of personalized support or adaptive learning, it seems to make sense that the AI assistant was useful for a specific group of participants for promoting motivation or maintaining engagement and that other participants did not work with the AI assistant at all or applied it only at the beginning of their reflection process. In general, a challenge for future research activities on adaptive learning will be to analyze which type of support is appropriate for particular student subgroups at specific time points. Nevertheless, in the present study a self-selection bias regarding the engagement of participants that used the assistant cannot be ruled out.

Lastly, our implementation is just one instantiation of the design principles we laid out. The observed effects might not always result from the general concept of the intervention but from issues in our implementation. Future work could investigate the existing features in isolation or require the participants to use them to gather more meaningful data on their effects, provide alternative implementations, and explore new tools. Another research avenue would be to investigate the effects of summarization on the reflection process. In this study we included them for all participants to improve the value of the chatbot and did not isolate their effects.

A further limitation concerns the operationalization of engagement. Our primary behavioral metric was response length in characters, complemented by journal-entry timing and notification reliance. This does not capture cognitive engagement, reflective depth, or the quality of the produced journal entries. Future work should therefore combine log-based measures with qualitative coding of reflection quality in addition to the measures presented here.

Conclusion

In this work, we investigated how the proposed design principles implemented in a mobile chatbot-based learning journaling systems fostered the creation of learning journals. In contrast to previous work, we focused on the motivational and engagement aspects of the journaling process. We demonstrate that the design principles presented in this paper can enhance the motivation of students to keep a learning journal and increase their engagement. With our findings, we also contribute a way to scaffold the effective creation of learning journals that do not require external SRL training. We also showed that usage patterns greatly depend on the implemented design principle. The course benefits were largely temporary, as diary entries were made broadly during the required period, with many students completing the minimum number of days but discontinuing soon after. In contrast, the assistant was used less frequently overall, but we saw indications of a feedback loop where repeated interaction with the assistant promoted users to write more and to create more elaborate entries. Taken together, the usage frequencies and effect patterns suggest that while the course provides short-term benefits to intrinsic motivation, the assistant offers the opportunity to foster more durable engagement dynamics by continuously modeling and reinforcing effective journaling behavior.

Author contributions

S.S. wrote the main manuscript text and prepared the figures. All authors reviewed the manuscript.

Data availability

The datasets, created, used and analyzed during this study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

According to our institutional rules, formal ethics approval was not required. At the same time great care was taken to conduct this research in an ethical manner. The study was conducted in accordance with the ethical principles for research involving human subjects and the specific guidelines for research and data protection of our university. To ensure data protection of the participants, the study was pre-discussed with the data protection officer at our institution. In addition, the recruitment of and all interaction with the participants (including payments) were handled by a panel that pseudonymizes all participants. At no point any of the researchers had access to identifying information of the participants except for voluntarily provided first names and ages. All interaction with OpenAI was done in a way to minimize the chance of identifying the individuals involved in the study, and no identifying information was shared with this service. Moreover, the participants were prohibited from including personal information about themselves and others in their journal entries. Finally, all participants provided written informed consent to voluntarily participate in the study and were paid once they fulfilled the required activities (pre-survey, introduction session, post-survey). The behavior of the participants beyond these requirements did not in any way influence the compensation. In order to fairly compensate the participants it was set in a way to exceed the local minimum wage when just doing the required steps of the study.

Competing interests
The authors declare no conflict of interest.

Received: 29 January 2026 / Accepted: 30 March 2026
Published online: 29 April 2026

References

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. *Journal of Medical Internet Research*. https://doi.org/10.2196/16021

Achtziger, A., & Gollwitzer, P. M. (2018). Motivation and volition in the course of action. In J. Heckhausen, H. Heckhausen (Eds.), *Motivation and action* (pp. 485–527). Springer International Publishing.

Alt, D., & Raichel, N. (2020). Reflective journaling and metacognitive awareness: Insights from a longitudinal study in higher education. *Reflective Practice, 21*(2), 145–158. https://doi.org/10.1080/14623943.2020.1716708

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*(2), 181–214.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*(3), 344–370. https://doi.org/10.1016/j.cedpsych.2003.09.002

Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C., & Greene, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science, 33*(5–6), 381–412. https://doi.org/10.1007/s11251-005-1273-8

Bain, J. D., Mills, C., Ballantyne, R., & Packer, J. (2002). Developing reflection on practice through journal writing: Impacts of variations in the focus and level of feedback. *Teachers and Teaching, 8*(2), 171–196. https://doi.org/10.1080/13540600220127368

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598. https://doi.org/10.1037/0022-3514.41.3.586

Barnard-Brak, L., Paton, V. O., & Lan, W. Y. (2010). Profiles in self-regulated learning in the online learning environment. *The International Review of Research in Open and Distributed Learning, 11*(1), 61–80. https://doi.org/10.19173/irrodl.v11i1.769

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Baumeister, R. F., & Vohs, K. D. (2007). Self-regulation, ego depletion, and motivation: Motivation and ego depletion. *Social and Personality Psychology Compass, 1*(1), 115–128. https://doi.org/10.1111/j.1751-9004.2007.00001.x

Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *Journal of Medical Internet Research*. https://doi.org/10.2196/14567

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology, 53*, 87–105. https://doi.org/10.1016/j.cedpsych.2018.01.002

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Bickmore, T., Schulman, D., & Yin, L. (2010). Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence, 24*(6), 648–666. https://doi.org/10.1080/08839514.2010.492259

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society.* (2nd ed., pp. 56–64). Worth Publishers.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*(6), 445–457. https://doi.org/10.1016/S0883-0355(99)00014-2

Boekaerts, M. (2011). Emotions, emotion regulation, and self-regulation of learning. *Handbook of self-regulation of learning and performance* (pp. 408–424). Routledge.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*(3), 199–210. https://doi.org/10.1007/s10648-006-9013-4

Brandtzæg, P. B., Skjuve, M., Dysthe, K. K., & Følstad, A. (2021). When the social becomes non-human: Young people's perception of social support in chatbots. *Proceedings of the 2021 CHI conference on human factors in computing systems* (p. 13). ACM.

Broadbent, J., Panadero, E., & Fuller-Tyszkiewicz, M. (2020). Effects of mobile-app learning diaries vs online training on specific self-regulated learning components. *Educational Technology Research and Development, 68*(5), 2351–2372. https://doi.org/10.1007/s11423-020-09781-6

Bråten, I., Latini, N., & Haverkamp, Y. E. (2022). Predictors and outcomes of behavioral engagement in the context of text comprehension: When quantity means quality. *Reading and Writing, 35*(3), 687–711. https://doi.org/10.1007/s11145-021-10205-x

Carr, M. (1996). *Teaching children to self-regulate: A resource for teachers. Instructional resource no. 34.* (Guides - Classroom Use Teaching Guides No. ED 400 517). National Reading Research Center.

Center for Self-Determination Theory n.d. *Intrinsic motivation inventory (IMI)*. https://selfdeterminationtheory.org/intrinsic-motivation-inventory/.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior* (1st ed.,). Springer.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Dincel, B. K., & Savur, H. (2019). Diary keeping in writing education. *Journal of Education and Training Studies, 7*(1), 48–59. https://doi.org/10.11114/jets.v7i1.3758

Dörrenbächer, L., & Perels, F. (2016). More is more? Evaluation of interventions to foster self-regulated learning in college. *International Journal of Educational Research, 78*, 50–65. https://doi.org/10.1016/j.ijer.2016.05.010

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Ewijk, C.-D.V., Fabriz, S., & Büttner, G. (2015). Fostering self-regulated learning among students by means of an electronic learning diary: A training experiment. *Journal of Cognitive Education and Psychology, 14*(1), 77–97. https://doi.org/10.1891/1945-8959.14.1.77

Fabriz, S., Dignath-van Ewijk, C., Poarch, G., & Büttner, G. (2014). Fostering self-monitoring of university students by means of a standardized learning journal—A longitudinal study with process analyses. *European Journal of Psychology of Education, 29*(2), 239–255. https://doi.org/10.1007/s10212-013-0196-z

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies, 132*, 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Fleckenstein, J., Jansen, T., Meyer, J., Trüb, R., Raubach, E. E., & Keller, S. D. (2024). How am I going? Behavioral engagement mediates the effect of individual feedback on writing performance. *Learning and Instruction, 93*, 101977. https://doi.org/10.1016/j.learninstruc.2024.101977

Fröscher, L., Friedrich, A.-K., Berentelg, M., Widmer, C., Gilbert, S. J., & Papenmeier, F. (2022). Framing cognitive offloading in terms of gains or losses: Achieving a more optimal use of reminders. *Cognitive Research: Principles and Implications, 7*(1), 61. https://doi.org/10.1186/s41235-022-00416-3

García, T., & Pintrich, P. R. (1991). *Student motivation and self-regulated learning: A LISREL model. Annual Meeting of the American Educational Research Association*. ERIC.

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology, 68*(5), 971–992. https://doi.org/10.1080/17470218.2014.972963

Greenleaf Brown, L., Briscoe, G. S., & Grabowsky, A. (2022). The Influence of journaling on nursing students: A systematic review. *Journal of Nursing Education, 61*(1), 29–35. https://doi.org/10.3928/01484834-20211203-01

Guan, R., Raković, M., Chen, G., & Gašević, D. (2025). How educational chatbots support self-regulated learning? A systematic review of the literature. *Education and Information Technologies, 30*(4), 4493–4518. https://doi.org/10.1007/s10639-024-12881-y

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction, 20*(1), 18–29. https://doi.org/10.1016/j.learninstruc.2008.12.001

Hiemstra, R. (2001). Uses and benefits of journal writing. *New Directions for Adult and Continuing Education, 90*, 19–26. https://doi.org/10.1002/ace.17

Hofer, B. K., Shirley, L. Y., & Pintrich, P. R. (1998). Teaching college students to be self-regulated learners. In D. H. Schunk, B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 57–85). Guilford Publications.

Huang, W., Jiang, J., King, R. B., & Fryer, L. K. (2025). Chatbots and student motivation: A scoping review. *International Journal of Educational Technology in Higher Education, 22*(1), 26. https://doi.org/10.1186/s41239-025-00524-2

Hubbs, D. L., & Brand, C. F. (2005). The paper mirror: Understanding reflective journaling. *Journal of Experiential Education, 28*(1), 60–71. https://doi.org/10.1177/105382590502800107

Hume, A. (2009). Promoting higher levels of reflective writing in student journals. *Higher Education Research and Development, 28*(3), 247–260. https://doi.org/10.1080/07294360902839859

Jarvis, M.-A., & Baloyi, O. B. (2020). Scaffolding in reflective journaling: A means to develop higher order thinking skills in undergraduate learners. *International Journal of Africa Nursing Sciences, 12*, 100195. https://doi.org/10.1016/j.ijans.2020.100195

Jarvis, P. (2001). Journal writing in higher education. *New Directions for Adult and Continuing Education, 2001*(90), 79–86. https://doi.org/10.1002/ace.23

Järvelä, S. (1995). The cognitive apprenticeship model in a technologically rich learning environment: Interpreting the learning interaction. *Learning and Instruction, 5*(3), 237–259. https://doi.org/10.1016/0959-4752(95)00007-P

Kasprabowo, T., Rahayu, E. Y., & Widyaningrum, A. (2021). Tell me about your day: Portraying students reflective practice through diary writing. *Language Circle: Journal of Language and Literature, 15*(2), 375–384. https://doi.org/10.15294/lc.v15i2.29252

Kawasaki, M., Yamashita, N., Lee, Y.-C., & Nohara, K. (2020). Assessing users' mental status from their journaling behavior through chatbots. *Proceedings of the 20th ACM international conference on intelligent virtual agents* (pp. 1–8). ACM.

Klingsieck, K. B. (2018). Kurz und knapp - die Kurzskala des Fragebogens, Lernstrategien im Studium (LIST). *Zeitschrift für Pädagogische Psychologie, 32*(4), 249–259. https://doi.org/10.1024/1010-0652/a000230

Kocielnik, R., Xiao, L., Avrahami, D., & Hsieh, G. (2018). Reflection companion: A conversational system for engaging users in reflection on physical activity. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*. https://doi.org/10.1145/3214273

Lally, P., & Gardner, B. (2013). Promoting habit formation. *Health Psychology Review, 7*(sup1), S137–S158. https://doi.org/10.1080/17437199.2011.603640

Lee, Y.-C., Yamashita, N., & Huang, Y. (2021). Exploring the effects of incorporating human experts to deliver journaling guidance through a chatbot. *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3449196

Lindroth, J. T. (2015). Reflective journals: A review of the literature. *Update Applications of Research in Music Education, 34*(1), 66–72. https://doi.org/10.1177/8755123314548046

Lipschitz, J. M., Pike, C. K., Hogan, T. P., Murphy, S. A., & Burdick, K. E. (2023). The engagement problem: A review of engagement with digital mental health interventions and recommendations for a path forward. *Current Treatment Options in Psychiatry, 10*, 119–135. https://doi.org/10.1007/s40501-023-00297-3

Liu, M., & Reinders, H. (2025). Do AI chatbots impact motivation? Insights from a preliminary longitudinal study. *System, 128*, 103544. https://doi.org/10.1016/j.system.2024.103544

Luft, F., Miyamoto, A., & Nückles, M. (2026). Writing or speaking? The role of the production medium in the effectiveness of learning journals. *Learning and Instruction, 101*, 102250. https://doi.org/10.1016/j.learninstruc.2025.102250

Lu, H., & Wang, Y. (2022). The effects of different interventions on self-regulated learning of pre-service teachers in a blended academic course. *Computers and Education, 180*, 104444. https://doi.org/10.1016/j.compedu.2022.104444

Lukmanji, A., Williams, J. V., Bulloch, A. G., & Patten, S. B. (2020). Seasonal variation in specific depressive symptoms: A population based study. *Journal of Affective Disorders, 261*, 153–159. https://doi.org/10.1016/j.jad.2019.10.003

Miller, L. B. (2017). Review of journaling as a teaching and learning strategy. *Teaching and Learning in Nursing, 12*(1), 39–42. https://doi.org/10.1016/j.teln.2016.10.004

Moon, J. A. (2019). *Learning journals: A handbook for reflective practice and professional development* (2nd ed.). Routledge.

Morrison, K. (1996). Developing reflective practice in higher degree students through a learning journal. *Studies in Higher Education, 21*(3), 317–332. https://doi.org/10.1080/03075079612331381241

Namkung, Y., & Kim, Y. (2024). Learner engagement in collaborative writing: The effects of SCMC mode, interlocutor familiarity, L2 proficiency, and task repetition. *System, 121*, 103251. https://doi.org/10.1016/j.system.2024.103251

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction, 19*(3), 259–271. https://doi.org/10.1016/j.learninstruc.2008.05.002

Nückles, M., Roelle, J., Glogger-Frey, I., Waldeyer, J., & Renkl, A. (2020). The self-regulation-view in writing-to-learn: Using journal writing to optimize cognitive load in self-regulated learning. *Educational Psychology Review, 32*(4), 1089–1126. https://doi.org/10.1007/s10648-020-09541-1

Nepal, S., Pillai, A., Campbell, W., Massachi, T., Heinz, M. V., Kunwar, A., & Campbell, A. T. (2024). MindScape study: Integrating LLM and behavioral sensing for personalized AI-driven journaling experiences. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies8*(4), 1–44. https://doi.org/10.1145/3699761

Ning, H. K., & Downing, K. (2014). A latent profile analysis of university students' self-regulated learning strategies. *Studies in Higher Education, 40*(7), 1328–1346. https://doi.org/10.1080/03075079.2014.880832

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems, 24*(1), 87. https://doi.org/10.17705/1CAIS.02428

OpenAI (2023). *New models and developer products announced at DevDay.* https://openai.com/blog/new-models-and-developer-products-announced-at-devday.

Papies, E. K., & Aarts, H. (2016). Automatic self-regulation: From habit to goal pursuit: Handbook of self regulation: Research, theory, and applications. In K. D. Vohs, R. F. Baumeister (Eds.) *Handbook of self-regulation: research, theory, and applications* (3rd ed., pp. 203–222). Guilford Press.

Park, C. (2003). Engaging students in the learning process: The learning journal. *Journal of Geography in Higher Education, 27*(2), 183–199. https://doi.org/10.1080/03098260305675

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Perels, F., Otto, B., Landmann, M., Hertel, S., & Schmitz, B. (2007). Self-regulation from a process perspective. *Zeitschrift für Psychologie / Journal of Psychology, 215*(3), 194–204. https://doi.org/10.1027/0044-3409.215.3.194

Pesonen, J. A., Ketonen, E. E., Kivimaki, V., & Ihantola, P. (2020). Does using structured learning diaries affect self-regulation or study engagement? An experimental study in engineering education. (2020). *IEEE frontiers in education conference (FIE)* (pp. 1–9). IEEE.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Tests/Questionnaires No. NCRIPTAL-91-B-004). National Center for Research to Improve Postsecondary Teaching and Learning.

Pitt, A., Oprescu, F., Tapia, G., & Gray, M. (2018). An exploratory study of students' weekly stress levels and sources of stress during the semester. *Active Learning in Higher Education, 19*(1), 61–75. https://doi.org/10.1177/1469787417731194

Ren, X., Zhang, X., Zou, R., Yan, R., & Yu, B. (2025). EmoVis: Exploring data-enabled analogue journaling to promote self-reflection for mental wellness among college students. *Behaviour and Information Technology, 44*(4), 859–881. https://doi.org/10.1080/0144929X.2024.2349182

Reynolds, L., & McDonell, K. (2021). Prompt programming for large language models: beyond the few-shot paradigm. *Extended abstracts of the 2021 CHI conference on human factors in computing systems* (pp. 1–7). ACM.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Scheu, S., Kuhlmeier, F., Müller, L., Benke, Ivo. (2023). Designing Amsl – A Mobile Digital Assistant for Self-Regulated Learning. ECIS 2023 Research-in-Progress Papers, 73. https://aisel.aisnet.org/ecis2023_rip/73/

Schick, A., Feine, J., Morana, S., Maedche, A., & Reininghaus, U. (2022). Validity of chatbot use for mental health assessment: experimental study. *JMIR mHealth and uHealth, 10*(10), e28082. https://doi.org/10.2196/28082

Schiefle, U., & Wild, K. P. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Différentielle und Diagnostische Psychologie, 15*(4), 185–200.

Schlarb, A., & Kulessa, G. (2012). Sleep characteristics, sleep problems, and associations of self-efficacy among German university students. *Nature and Science of Sleep, 4*, 1–7. https://doi.org/10.2147/NSS.S27971

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*(1), 64–96. https://doi.org/10.1016/j.cedpsych.2005.02.002

Schunk, D. H. (1999). Social-self interaction and achievement behavior. *Educational Psychologist, 34*(4), 219–227. https://doi.org/10.1207/s15326985ep3404_3

Schunk, D. H., & Greene, J. A. (2018). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk, J. A. Greene (Eds.) *Handbook of self-regulation of learning and performance* (pp. 1–15). Routledge/Taylor & Francis Group.

Schunk, D. H., & Rice, J. M. (1987). Enhancing comprehension skill and self-efficacy with strategy value information. *Journal of Reading Behavior, 19*(3), 285–302. https://doi.org/10.1080/10862968709547605

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*(4), 195–208. https://doi.org/10.1207/s15326985ep3204_1

Siriwardhana, S., Gupta, C., Kaluarachchi, T., Dissanayake, V., Ellawela, S., & Nanayakkara, S. (2023). Can AI models summarize your diary entries? Investigating utility of abstractive summarization for autobiographical text. *International Journal of Human-Computer Interaction, 40*(24), 1–19. https://doi.org/10.1080/10447318.2023.2286090

Smyth, J. M., Johnson, J. A., Auer, B. J., Lehman, E., Talamo, G., & Sciamanna, C. N. (2018). Online positive affect journaling in the improvement of mental distress and well-being in general medical patients with elevated anxiety symptoms: A preliminary randomized controlled trial. *JMIR Mental Health, 5*(4), e11290. https://doi.org/10.2196/11290

Stevens, D. D., & Cooper, J. E. (2009). *Journal keeping: How to use reflective writing for effective learning, teaching, professional insight, and positive change* (1st ed.,). Stylus Pub.

Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers and Education, 89*, 53–74. https://doi.org/10.1016/j.compedu.2015.08.004

Usher, E. L., & Schunk, D. H. (2018). Social cognitive theoretical perspective of self-regulation. In D. H. Schunk, J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 19–35). Routledge/Taylor & Francis Group.

Vaishnavi, V., & Kuechler, B. (2008). *Design research in information systems* https://web.archive.org/web/20081015015817/http://www.isworld.org/Researchdesign/drisISworld.htm.

Valenzuela, R., Codina, N., Castillo, I., & Pestana, J. V. (2020). Young university students' academic self-regulation profiles and their associated procrastination: Autonomous functioning requires self-regulated operations. *Frontiers in Psychology, 11*, 354. https://doi.org/10.3389/fpsyg.2020.00354

Wallin, P., & Adawi, T. (2018). The reflective diary as a method for the formative assessment of self-regulated learning. *European Journal of Engineering Education, 43*(4), 507–521. https://doi.org/10.1080/03043797.2017.1290585

Wolfbauer, I., Pammer-Schindler, V., Maitz, K., & Rosé, C. P. (2022). A script for conversational reflection guidance: A field study on developing reflection competence with apprentices. *IEEE Transactions on Learning Technologies, 15*(5), 554–566. https://doi.org/10.1109/TLT.2022.3207226

Wolfbauer, I., Pammer-Schindler, V., & Rosé, C. P. (2020). Rebo junior: Analysis of dialogue structure quality for a reflection guidance chatbot. *Proceedings of the Impact Papers at EC-TEL 2020.*

Wong, J., Khalil, M., Suschevskiy, V., Baars, M., De Koning, B., & Paas, F. (2026). Student engagement profiles in a mobile app: Links to self-regulated learning and performance. *Educational technology research and development*. https://doi.org/10.1007/s11423-026-10586-2

Zhao, X., Jin, Y., Wang, A. Y., & Zhang, M. (2026). From tracking to thinking: Facilitating post-exercise reflection by a large language model-mediated journaling system. *Information Processing and Management, 63*(4), 104574. https://doi.org/10.1016/j.ipm.2025.104574

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Elsevier.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J. (2008). Motivation: An essential dimension of self-regulated learning. In D. H. Schunk, B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (1st ed., pp. 1–30). Routledge.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds.), *Handbook of metacognition in education* (1st ed., pp. 299–315). Routledge.

Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614–628. https://doi.org/10.3102/00028312023004614

Publisher's Note