frontiers | Frontiers in Education

TYPE Brief Research Report
PUBLISHED 17 September 2025
DOI 10.3389/feduc.2025.1621151

Check for updates

OPEN ACCESS

EDITED BY
Catalina Lomos,
Luxembourg Institute of Socio-Economic Research, Luxembourg

REVIEWED BY
Hassan Banaruee,
University of Education Weingarten, Germany
Margaret Thomas,
Earlham College, United States

*CORRESPONDENCE
Lars van Rijn
✉ lars.vanrijn@fernuni-hagen.de
Silke Elisabeth Wrede
✉ silke.wrede@fernuni-hagen.de

RECEIVED 30 April 2025
ACCEPTED 01 September 2025
PUBLISHED 17 September 2025

CITATION
van Rijn L, Wrede SE, de Witt C, Wang X, Duong-Trung N, Jose AM and Pinkwart N (2025) Scaffolding recommendations for students' essay planning.
Front. Educ. 10:1621151.
doi: 10.3389/feduc.2025.1621151

COPYRIGHT
© 2025 van Rijn, Wrede, de Witt, Wang, Duong-Trung, Jose and Pinkwart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Scaffolding recommendations for students' essay planning

Lars van Rijn[1]*, Silke Elisabeth Wrede[1]*, Claudia de Witt[1], Xia Wang[2], Nghia Duong-Trung[2], Angelin Mary Jose[2] and Niels Pinkwart[2]

[1]Center of Advanced Technology for Assisted Learning and Predictive Analytics, University of Hagen, Hagen, Germany, [2]Educational Technology Lab, German Research Center for Artificial Intelligence, Berlin, Germany

Planning scientific essays is a fundamental activity for students to acquire and demonstrate research competence in higher education. The planning and writing process is complex and time-consuming, requiring significant effort from both students and tutors. Consequently, recommender systems have been used to support the writing process, but the evidence regarding their pedagogical effectiveness is sparse. In this study we conceptualize and evaluate an approach to closely align a recommender system with pedagogical principles of the specific use case. In our Wizard-of-Oz experiment we therefore used scaffolded recommendations to explore whether this approach can assist students in planning a scientific essay for a course in media education. The intervention resulted in positive effects on the time students' need to be tutored for and the intention of students to reuse the system. We discuss implications of these results for the design of supporting essay planning through recommender systems, considering limitations of the method applied.

KEYWORDS

essay planning, research competence, students, higher education, scaffolding, recommendations

1 Introduction

The ability to write scientific essays is a key skill in higher education (Dahl et al., 2023). While essay writing is widely used in higher education, tasks specifically emulating scientific essay writing follow very high rigor and incorporate academic standards (Halliday and Martin, 2003). Such tasks positively impact students' writing proficiency (Haswell, 2000; Oppenheimer et al., 2017; Dahl et al., 2023). Research competence and scientific writing require a complex skillset (Dahl et al., 2023; Guilford, 2001). Therefore, writing essays in higher education confronts students with affective, cognitive, vocabulary and structure related challenges (Bulquiyah et al., 2021). One key part of the writing process is the planning phase of the essay (Hounsell, 1984). Here, students explore a subdomain of their field, identify a research gap, define their essay topic and form corresponding research questions (Alley, 2018; Singh and Mayer, 2014). This process could be supported through recommender systems (RecSys). RecSys guide users to a useful object when a magnitude of options is possible. This is done through diverse techniques to automate finding optimal matches between users and potential objects of interest (Ricci et al., 2011). An overview of such techniques can be found in Joy and Pillai (2022). Educational RecSys recommend further learning content, activities or sequences to learners or provide recommendations on learning partners (Drachsler et al., 2015). RecSys are used in higher education (Maphosa and Maphosa, 2023), in research processes and scientific writing (Kreutz and Schenkel, 2022; Sun et al., 2024; Zhang et al., 2023) even though evidence for their pedagogical effectiveness is scarce (da Silva et al., 2023). A related

meta-analysis has shown high heterogeneity in the effects of automated feedback tools on students' writing (Fleckenstein et al., 2023), indicating dependence on the use case. However, we could not identify studies where RecSys support students' essay planning. We explore this gap by addressing the challenge of designing recommendations to facilitate essay planning. Since recommendations could be given on several concepts related to scientific essays, reducing the information provided to students at once will be key for achieving a match between student competence and task challenge. This should facilitate students to reach a zone of proximal development (ZPD). ZPD was conceptualized by Vygotsky (1978) and defines a state of development, where learners can achieve something with guidance from more knowledgeable others (e.g., teachers) that they would be unable to independently. Such support builds task competence which increases the pace at which learning occurs. Recommendations could achieve this if their content is structured to enable students to more easily focus on each subtask necessary for essay planning. We propose to realize this by scaffolding recommendation content. Scaffolding is built on the idea of ZPD and was first coined by Wood et al. (1976), introducing it as controlling elements of a task to enable learners to engage in tasks that otherwise would be outside of their capacity. It describes any support given that bridges the gap between a student's current competences and necessary competences, while fading out support over time (van de Pol et al., 2010). Scaffolding strategies aim at support for cognition, metacognition and affect. Cognitive scaffolding focuses on taking over parts of the task, showing model behavior or offering feedback (van de Pol et al., 2010). Metacognitive Scaffolding supports students in staying on task and in pursuit of learning goals, while affective scaffolding supports a conducive affective state (van de Pol et al., 2010). This guidance is effective if done strongly instructional (Hmelo-Silver et al., 2007; Kirschner et al., 2006), requiring one-to-one tutoring, which is highly demanding in higher education with the given teacher-to-student ratios. In a recent meta-analysis, scaffolding has been shown to be highly effective in online learning (Doo et al., 2020). Scaffolding positively affects metacognitive processes in writing, (Jafarigohar and Mortazavi, 2017), reduces cognitive load (Lee and Tan, 2010), and supports self-regulated learning processes (Li et al., 2023). Scaffolding supports science text comprehension and writing (Lin et al., 2014; Lo et al., 2009; Ouyang et al., 2024), showing potential in online learning (Ouyang et al., 2024; Ucar-Longford et al., 2024). Scaffolding can be realized through feedback (Lee and Tan, 2010; Panadero et al., 2016). Feedback is most effective if it incorporates information on the correct result, providing a cognitive scaffold, while including information on potential next steps and self-regulation, incorporating metacognitive and affective scaffolds, so called high information feedback (HIF) (Hattie and Timperley, 2007; Wisniewski et al., 2020). Recommendations can offer information on next steps and potentially relevant learning resources to students (Gloerfeld et al., 2020). The combination of such recommendations with corrective feedback on essay planning would point out potential improvements and offer information on how to proceed, achieving attributes of HIF.

Abbreviations: RecSys, Recommender System; HIF, high information feedback; RQ, Research Question; HXa), Hypothesis n (number) a (letter); DPO, data protection officer; LMS, learning management system; DTAS, Skala zur Erfassung der Digitalen Technologieakzeptanz.

2 Method

In this study we investigate whether scaffolding is a viable approach for recommendations to support students in planning scientific essays. We assume that an appropiate RecSys (addressing the ZPD) can reduce the time spent by instructors in 1-on-1 tutoring, as individual tutoring processes are transferred to the RecSys and essay planning reaches its final stage faster. This could reduce the workload for both students and tutors, enabling them to reallocate resources for more in-depth tutoring or other challenges. To examine the pedagogical effectiveness of the RecSys, we define it as the system's ability to promote student's research competence and increase their satisfaction with the system itself. We aim to determine how the use of the RecSys affects student's self-assessment of their own research competence and how scaffolded recommendations influence students' satisfaction with the RecSys. The corresponding research questions are:

> *RQ1:* How do scaffolded recommendations affect the time students need until they reach sufficient quality of their scientific essay concept?
>
> *H1:* Students provided with scaffolded recommendations need shorter time to reach sufficient quality compared.
>
> (a) to students who are exclusively tutored.
> (b) to those provided with generalized recommendations.
>
> *RQ2:* How do scaffolded recommendations affect the time students need to be tutored to reach sufficient quality of their scientific essay concept?
>
> *H2:* The tutoring process for students provided with scaffolded recommendations is significantly shorter than for students.
>
> (a) receiving generalized recommendations.
> (b) who are exclusively tutored.
>
> *RQ3:* Do students perceive their research competence differently with and without scaffolded recommendations?
>
> *H3:* Changes in students' perception of their own research competence increase if provided with scaffolded recommendations compared to generalized recommendations.
>
> *RQ4:* How do scaffolded recommendations affect the students' perception of the system's ability to provide on-task support?
>
> *H4:* Students provided with scaffolded recommendations perceive a higher ability of the RecSys to support them in their planning task compared to generalized recommendations.

A mixed within-between-subject field experiment was conducted in a B. A. educational science course on media education and communication at a distance university. The final examination is a literature-based scientific essay. We used a Wizard-Of-Oz experimental design, where educational scientists simulated system outputs on a RecSys mockup frontend. Scientists prepared sample responses, which were sent to students via the system, imitating functionality of a fully

developed recommender engine. This ensured a controlled output, separating recommendation design aspects from unexpected system outputs. Because of the asynchronous nature of the distance learning course, recommendations were sent out 3 times a day to students. A faster response time could not be achieved with the given resources. To enhance the output quality, we developed recommendation texts for all quality criteria and used them to provide the recommendations.

2.1 Designing scaffolded recommendations

To adhere to theoretical and empirical implications of effective scaffolding, we defined the content of provided recommendations as follows: to offer cognitive scaffolding, corrective feedback should support students in identifying which parts of their current solution were adhering to quality criteria. Therefore, we designed an assessment rubric with the teachers, from which we could identify what to focus on in the feedback (e.g., "Are one or multiple topics provided by the student?"). Direct feedback towards the achievement of each criterion was provided (e.g., "Nice to see that you have stated one topic and corresponding research question."). Metacognitive scaffolding was achieved through two principles. Firstly, criteria were organized sequentially, offering step-by-step guidance to support progress towards the task goal. Secondly, multiple recommended next steps were outlined for each criterion to enhance the solution or advance the task. (e.g., "You may want to think about how the chosen topic fits into the course's focus domain."). To incorporate affective scaffolding, each criterion was coupled with motivational text related to its achievement (e.g., "Thank you for the work put into your proposed topic, […]") and each potential next step with one that motivated to persist (e.g., "You are approaching your goal!").

2.2 Course design and recommender system mockup

In the course, students prepare their essays throughout the semester in a 1-on-1 tutoring setting. They are given the initial task to derive their research interest and formulate a corresponding essay topic and a research question. In an online forum tutors give corrective feedback on the progress, motivating comments and recommend how to proceed with the planning, providing scaffolding on the cognitive, metacognitive and affective level. Since essay topics are chosen according to students' interests, the task is open-ended. Fittingly, recommendations provide enough leeway for students not to be guided to a specific answer. Rather, they induce self-evaluation of solutions while considering the proposed next steps. However, because empirical evidence for pedagogical effectiveness of RecSys is scarce, pedagogical principles are implemented into the system's design. We identified scaffolding strategies as a mechanism to facilitate positive effects. To simulate the RecSys' functions effectively, essential features were developed as a mockup, providing students with a frontend to enter their essay concepts (Figure 1). A frontend for teachers with an overview of students' inputs and the possibility to send recommendations was provided (Figure 2). The mockup was implemented by using a modular architecture of the MERN stack (i.e., MongoDB, Express.js, React.js, and Node.js), which is suited for scalable and easy-to-maintain web applications. This system allows tutors and students to simultaneously manage and collaborate online on a variety of operations (submission, revisions, feedback, recommendations, status logs, updates, notifications). MongoDB was used to efficiently store each operation on a student's essay concept with an id and timestamp, which allows tracking and analyzing student's responses to recommendations.

2.3 Experimental conditions

Azevedo et al. (2005) found that in a knowledge acquisition setting, scaffolding that adapts to the students' progress is more effective than generalized scaffolding that provides predefined information which is the same for each student. Since the given scenario differs quite significantly from that setting, we designed two conditions. Students were randomly allocated to the conditions, which provided either three consecutive scaffolded recommendations on their solution (Group 1) or a generalized recommendation on the quality criteria for essay concepts (Group 2). This was done to check whether the scaffolded recommendations retain the positive effect found by (Azevedo et al., 2005) when applied to

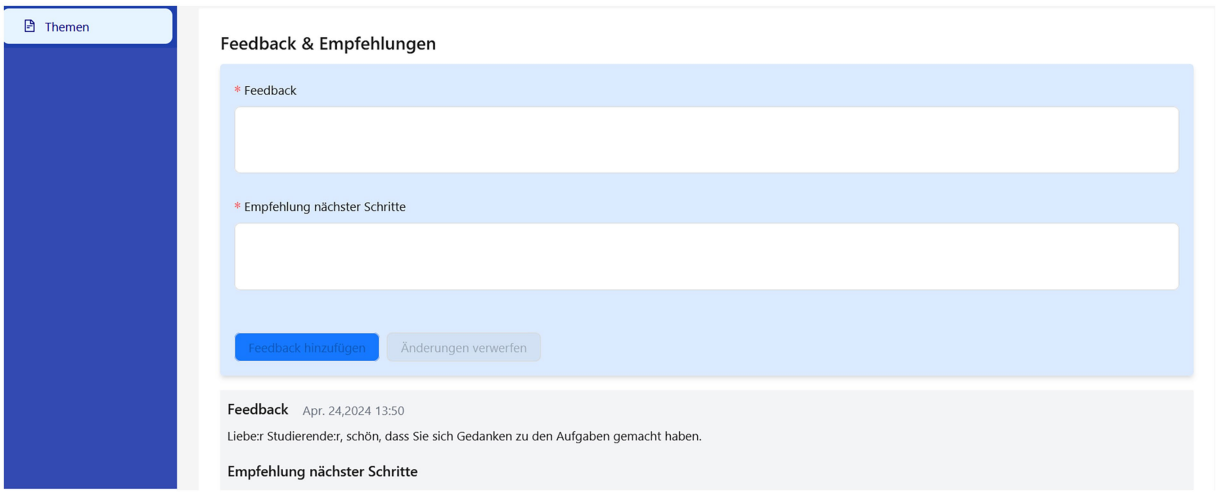

FIGURE 1
Student-facing web-interface.

FIGURE 2
Teacher-facing web-interface.

supporting students' essay planning. To ensure that students did not miss out on additional support, both conditions had comparable content, differing only in the method. A quasi-control group with no interventions was derived from historical data of a previous cohort (Group 3).

2.4 Procedure

To ensure that the study was understood, the study design including information on the experimental conditions were explained at two online seminars prior to the initial survey. Students were informed that which conditions existed and that allocation to each was random. The experiment took place from April to June 2024. Firstly, participants answered the pre-intervention survey. Here, informed consent for participation was requested via consent forms that are developed together with the university's data protection officers (DPO) and the survey was administered through the learning management system (LMS). Thereafter, access to the RecSys-mockup was granted. Students interacted with the mockup as if it was an actual RecSys. Students in the generalized recommendation condition received a single recommendation encompassing all quality criteria, along with guidance to reflect upon them. Students in the scaffolded recommendations condition received feedback on their initial essay concept regarding the first quality criterion, identifying issues and providing recommendations for improvement. A maximum of three resubmissions were allowed. Finally, students were informed about which condition they were allocated to with the final message and received an invitation to the post-intervention survey. All data collection efforts were documented throughout the study and monitored by the university's DPOs.

2.5 Participants

57 students were recruited for the intervention study and took part in the first survey. Participation was optional for each survey. 42 students were retained until the second survey. Additionally, as a control condition an equivalent random sample of students ($n = 42$) from the previous cohort were drawn. Students from the experimental conditions were only included in the final analysis if they started the writing process after tutoring, took part in the intervention and both surveys. 7 students in the experimental conditions did not start their writing process and were excluded. The students included ($n = 35$) were randomly assigned to the two conditions: generalized ($n = 17$) and scaffolded recommendations ($n = 18$). Participants were between 24 and 60 years old (M = 39.2; sd = 9.9), predominantly female (90.1%) and had German as a first language (88.2%).

2.6 Measurements

For RQ 1 we calculated the difference between provision of the task and the time sufficient quality was achieved. Additionally, we calculated the difference between the provision of the task and the deadline for starting the writing process. The time students used (time_stud) was then calculated as the percentage value of time students had used of the longest possible time frame to achieve sufficient quality. This accounts for any additional time that the use of the RecSys may have introduced. For RQ 2 we calculated the difference between the time students started tutoring in the forum and the time sufficient quality was achieved. Additionally, we calculated the difference between the time students started tutoring in the forum and the deadline for starting the writing process. The time students were tutored for (time_tut) was calculated as the percentage value of time students were tutored for to achieve sufficient quality compared to the longest possible time frame. This indicates efficiency in terms of workload for teachers and potential time freed in the tutoring process to go into further depth with students. Since participation in the study was optional, we reduced the used questionnaires as much as possible in favor of minimizing workload for students to retain participants. For RQ 3 we administered the F-Komp questionnaire (Hauser et al., 2018) in its German version

(Böttcher-Oschmann et al., 2019). We reduced the questionnaire to subscales relevant to the scenario, which are skills in reviewing the state of research (sr), methodological skills (me) and content knowledge (ck). The F-Komp was administered in both surveys. The difference between the pre- and post-intervention survey was calculated for the global value (δ_FKomp) and all subscales (δ_sr; δ_me; δ_ck) to measure changes in students' perceptions of their research competence. At the time, no validated instrument to measure students´ perceptions of educational RecSys was available (RQ4). For that reason, we adapted a questionnaire for the acceptance of digital technologies (DTAS) to target RecSys (Schorr, 2020). To keep the questionnaire as close to original as possible, we only replaced the term "digital technologies" with "recommender systems." The subscales of the instrument are perceived usefulness (DTAS_uf), behavioral intention to use (DTAS_in), perceived ease of use (DTAS_ea) and attitude towards use. Since the focus of our study was on aspects of the RecSys' on-task functionality, we excluded the last subscale in favor of survey time efficiency. This questionnaire was administered after the intervention.

2.7 Data collection

We gathered forum interaction and survey data from the LMS database. Raw forum data for the given course and survey data was collected through an SQL-query and imported into the data analytics software KNIME (v5.1.3). There, data was filtered for students' interactions with teachers in the tutoring forum, including EPOCH-timestamps for messages. These were filtered for the first submission and the submission, that achieved the necessary quality for starting the writing process. For each student that did not achieve this quality throughout the tutoring phase, the timestamp was set to the deadline for starting writing. Sociodemographic information was provided by the university's statistics department. Data on forum interactions and sociodemographic information for the quasi control group was collected and transformed accordingly. However, no survey data for this historical dataset exists. All data was imported into the statistics software SPSS (v29.0). Here, the differences from between both surveys for the F-Komp as well as the percentage value for time_stud and time_tut were calculated.

3 Data analysis and results

The descriptive statistics and Shapiro–Wilk tests for normal distribution can be seen in Table 1 separated by group, including subscales for each questionnaire. No significant outliers were found. Apart from DTAS_in and time_tut for Group 3, all measurements were normally distributed ($p \geq 0.05$). The adapted scales were tested for their reliability at the time of first application. The F-Komp achieved good ($n = 57$; $\alpha = 0.80$) and the DTAS excellent reliability ($n = 35$; $\alpha = 0.92$). Participation was insufficient to validate the psychometric structure. However, changes were kept to a necessary minimum, and the included subscales were kept intact to minimize any potential bias introduced through these changes (see also section 2.6). Results of the hypothesis testing as described in section 3.1 to 3.4 are presented in Table 2.

3.1 Research question 1

We expected but did not find that students provided with scaffolded recommendations needed shorter time to achieve sufficient quality compared to students with generalized feedback (H1a) and students who are exclusively tutored (H1b). According to the one-sided unpaired t-tests, Group 1 neither outperformed Group 2 ($p = 0.316$) nor Group 3 ($p = 0.290$). Therefore, we reject both hypotheses.

3.2 Research question 2

We expected that students provided with scaffolded recommendations would need shorter tutoring time compared to students with generalized feedback (H2a) and students who are exclusively tutored (H2b). Group 1 (M = 59.97; sd = 20.90) outperformed both group 2 (M = 73.52; sd = 22.20) and group 3 (M = 72.22; sd = 22.95) on the time students were tutored to achieve sufficient quality (time_tut). No correlations between age, gender or first language and time tutored were found, indicating no need to

TABLE 1 Descriptive statistics and Shapiro–Wilk test.

Variable	Group 1			Group 2			Group 3		
	M	sd	*p*	M	sd	*p*	M	sd	*p*
δ_F-Komp	0.07	0.32	-	−0.02	0.25	-			
δ_sr	0.00	0.39	-	0.00	0.44	-			
δ_me	0.19	0.48	-	−0.14	0.70	-			
δ_ck	0.00	0.496	-	0.08	0.52	-			
DTAS_all	4.59	0.93	-	4.09	1.24	-			
DTAS_uf	4.03	1.62	-	3.79	1.33	-			
DTAS_in	4.90	0.96	*	4.15	1.40	-			
DTAS_ea	4.85	1.16	-	4.32	1.37	-			
time_stud	48.38	18.17	-	54.35	16.40	-	51.29	18.71	-
time_tut	59.97	20.90	-	73.52	22.20	-	72.22	22.95	***

*$p < 0.5$; ***$p < 0.001$.

TABLE 2 Results of hypothesis testing.

Hypothesis	Variable	Comparison	*t*-test						Mann–Whitney U test			
			n	df	t	CI	*p*	g	U	Z	*p*	r
H1a)	time_stud	1 < 3	59	58	−0.56	[,]	0.290	-	-	-	-	-
H1b)	"	1 < 2	35	33	−1.02	[,]	0.158	-	-	-	-	-
H2a)	time_tut	1 < 2	35	33	−1.86	[,]	0.036*	−0.615	99.000	−2.079	0.038*	0.351
H2b)	"	1 < 3	59	58	−1.94	[,]	0.028*	−0.541	245.000	−2.145	0.032*	0.279
H3)	*δ_F-Komp*	1 > 2	35	33	0.94	[−0.11, 0.29]	0.177	-	-	-	-	-
	δ_sr	1 > 2	35	33	0.00	[−0.29, 0.29]	0.500	-	-	-	-	-
	δ_me	1 > 2	35	33	1.61	[−0.09, 0.73]	0.059	-	-	-	-	-
	δ_ck	1 > 2	35	33	−0.29	[−0.40, 0.30]	0.388	-	-	-	-	-
H4)	*DTAS_all*	1 > 2	35	33	1.36	[−0.25, 1.25]	0.092	-	-	-	-	-
	DTAS_uf	1 > 2	35	33	1.23	[−0.79, 1.25]	0.322	-	-	-	-	-
	DTAS_in	1 > 2	35	33	1.84	[−0.08, 1.56]	0.038*	0.621	102,500	−1,729	0.084	-
	DTAS_ea	1 > 2	35	33	1.23	[−0.35, 1.40]	0.155	-	-	-	-	-

*$p < 0.5$.

control for covariates. Divergence from normal distribution was found for Group 3 ($p < 0.001$). We performed a one-sided unpaired t-test for both hypotheses, since it is robust against this. Levene-tests revealed no difference in variance. Group 1 needed less tutoring time than Group 2 ($p = 0.036$) and Group 3 ($p = 0.028$). Estimation of the effect size (Hedge's g) showed moderate effects when compared to both Group 2 (g = −0.615) and Group 3 (g = −0.541), indicating that scaffolded recommendations reduced teacher workload by reducing the time students are tutored to achieve sufficient quality. However, one-sided t-tests are controversial in their interpretability and inflate chances for false positives (Trafimow, 2023). Therefore, we explored the data further by performing the non-parametric Mann Whitney U test for both H2a and H2b, which is robust with small groups. The test revealed significant differences for both H2a ($p = 0.038$, r = 0.351) and H2b ($p = 0.032$, r = 0.279) with Group 1 outperforming both groups, confirming initial results.

3.3 Research question 3

We expected but did not find that students' perception of their own research competence increased if they received scaffolded recommendations compared to generalized recommendations (H3). Levene-Tests revealed no difference in variance for any variable ($p > 0.05$). The one-sided unpaired t-test revealed no significant differences between Group 1 and 2 for δ_FKomp ($p = 0.177$), for δ_sr ($p = 0.500$), for δ_me ($p = 0.059$) or for δ_ck ($p = 0.388$). Therefore, we reject the hypothesis that students' perception of their own research competence increased significantly if they received scaffolded recommendations compared to generalized recommendations.

3.4 Research question 4

We expected that students provided with scaffolded recommendations would perceive a higher ability of the RecSys to support them than students provided with generalized recommendations (H4). To analyze these differences, a two-sided unpaired t-test was performed for the global DTAS value and all subscales. Levene-tests revealed no difference in variance for any variable (p > 0.05). The one-sided unpaired t-test revealed no significantly higher value for Group 1 compared to Group 2 for DTAS_all ($p = 0.092$), for DTAS_uf ($p = 0.322$) or for DTAS_ua ($p = 0.115$). However, a significantly higher value for DTAS_in was found ($p = 0.038$). Estimation of the effect size revealed a moderate effect (g = 0.621). This indicates scaffolded recommendations having a positive effect on students' intention to use such a system in the future. The same issues with interpretability as in RQ 2 are applicable. Accordingly, the Mann–Whitney U test was used, revealing no significant differences between Group 1 and 2 ($p = 0.084$), contradicting initial results.

4 Discussion

The goal of this study was to investigate the feasibility of scaffolding recommendations to support students in essay planning. Our results indicate the potential of facilitating essay planning through recommendations. Although students do not improve the quality of their essays significantly faster, the time needed for tutoring is significantly reduced. This is supported by research on scaffolding effectiveness in online learning (Doo et al., 2020). The strict scaffolding in the experimental condition resulted in time efficiency gains. This may be rooted in the fact that students inexperienced with essay planning have a high need to support in acquiring necessary skills. In the asynchronous tutoring format, this experience gap was closed earlier by the given RecSys, matching students' ability to perform the task, effectively reaching ZPD prior to tutoring. This is supported by the claim that adaptive scaffolding often outperforms other forms of scaffolding (Azevedo et al., 2005), allowing the scaffolded recommendations to more closely adapt to the student's current competence level in essay planning. Additionally descriptive statistics show the scaffolded recommendations condition outperforming other conditions on nearly all variables (see Table 1), whilst consistently receiving above average ratings for satisfaction with the RecSys (DTAS) and showing a small positive trend for

research competence (F-Komp). While not all differences are statistically significant, this may indicate an overall positive trend for the proposed method. This is in line with prior evidence supporting the effects of scaffolding in general (Azevedo et al., 2005; Doo et al., 2020; Panadero et al., 2016) and for other settings using automatic writing support (Fathi and Rahimi, 2024; Kim et al., 2022; Rapp and Kauf, 2018). Therefore, RecSys implementing HIF to provide scaffolded recommendations on students' essay planning can complement personal tutoring and provide flexible additional support. This seems to reduce workload for tutors while allowing students for more autonomy in the planning process. However, the interpretability of the results is somewhat limited. While directional tests revealed significant effects, more robust tests revealed no significant difference for the intention to reuse the system, indicating that the satisfaction with the system may not be dependent on the method through which recommendations are provided. Also, due to the skewed demographics and the specific domain of media education, the results should not be generalized. No correlations between demographic information and the chosen measurements were found, but this may differ with larger sample size or in other educational settings.

5 Conclusion and implications

In summary, the results provide novel insights into the potential of scaffolding as a driver of positive effects of recommendations for students' essay planning and provide evidence for its feasibility. The implemented RecSys has led to shorter tutoring time, a high satisfaction with the system and an overall positive trend for measured variables in favour of scaffolding recommendations. It may be interesting to analyse, whether these results are a consequence of raising the quality of initial contributions. Therefore, we aim to analyse the quality of students' essay concepts from the experiment that were submitted to the RecSys before tutoring with those that were submitted to the tutoring process afterwards. Additionally, when an recommender engine is applied, the response time of the system is reduced through automation. The immediacy of the RecSys could allow students to independently work on their submissions faster compared to the conditions in the simulation. This may in turn facilitate the effects of guiding students towards a ZPD earlier. In addition, broadening the scope of the experimental setting to other domains and beyond distance education would provide clarify whether the effects are retained in different settings.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for the studies involving humans because research participation and use of the system was entirely voluntary and participation as well as use of the system was only disclosed to researchers in pseudonymized form. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing, Software. SW: Conceptualization, Data curation, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. ND-T: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Investigation, Software. XW: Conceptualization, Data curation, Investigation, Software, Writing – original draft, Writing – review & editing. AJ: Data curation, Software, Visualization, Writing – original draft, Writing – review & editing. CW: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. NP: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research is part of the AI.EDU Research Lab 2.0 project led by Claudia de Witt and Prof. Niels Pinkwart and funded by the Center of Advanced Technology for Assisted Learning and Predictive Analytics (CATALPA) at the FernUniversität in Hagen.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alley, M. (2018). The craft of scientific writing. New York, NY: Springer New York.

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

Böttcher-Oschmann, F., Groß Ophoff, J., and Thiel, F. (2019). Validierung eines Fragenbogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen – Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements. *Unterrichtswiss* 47, 495–521. doi: 10.1007/s42010-019-00053-8

Bulquiyah, S., Mahbub, M. A., and Nugraheni, D. A. (2021). Investigating writing difficulties in essay writing: tertiary students' perspectives. *English Lang. Teach. Educ. J.* 4, 61–73. doi: 10.12928/eltej.v4i1.2371

da Silva, F. L., Slodkowski, B. K., da Silva, K. K. A., and Cazella, S. C. (2023). A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities. *Educ. Inf. Technol.* 28, 3289–3328. doi: 10.1007/s10639-022-11341-9

Dahl, B. M., Vasset, F., and Frilund, M. (2023). Students' approaches to scientific essay writing as an educational method in higher education: a mixed methods study. *Soc. Sci. Humanit. Open* 7:100389. doi: 10.1016/j.ssaho.2022.100389

Doo, M. Y., Bonk, C., and Heo, H. (2020). A meta-analysis of scaffolding effects in online learning in higher education. *Int. Rev. Res. Open Dist. Learn.* 21:1 21. doi: 10.19173/irrodl.v21i3.4638

Drachsler, H., Verbert, K., Santos, O. C., and Manouselis, N. (2015). "Panorama of recommender systems to support learning" in *Recommender systems handbook*, eds. F. Ricci, L. Rokach and B. Shapira (Boston, MA: Springer US).

Fathi, J., and Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Comput. Assist. Lang. Learn.* 1–40, 1–40. doi: 10.1080/09588221.2024.2374772

Fleckenstein, J., Liebenow, L. W., and Meyer, J. (2023). Automated feedback and writing: a multi-level meta-analysis of effects on students' performance. *Front. Artif. Intellig.* 6:1162454. doi: 10.3389/frai.2023.1162454

Gloerfeld, C., Wrede, S., de Witt, C., and Wang, X. (2020). Recommender – potentials and limitations for self-study in higher education from an educational science perspective. *iJAI* 2:34. doi: 10.3991/ijai.v2i2.14763

Guilford, W. H. (2001). Teaching peer review and the process of scientific writing. *Adv. Physiol. Educ.* 25, 167–175. doi: 10.1152/advances.2001.25.3.167

Halliday, M. A. K., and Martin, J. R. (2003). *Writing science: Literacy and discursive power*. Routledge.

Haswell, R. H. (2000). Documenting improvement in college writing: a longitudinal approach. *Writ. Commun.* 17, 307–352. doi: 10.1177/074108830001700301

Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi: 10.3102/003465430298487

Hauser, F., Reuter, R., Gruber, H., and Mottok, J. (2018). Research competence: Modification of a questionnaire to measure research competence at universities of applied scienes, in Proceedings of 2018 IEEE Global Engineering Education Conference (EDUCON): Date and venue: 17–20 April, 2018, Santa Cruz de Tenerife, Canary Islands, Spain (Piscataway, NJ: IEEE). pp. 109–117.

Hmelo-Silver, C. E., Duncan, R. G., and Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). *Educ. Psychol.* 42, 99–107. doi: 10.1080/00461520701263368

Hounsell, D. (1984). Essay planning and essay writing. *High. Educ. Res. Dev.* 3, 13–31. doi: 10.1080/0729436840030102

Jafarigohar, M., and Mortazavi, M. (2017). The impact of scaffolding mechanisms on EFL learners' individual and socially shared metacognition in writing. *Read. Writ. Q.* 33, 211–225. doi: 10.1080/10573569.2016.1154488

Joy, J., and Pillai, R. (2022). Review and classification of content recommenders in E-learning environment, *Journal of King Saud University - Computer and Information Sciences*, 34, 7670–7685. doi: 10.1016/j.jksuci.2021.06.009

Kim, M. K., Kim, N. J., and Heidari, A. (2022). Learner experience in artificial intelligence-scaffolded argumentation. *Assess. Eval. High. Educ.* 47, 1301–1316. doi: 10.1080/02602938.2022.2042792

Kirschner, P. A., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86. doi: 10.1207/s15326985ep4102_1

Kreutz, C. K., and Schenkel, R. (2022). Scientific paper recommendation systems: a literature review of recent publications. *Int. J. Digit. Libr.* 23, 335–369. doi: 10.1007/s00799-022-00339-w

Lee, C. C., and Tan, S. C. (2010). Scaffolding writing using feedback in students' graphic organizers – novice writers' relevance of ideas and cognitive loads. *Educ. Media Int.* 47, 135–152. doi: 10.1080/09523987.2010.492678

Li, T., Fan, Y., Tan, Y., Wang, Y., Singh, S., Li, X., et al. (2023). Analytics of self-regulated learning scaffolding: effects on learning processes. *Front. Psychol.* 14:1206696. doi: 10.3389/fpsyg.2023.1206696

Lin, T.-J., Horng, R.-Y., and Anderson, R. C. (2014). Effects of argument scaffolding and source credibility on science text comprehension. *J. Exp. Educ.* 82, 264–282. doi: 10.1080/00220973.2013.769423

Lo, J.-J., Chen, Y.-R., and Yeh, S.-W. (2009). Scaffolding Support Mechanism for Writing Cause and Effect Essays, in International Conference on Computational Intelligence and Software Engineering: CiSE 2009; Wuhan, China, 11–13 December 2009 (Piscataway, NJ: IEEE). pp. 1–4.

Maphosa, V., and Maphosa, M. (2023). Fifteen years of recommender systems research in higher education: current trends and future direction. *Appl. Artif. Intell.* 37:2175106. doi: 10.1080/08839514.2023.2175106

Oppenheimer, D., Zaromb, F., Pomerantz, J. R., Williams, J. C., and Park, Y. S. (2017). Improvement of writing skills during college: a multi-year cross-sectional and longitudinal study of undergraduate writing performance. *Assess. Writ.* 32, 12–27. doi: 10.1016/j.asw.2016.11.001

Ouyang, F., Zhang, N., Chen, Y., and Shao, X. (2024). The use and effect of cognitive scaffoldings on facilitating collaborative discussion-supported writing. *Technol. Pedagog. Educ.* 34, 275–295. doi: 10.1080/1475939X.2024.2437036

Panadero, E., Jonsson, A., and Strijbos, J.-W. (2016). "Scaffolding self-regulated learning through self-assessment and peer assessment: guidelines for classroom implementation" in Assessment for learning: Meeting the challenge of implementation. eds. D. Laveault and L. Allal (Cham: Springer International Publishing), 311–326.

Rapp, C., and Kauf, P. (2018). Scaling academic writing instruction: evaluation of a scaffolding tool (thesis writer). *Int. J. Artif. Intell. Educ.* 28, 590–615. doi: 10.1007/s40593-017-0162-z

Ricci, F., Rokach, L., and Shapira, B. (2011). "Introduction to Recommender Systems Handbook," in Recommender Systems Handbook, ed. B. Shapira (New York, NY: Springer), 1–35.

Schorr, A. (2020). (GfA Frühjahrskongress 2020). "Skala zur Erfassung der Digitalen Technologieakzeptanz – Weiterentwicklung zum testtheoretisch geprüften Instrument,". Beitrag B.20.5. Digitaler Wandel, digitale Arbeit, digitaler Mensch? Frühjahrskongress der Gesellschaft für Arbeitswissenschaft. Available online at: https://gfa2020.gesellschaft-fuer-arbeitswissenschaft.de/inhalt/B.20.5.pdf (Accessed February 26, 2024).

Singh, V., and Mayer, P. (2014). Scientific writing: strategies and tools for students and advisors. *Biochem. Mol. Biol. Educ.* 42, 405–413. doi: 10.1002/bmb.20815

Sun, L., Tao, S., Hu, J., and Dow, S. P. (2024). MetaWriter: Exploring the Potential and Perils of AI Writing Support in Scientific Peer Review. *Proc. ACM Hum.-Comput. Interact.* 8:CSCW1. doi: 10.1145/3637371

Trafimow, D. (2023). The power of directional predictions in psychology. *J. Theory Soc. Behav.* 53, 62–84. doi: 10.1111/jtsb.12343

Ucar-Longford, B., Hosein, A., and Heron, M. (2024). The use of online scaffolding to develop argumentation skills: a scoping review. *Technol. Pedagog. Educ.* 33, 437–453. doi: 10.1080/1475939X.2024.2352040

van de Pol, J., Volman, M., and Beishuizen, J. (2010). Scaffolding in teacher–student interaction: a decade of research. *Educ. Psychol. Rev.* 22, 271–296. doi: 10.1007/s10648-010-9127-6

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Massachusetts: Harvard University Press.

Wisniewski, B., Zierer, K., and Hattie, J. (2020). The power of feedback revisited: a meta-analysis of educational feedback research. *Front. Psychol.* 10:3087. doi: 10.3389/fpsyg.2019.03087

Wood, D., Bruner, J. S., and Ross, G. (1976). The role of tutoring in problem solving. *J. Child Psychol. Psychiatry* 17, 89–100. doi: 10.1111/j.1469-7610.1976.tb00381.x

Zhang, Z., Patra, B. G., Yaseen, A., Zhu, J., Sabharwal, R., Roberts, K., et al. (2023). Scholarly recommendation systems: a literature survey. *Knowl. Inf. Syst.* 65, 4433–4478. doi: 10.1007/s10115-023-01901-x