Complementing Educational Recommender Systems with Open Learner Models

Solmaz Abdi
The University of Queensland
Brisbane, QLD, Australia
s.abdi@uq.edu.au

Hassan Khosravi
The University of Queensland
Brisbane, QLD, Australia
h.khosravi@uq.edu.au

Shazia Sadiq
The University of Queensland
Brisbane, QLD, Australia
shazia@itee.uq.edu.au

Dragan Gasevic
Monash University
Clayton, Vic, Australia
Dragan.Gasevic@monash.edu

ABSTRACT

Educational recommender systems (ERSs) aim to adaptively recommend a broad range of personalised resources and activities to students that will most meet their learning needs. Commonly, ERSs operate as a “black box" and give students no insight into the rationale of their choice. Recent contributions from the learning analytics and educational data mining communities have emphasised the importance of transparent, understandable and open learner models (OLMs) that provide insight and enhance learners' understanding of interactions with learning environments. In this paper, we aim to investigate the impact of complementing ERSs with transparent and understandable OLMs that provide justification for their recommendations. We conduct a randomised control trial experiment using an ERS with two interfaces (“Non-Complemented Interface" and “Complemented Interface") to determine the effect of our approach on student engagement and their perception of the effectiveness of the ERS. Overall, our results suggest that complementing an ERS with an OLM can have a positive effect on student engagement and their perception about the effectiveness of the system despite potentially making the system harder to navigate. In some cases, complementing an ERS with an OLM has the negative consequence of decreasing engagement, understandability and sense of fairness.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **User models**; **Visualization application domains**.

KEYWORDS

Educational Recommender Systems, Open Learner Models

LAK '20, March 23–27, 2020, Frankfurt, Germany

ACM ISBN 978-1-4503-7712-6/20/03…$15.00
https://doi.org/10.1145/3375462.3375520

1 INTRODUCTION

Recommender systems are one of the most common applications of big data, personalising many aspects of our lives [22]. In the educational setting, researchers have tailored exemplary techniques from recommender systems to adaptively recommend a broad range of resources such as learning activities and videos to students that will best meet their learning needs [20]. To recommend personalised learning resources to students, educational recommender systems (ERSs) often rely on a component–commonly referred to as a learner model– which plays an important role in many educational systems. Learner models capture an abstract representation of a student's ability and behaviour based on their performance and interactions with an educational system [7].

Commonly in ERSs, the learner model is not shared with the students. Therefore, from a student's perspective, the recommender system operates as a “black-box”, giving them no insight into the rationale of their choice [25]. Black-box recommendations, which are commonly used outside of education as well, can lead to trust issues, especially when users do not agree with the recommendations [11, 29]. The use of black-box recommendations seems to be particularly inadequate for educational settings where educators strive to enable students to develop their own vision, reasoning, and appreciation for inquiry and investigation.

Development of strategies and approaches that assist students in better understanding of how their learning is captured and approximated in educational systems has been studied in a field–commonly referred to Open Learner Models (OLMs) [4]. Simply put, OLMs are learner models that are externalised and made accessible to students or other stakeholders such as instructors. They are often opened through visualisations, as an important means of supporting learning [4]. OLMs have been integrated into a variety of educational tools such as learning analytics dashboards [2, 3], intelligent tutoring systems [23, 24], and adaptive learning platforms [8] to help students and instructors in monitoring, reflecting and regulating learning [4]. However, to date, empirical evaluation of their effectiveness has received little attention [12].

This paper aims to investigate the impact of complementing ERSs with transparent and understandable OLMs that provide justification for their recommendations. The following research questions guide our investigation:

- *RQ1.* How does complementing an educational recommender system with the visualisation of an open learner model impact student engagement with the system?
- *RQ2.* How does complementing an educational recommender system with the visualisation of an open learner model impact students' perceived effectiveness of the system?

To answer these research questions, we conducted a between-subject, randomised controlled study in an on-campus course at The University of Queensland using a course-level ERS called RiPPLE which uses a learner model for recommending resources. Two interfaces of RiPPLE ("Non-Complemented Interface" and "Complemented Interface") were used in which only the complementary OLM feature differed between the two versions. Data from participants' survey results and interaction logs with RiPPLE as well as insights from additional 3000 RiPPLE users, collected through discussion forums, survey results and feedback forms, are used in our analysis. Our results suggest that complementing an ERS with an OLM improves student engagement, as well as improving students' understanding and trust in recommendations and enhances their acceptance of the system. However, this may come at the cost of making the system harder to navigate. In some cases, there may also be other negative consequences from complementing an ERS with an OLM including (1) reducing understandability of the system due to a lack of understanding of the OLM; (2) increasing a sense of unfairness due to students not agreeing with how the OLM models their competency; and (3) decreasing engagement due to student concerns about "risking a hit to their rating", as represented in the OLM, in case they answer questions incorrectly.

2 RELATED WORK

In this section, we position our work with regards to previous research that has focused on providing transparency in recommender systems and learner models. In both cases, we discuss the (1) context in which they have been applied (2) proposed approaches for providing transparency (3) experimental designs used for their evaluation and (4) general reported findings from these fields.

Transparent recommender systems. In recent years, the research on the importance of addressing the "black-box" issue of recommender systems has been widely studied in different contexts [31] such as critiquing-based recommender systems [30], music and movie recommender systems [11, 25], art recommender systems [6] and social recommender systems [9]. Previous studies have attempted to expose the reasoning behind recommendations with various approaches ranging from textual interfaces to explain "why" a certain item is recommended [6] and simple icon-based visual interfaces [30] to more complicated interactive visual interfaces [9, 29]. A wide range of evaluation methods including controlled trials studies based on users' interaction log data with the recommender system (e.g. [6, 30]) and survey questionnaire (e.g. [11, 25]) have been investigating the impact of these systems. Findings from previous works in this field suggest that exposing the reasoning behind recommendations can help to produce a higher recommendation accuracy, gives users a sense of control over the recommendation process, and improves users' satisfaction and trust in recommendation and enhances the acceptance of the system [9, 11, 25].

Open Learner Models (OLMs). OLMs have been studied predominantly in the educational data mining (EDM), computer-supported collaborative learning (CSCL) and learning analytics communities. To date, they have mostly been embedded in learning analytics dashboards [2, 3], intelligent tutoring systems [23, 24], and adaptive learning platforms [8]. Typically, a variety of approaches are used to visualise the findings of OLMs to students ranging from simple skill-meters [28] to complex graphs such as concept maps [17, 19], and hierarchical tree structures [5]. The previous studies on OLMs have either presented their models conceptually (e.g. [5, 13]) or have used qualitative analysis based on focus groups and surveys (e.g. [3, 19]) and quantitative analysis (e.g. [17, 18]) to investigate the benefits of the provided transparency by OLMs. However, only a very few of them, such as [18] have conducted empirical studies using randomised controlled trials. The findings of these studies have framed the benefits of OLMs mostly around promoting metacognitive activities (e.g. reflection, planning and self-regulation), allowing the student to take greater control over their learning, improve the accuracy of the learner model, and increasing student's trust in the educational system [3–5, 12, 18].

3 METHOD

In what follows, Section 3.1 introduces the system called RiPPLE[1] followed by Section 3.2 that provides a description for each of the two interfaces of RiPPLE used by this study. Finally, Section 3.3 presents the experimental design used for this study[2].

3.1 RiPPLE

This study uses a course-level, discipline-agnostic adaptive learning system called RiPPLE that uses crowdsourcing for development of learning activities [16]. RiPPLE utilises crowdsourcing to enable students to create learning activities and complete, rate and discuss learning activities created by their peers [16]. In its current version, RiPPLE generates a learner model based on Elo rating system to estimate a student's knowledge state [1] based on their performance on the learning activities available in the repository of the platform. This model is used by the recommender system of RiPPLE to recommend new personalised learning activities to students tailored towards their learning needs. For a detailed description of RiPPLE please see [14].

3.2 Two Experiment Interfaces

This section provides an overview of the two interfaces that are used for the study. The interfaces relate to one of the pages of RiPPLE, called "VIEW AND RESPOND" page which enables students to select learning activities using search and recommendation functionalities. A detailed description of the design choices made in the development of RiPPLE is provided in [16] and it is outside of the scope of this paper. Here, we only discuss the principles and design choices that were considered for the addition of the OLM, which are discussed in this section when we present the Complemented Interface.

[1] http://ripplelearning.org/
[2] Approval from our Human Research Ethics Committee #2018000125 was received for conducting this experiment

Interface 1: Non-Complemented Interface Figure 1 shows "VIEW AND RESPOND" page of RiPPLE based on the Non-Complemented Interface. A set of filters that are available to help students search the resource repository are demonstrated at the top of the figure. The "Sort By" option allows students to sort the returned resources based on their difficulty, quality, number of responses, or personal fit ("Recommended"). By selecting "Recommended", the platform sorts the resources based on their learning benefits to the student. The "Filter" option enables students to filter the learning activities that are included in the results. They can request all learning activities (default), uncompleted learning activities, completed learning activities, or wrongly answered learning activities to be included in the results. The "Search" option enables students to search for learning activities based on specific content that may be present in the learning activities.

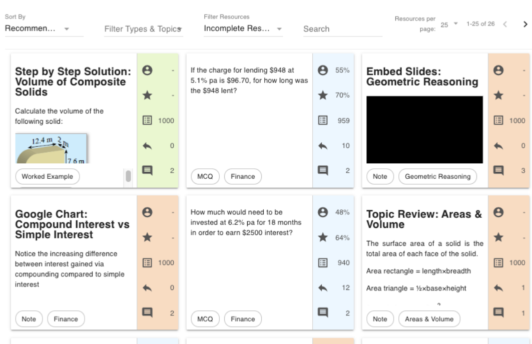

Figure 1: Overview of "VIEW AND RESPOND" page of RiPPLE in the Non-Complemented Interface.

The results of the search are presented as a list of learning activity cards, allowing students to engage with learning activities that best suit their needs. Each learning resource card includes an overview of the learning activity content, the knowledge units (concepts) associated with the learning activity, and a sidebar demonstrating some additional information about the learning activity.

Interface 2: Complemented Interface Figure 2 shows the "VIEW AND RESPOND" page of RiPPLE based on the second interface, in which recommendations are complemented with the visualisations of learner model. The top section of this page provides an interactive visualisation widget that enables students to view an abstract representation of their knowledge state based on the knowledge units (concepts) that are present in the domain model. The proposed OLM was designed based on the following two principles: (1) The OLM and the recommendation results need to be placed close to one another on the interface. This was to ensure that the effect of complementing the ERS with an OLM is maximised. (2) The utilised visualisation must be easy to understand by the majority of the users. A range of visualisations – including bar charts, line charts that demonstrate progress over time, zoomable Treemap, and Topic Dependency Models [15] – have been incorporated and tested by earlier versions of RiPPLE. Based on the results of multiple rounds of usability tests, bar charts and line charts have been adopted in the latest version of the system. The"Visualisation Data" drop-down enables students to select between two models visualising their knowledge state: viewing their current knowledge state or track changes to their knowledge state over time.

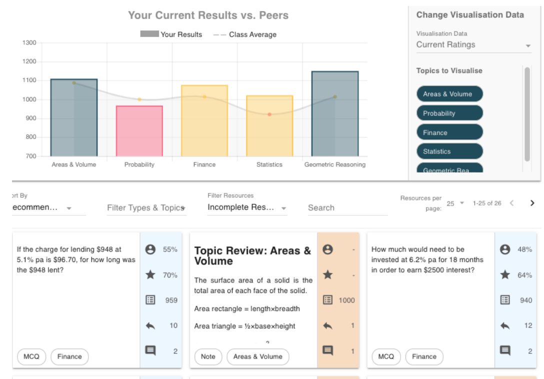

Figure 2: Overview of the "VIEW AND RESPOND" page of RiPPLE in the Complemented Interface.

Colour of the bars, determined by the underlying algorithm modelling the student, categorises competencies into three levels: red demonstrates inadequate competency in a knowledge unit, yellow demonstrates adequate competency with room for improvement, and blue demonstrates mastery in a knowledge unit. The model also shows the average competency of the entire cohort over each knowledge unit using a line graph. The "Topic to Visualise" option enables students to select the knowledge units that they would like to be included in the visualisation. By default, all of the knowledge units are selected. Upon attempting a learning activity by a student, their knowledge state is updated by the learner model, which is also reflected on the visualisation widget on a real-time basis. The bottom section of the interface is exactly the same as the Non-Complemented Interface.

3.3 Experimental Design

For this study, RiPPLE was used in an on-campus undergraduate course at The University of Queensland for five weeks. During this period, the 50 students that participated in this study, made 2,534 attempts on 280 learning activities which were available in the platform repository for this course. The experiment used a between-subject design where participants were randomly assigned to one of the two interfaces. A randomised control trial was chosen because of the following two reasons, which were earlier explored and discussed in Section 2: (1) Many of the recommender systems that have attempted to provide insights on their recommendations have used this experimental setting for their evaluation (2) While OLMs have been widely studied, there are very limited studies that have used randomised controlled trials for evaluating the effect of OLMs. The proposed experimental setting seems to hold potential for investigating our proposed research questions while providing novel empirical contribution on OLMs.

Ethical Considerations. There have been fiery debates about the opportunities and challenges of using randomised controlled trials in education [26, 27]. While they remain a gold standard test for establishing causality in many fields, in the educational setting, they are often subject to threats to unethically disadvantaging the learning opportunities of students in one of the experimental groups. The following steps were taken to seek consent from students and minimise the risk of harming students' learning outcomes in this study: (1) upon the first use of the platform, users were presented with a consent form seeking their permission for running educational experiments and using their data. All users regardless of their response can use RiPPLE and only data collected from users that have provided consent were used for the study; (2) the use of the platform was not tied to students' assessment in any way; (3) the scope of the study was designed around engagement and perception about the platform rather than learning gains; and finally, (4) the duration of the study was limited to 5 weeks instead of the entire semester.

Impact on Engagement. For investigating RQ1, we concentrated on the following objective measures: the average number of learning activities that are attempted by students, the total session time on the platform and the average time of each recommendation session of the users that were participating in the study. These measures were obtained from logs of students' interaction with the system. In each group, students who had attempted less than five learning activities were considered inactive and were excluded from further analysis, leaving 13 students in the non-complemented condition and 13 students in the complemented condition. Results of this investigation are reported in Section 4.1.

Impact on Students' Perceptions. For investigating RQ2, we concentrated on perceived measures by conducting a survey asking participants for feedback related to their understanding and trust in recommendations, their perception about the usability of the system as well as the acceptance of the system using the 5 statements listed in table 1. The survey was administered on paper at the beginning of a face-to-face session where students had the option of completing the survey. The survey was conducted anonymously, so, to identify students' experimental condition, at the beginning of the survey the screenshot of the both "View and RESPOND" interfaces were presented to the students, and they were asked to select the RiPPLE interface that was accessible to them during the study. However, because of this anonymity, it was not possible to exclude students based on the number of activities they had attempted. Instead, in the survey, the students were asked to determine their level of engagement with RiPPLE throughout the study from the following options: less than one hour, 1 to 4 hours, 4 to 8 hours, 8 to 12 hours, and more than 12 hours. Based on student responses to this question, responses from students who had spent less than one hour in RiPPLE where excluded from further evaluations, leaving 22 students in the non-complemented condition and 19 students in the complemented condition. Responses were captured using a five-point Likert-scale were 1 represents strongly disagree and 5 represents strongly agree. We used a Mann-Whitney test to perform statistical analysis of the reported results. Considering the small sample size of this study, we use 0.1 as the criterion for assessing statistical significance. Results of this investigation are reported in Section 4.2.

Additional Findings and Lessons Learned. We also reflect on findings and lessons learned from running the complemented version of the interface of RiPPLE in 15 courses with over 3,000 registered users. These insights are derived from data collected through surveys conducted in some of these courses, completion of the feedback form which is available on RiPPLE, discussion forum comments, as well as focus groups and usability studies that have been conducted by the developers of RiPPLE. These insights are reported in Section 4.3.

Table 1: Survey statements

S1	I understand the rationale behind recommendations.
S2	The method used by RiPPLE in estimating my knowledge state is fair.
S3	RiPPLE correctly adapts its recommendations on the basis of my current learning needs.
S4	The RiPPLE platform is easy to navigate.
S5	I would like to use RiPPLE in my other courses.

Table 2: Average of Students' engagement with each of the two RiPPLE interfaces (*Mean* ± *SD*)

Interface	# of attempted activities	Total sessions time (min)	Average of each session time (min)
Non-Complemented	30 ± 24	94 ± 62	20 ± 14
Complemented	37.46 ± 28	127 ± 82	25 ± 11

4 RESULTS AND DISCUSSION

This section investigates and answers the two research questions that were introduced in Section 1. In what follows, Section 4.1 reports our findings on the impact of complementing an ERS with an OLM on student engagement. Section 4.2 reports our findings on the impact of complementing an ERS with an OLM on students' perception of the effectiveness of the system. Finally, Section 4.3 shares additional insights which have been obtained from using the complementing version of RiPPLE in 15 courses with over 3,000 users.

4.1 Impact on Engagement

Table 2 reports the engagement results obtained from the two RiPPLE interfaces. With regards to the number of attempted learning activities, the reported results indicate that on average, students who used the Complemented Interface attempted 23% more learning activities than students who used the Non-Complemented Interface (U = 70.5, Mean'Complemented' = 37.14±28, Mdn'Complemented' = 35, Mean'Non-Complemented' = 30 ± 24, Mdn'Non-Complemented' = 23, p>0.1). With regards to the total session time, the reported results indicate that, the average total session time for students who used the Complemented Interface was 32 minutes higher than the total session time spent by students using the Non-Complemented Interface (U = 72, Mean'Complemented' = 127±82, Mdn'Complemented' = 98, Mean'Non-Complemented' = 94±62, Mdn'Non-Complemented' = 92, p>0.1). In addition, the average of each session time for students who used the Complemented Interface was approximately 25% higher than the students who used the Non-Complemented Interface (U = 59, Mean'Complemented' = 25±11, Mdn'Complemented' = 29, Mean'Non-Complemented' = 20±14, Mdn'Non-complemented' = 13, p<0.1).

The reported results suggest that complementing an ERS with an OLM can increase student engagement. It is particularly worth noting that students who used the Complemented Interface showed

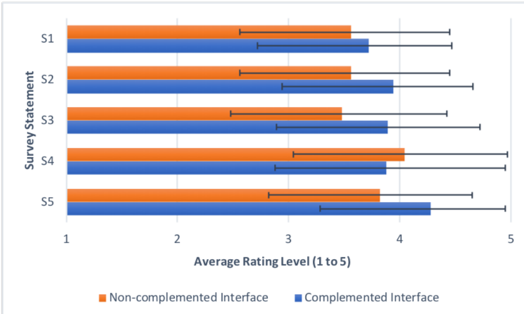

Figure 3: Survey results for both interfaces.

a significantly higher average session time compared to students who used the Non-Complemented Interface. The increase in the number of attempted activities and the total session time seem to demonstrate practical significance, providing further evidence for an increase in engagement. However, this increase in engagement was not statistically significant, which may be due to the small sample size of the experiment.

Based on the existing literature on OLMs, it is possible to speculate that the additional time spent on RiPPLE by users in the Complemented Interface was towards self-monitoring, self-regulation and reflecting on the visualisations of OLM. However, further experiments are required to validate these interpretations. An interesting direction in the future work would be to use eye-tracking devices to run a more comprehensive study to investigate the difference in the behaviour of users in both versions of the interface.

4.2 Impact on Students' Perceptions

Figure 3 represents the results of the survey, discussed in Section 3.3, for the Non-Complemented Interface and Complemented Interface. With regards to "understanding" (*S*1), the students who used the Complemented Interface reported a higher understanding of the rationale behind recommendations compared to the students who used the Non-Complemented Interface (U = 181, Mean'Complemented' = 3.72±0.75, Mdn'Complemented = 4, Mean'Non-Complemented' = 3.56 ± 0.89, Mdn'Non-Complemented' = 3, p>0.1). In terms of "Fairness" (*S*2), the students who used the Complemented Interface reported a statistically significant higher confidence in the fairness of the method used by RiPPLE in estimating their knowledge state compared to the students who used the Non-Complemented Interface (U = 155.5, Mean'Complemented' = 3.94 ± 0.72, Mdn'Complemented = 4, Mean'Non-Complemented' = 3.56 ± 0.79, Mdn'Non-Complemented' = 3, p<0.1). Furthermore, the students who used the Complemented Interface reported a statistically significant higher confidence in RiPPLE to "correctly adapt" (*S*3) its recommendations on the basis of their current learning needs compared to the students who used the Non-Complemented Interface (U = 156.5, Mean'Complemented' = 3.88 ± 0.83, Mdn'Complemented = 4, Mean'Non-Complemented' = 3.47±0.94, Mdn'Non-Complemented' = 4, p<0.1). Interestingly, with regards to the "ease of navigation" (*S*4), the students who used the Non-Complemented Interface perceived RiPPLE easier to navigate compared to the students who used the Complemented Interface (U = 194, Mean'Complemented' = 3.83 ± 1.04, Mdn'Complemented = 4, Mean'Non-Complemented' = 4.04 ± 0.93, Mdn'Non-Complemented' = 4, p>0.1). Finally, in terms of acceptance of the system (*S*5), the students who used the Complemented Interface reported statistically significant higher inclination towards using RiPPLE in their other courses compared to the students who used the Non-Complemented Interface (U = 144.5, Mean'Complemented' = 4.27 ± 0.67, Mdn'Complemented = 4, Mean'Non-Complemented' = 3.83 ± 0.83, Mdn'Non-Complemented' = 4, p<0.1).

The survey results from both groups indicate that in general, the students had a positive attitude towards RiPPLE and that complementing RiPPLE with an OLM has an overall positive effect on the students' perceived effectiveness of the system. In general, our results are aligned with findings from previous studies on the impact of using a visual interface for recommender systems in enhancing users' understanding of the reasoning behind recommendations and trust in them [9, 30]. In particular, findings from this section coupled with findings from Section 4.1 suggest that accurate recommendations are not enough to make a system acceptable to users; other human factors such as transparency, understanding, fairness and trust have a paramount effect on users' decision to further use the system [6, 10, 21].

We posit that the reason why the students found the Complemented Interface harder to navigate is because of the additional complexity involved in understanding the OLM itself, which requires some learning efforts from students to learn how to work with it. Embedding explanatory examples or detailed instructions into the system, as suggested by [30] may assist students with the navigation and increase the overall usability of the system.

4.3 Additional Findings and Lessons Learned

By and large, the students provided positive feedback about the OLM integrated into RiPPLE. However, we also encountered some challenges and drawbacks regarding the addition of an OLM to RiPPLE. We have shared these challenges and potential ways of addressing them below.

(1) The addition of the OLM has generally helped students better understand the rationale behind recommendations. However, this addition has changed the problem of understanding the rationale behind the recommendations to two new problems for some students: (1) what is the OLM precisely showing and (2) how is my competency calculated by the OLM? Our recent attempts in creating videos and hints on the platform for further describing how the OLM works seem to help with this issue.

(2) By and large, the students feel that their competency is fairly and accurately represented in RiPPLE. However, some students have raised concerns about the fairness of the OLMs. Commonly, these concerns have been raised under the following two conditions: (1) a student with high competency rating on a knowledge unit answering an easy question (low difficulty rating) incorrectly. Based on the Elo-based OLM which is used in RiPPLE [1], this leads to a significant drop in the competency rating of the student; (2) a student answering a question which they believe was a poor question incorrectly. The following quote provides an example of a student raising their disagreement. "I'm finding that 20-30 minutes of considered question answering, to try and achieve a particular

competency, can be negated by one dodgy question... I would liken it to the experience of a child dropping their ice cream on the floor, only with valuable time (and marks) in place of the ice cream“. One potential solution is to make the visualisation widget more interactive and allow students to negotiate with the system about their estimated knowledge states as proposed by [3]. This approach is also closely related to the existing literature on the importance of incorporating users' input and feedback into the recommendation process to improve recommendations [10].

(3) Most students have reported that the availability of the OLM has increased their engagement with RiPPLE as it motivates them to improve their competency. However, feedback from some students has revealed that the availability of the OLM can also act as a source of disengagement. In particular, students have referred to engaging with more challenging practice questions as "risking a hit to your rating". Based on the request of the students, we are considering adding a practice space on RiPPLE where answering questions is decoupled from students' OLMs.

5 CONCLUSION AND FUTURE DIRECTION

This paper investigates the impact of complementing educational recommender systems (ERSs) with open learner models (OLMs). Theoretically, ERSs which tend to operate as a "black-box", should benefit enormously from OLMs. Indeed results of our randomised controlled experiments from a course at The University of Queensland suggest that complementing an ERS with an OLM can provide students with a better sense of their learning, leading to an overall positive impact on their engagement and perception of the effectiveness of ERSs; however, in practice, there may be some drawbacks. Therefore, integration of OLMs into ERSs and more broadly educational technologies need to be done with care providing: (1) students with sufficient digital literacy to understand visualisations and OLMs; (2) sufficient information and transparency about how the OLM approximates students' abilities; (3) mechanisms for students to voice their disagreement about their OLMs; and (4) safe space for students to practice without being concerned about a hit to their OLMs. A major limitation of the presented study, which restricts the generalisablity of the presented findings, is that the study was conducted only within one course with a relatively small number of participants. Future directions include replicating this study across different disciplines with a larger number of students to evaluate the generalizability of our current findings.

REFERENCES

[1] Solmz Abdi, Hassan Khosravi, Shazia Sadiq, and Dragan Gasevic. 2019. A Multivariate ELO-based Learner Model for Adaptive Educational Systems. In *Proceedings of the Educational Data Mining Conference.* 462–467.

[2] Robert Bodily, Judy Kay, Vincent Aleven, Ioana Jivet, Dan Davis, Franceska Xhakaj, and Katrien Verbert. 2018. Open learner models and learning analytics dashboards: a systematic review. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge.* ACM, 41–50.

[3] Susan Bull, Blandine Ginon, Clelia Boscolo, and Matthew Johnson. 2016. Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge.* ACM, 30–39.

[4] Susan Bull and Judy Kay. 2010. Open learner models. In *Advances in intelligent tutoring systems.* Springer, 301–322.

[5] Ricardo Conejo, Monica Trella, Ivan Cruces, and Rafael Garcia. 2011. INGRID: A web service tool for hierarchical open learner model visualization. In *International conference on user modeling, adaptation, and personalization.* Springer, 406–409.

[6] Henriette Cramer, Vanessa Evers, Satyan Ramlal, Maarten Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga. 2008. The effects of transparency on trust in and acceptance of a content-based art recommender. *User modeling and user-adapted interaction* 18, 5 (2008), 455–496.

[7] Michel C Desmarais and Ryan S Baker. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22, 1-2 (2012), 9–38.

[8] Briana B Morrison Betsy DiSalvo et al. 2014. Khan academy gamifies computer science. In *Proc. 45th ACM Techn. Symp. Comput. Sci. Educ.* 39–44.

[9] Brynjar Gretarsson, John O'Donovan, Svetlin Bostandjiev, Christopher Hall, and Tobias Höllerer. 2010. Smallworlds: visualizing social recommendations. In *Computer Graphics Forum*, Vol. 29. Wiley Online Library, 833–842.

[10] Chen He, Denis Parra, and Katrien Verbert. 2016. Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications* 56 (2016), 9–27.

[11] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work.* ACM, 241–250.

[12] Ioana Jivet, Maren Scheffel, Marcus Specht, and Hendrik Drachsler. 2018. License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge.* ACM, 31–40.

[13] Judy Kay and Bob Kummerfeld. 2019. From data to personal user models for life-long, life-wide learners. *British Journal of Educational Technology* 50, 6 (2019), 2871–2884.

[14] Hassan Khosravi. 2017. Recommendation in Personalised Peer-Learning Environments. *arXiv preprint arXiv:1712.03077* (2017).

[15] Hassan Khosravi and Kendra Cooper. 2018. Topic Dependency Models: Graph-Based Visual Analytics for Communicating Assessment Data. *Journal of Learning Analytics* 5, 3 (2018), 136–153.

[16] Hassan Khosravi, Kirsty Kitto, and Williams Joseph. 2019. RiPPLE: A Crowd-sourced Adaptive Platform for Recommendation of Learning Activities. *Journal of learning analytics* 6, 3 (2019), 91–105.

[17] Amruth Kumar and Adrian Maries. 2007. The effect of open student model on learning: A study. *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS* 158 (2007), 596.

[18] Yanjin Long and Vincent Aleven. 2013. Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In *International conference on artificial intelligence in education.* Springer, 219–228.

[19] Andrew Mabbott and Susan Bull. 2006. Student preferences for editing, persuading, and negotiating the open learner model. In *International conference on intelligent tutoring systems.* Springer, 481–490.

[20] Nikos Manouselis, Hendrik Drachsler, Riina Vuorikari, Hans Hummel, and Rob Koper. 2011. Recommender systems in technology enhanced learning. In *Recommender systems handbook.* Springer, 387–415.

[21] Sean M McNee, John Riedl, and Joseph A Konstan. 2006. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems.* ACM, 1097–1101.

[22] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. *Introduction to Recommender Systems Handbook.* Springer US, Boston, MA, 1–35. https://doi.org/10.1007/978-0-387-85820-3_1

[23] Steven Ritter, John R Anderson, Kenneth R Koedinger, and Albert Corbett. 2007. Cognitive Tutor: Applied research in mathematics education. *Psychonomic bulletin & review* 14, 2 (2007), 249–255.

[24] Vasile Rus, Sidney D'Mello, Xiangen Hu, and Arthur Graesser. 2013. Recent advances in conversational intelligent tutoring systems. *AI* 34, 3 (2013).

[25] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems.* ACM, 830–831.

[26] Ben Styles and Carole Torgerson. 2018. Randomised controlled trials (RCTs) in education research–methodological debates, questions, challenges.

[27] Gail M Sullivan. 2011. Getting off the âĂIJgold standardâĂİ: randomized controlled trials and education research. *Journal of graduate medical education* 3, 3 (2011), 285–289.

[28] BULL Susan and Adam Thomas McEVOY. 2003. An intelligent learning environment with an open learner model for the Desktop PC and Pocket PC. *Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies* 97 (2003), 389.

[29] Katrien Verbert, Denis Parra, Peter Brusilovsky, and Erik Duval. 2013. Visualizing recommendations to support exploration, transparency and controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces.* ACM, 351–362.

[30] Jiyong Zhang, Nicolas Jones, and Pearl Pu. 2008. A visual interface for critiquing-based recommender systems. In *Proceedings of the 9th ACM conference on Electronic commerce.* ACM, 230–239.

[31] Yongfeng Zhang and Xu Chen. 2018. Explainable recommendation: A survey and new perspectives. *arXiv preprint arXiv:1804.11192* (2018).