Design and Application of Automatic Feedback Scaffolding in Forums to Promote Learning

Qi Wang, Carolyn P. Rose, *Senior Member, IEEE*, Ning Ma, Shiyan Jiang, Haogang Bao, and Yanyan Li

***Abstract*—Forums are essential components facilitating interactions in online courses. However, in large-scale courses, many posts generated, which results in learners' difficulties. First, the posts are poorly organized and some deviate from the topic, making it difficult for learners' knowledge acquisition. Second, learners cannot receive timely feedback and guidance, making the learning progress unclear for them. Well-designed scaffoldings should be built based on challenges of forums to improve learners' learning outcomes, knowledge construction, and completion rate. While targeting the problems in online forums, this article proposed principles for the design of online scaffolding after analyzing the requirements of online learning scaffolding or scripts. Subsequently, in this article, we designed an automatic feedback scaffolding based on the principles and a knowledge construction model. The scaffolding provided learners with timely feedback and related learning guidance. Tags were used to assist learners in acquiring relevant information more easily. The scaffolding was then integrated into the Learning Cell Knowledge Community and used in an online course for 955 learners. The results showed that automatic feedback scaffolding positively affected learners' learning and promoted positive knowledge transformation. Furthermore, we found that the scaffolding could help learners induce more constructive behaviors defined in the Interactive, Constructive, Active, and Passive deep learning framework that demonstrated the reason for learners' knowledge transformation. At last, learners' course completion rate also increased with the help of the scaffolding, which provided evidence that well-designed scaffolding can result in positive educational outcomes. In addition, the principles proposed could also contribute to further scaffolding design and practices.**

***Index Terms*—Automatic feedback, online forum, scaffolding, system design.**

Manuscript received February 1, 2022; revised February 24, 2022; accepted March 2, 2022. Date of publication March 7, 2022; date of current version June 20, 2022. This work was supported in part by the Project for the 14th 5 Year Plan of Beijing Education Sciences (Priority Topics 2021) under Project CEEA21011, in part by the National Natural Science Foundation of China under Projects 62077007, 61877003, and 62007025, and in part by the Humanities and Social Science Research Projects, Ministry of Education of China under Project 20YJC880088. *(Corresponding author: Ning Ma.)*

Qi Wang is with Artificial Intelligence and Human Languages Lab, Beijing Foreign Studies University, Beijing 100089, China (e-mail: wangqi.20080906@163.com).

Carolyn P. Rose is with the Language Technologies Institute and Human Computer Interaction Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: cprose@cs.cmu.edu).

Ning Ma is with the Advanced Innovation Center for Future Education, Beijing Normal University, Beijing 100875, China (e-mail: horsening@bnu.edu.cn).

Shiyan Jiang is with the Department of Teacher Education and Learning Sciences, Learning Design and Technology, College of Education, North Carolina State University, Raleigh, NC 27695 USA (e-mail: sjiang24@ncsu.edu).

Haogang Bao is with the National Institute of Education Sciences, Beijing 100088, China (e-mail: bestbaohg@126.com).

Yanyan Li is with Beijing Normal University, Beijing 100875, China (e-mail: liyy@bnu.edu.cn).

Digital Object Identifier 10.1109/TLT.2022.3156914

I. Introduction

MOOCs and online courses provide valuable opportunities for learners to acquire knowledge and advance themselves. However, there is a high possibility that learners drop out and cannot achieve satisfactory learning outcomes in online courses [1] because the unregulated spontaneous learning process lacks timely and useful feedback that could help solve problems and maintain interest [2], [3]. Some researchers have revealed that feedback is essential in promoting learners' learning [4], especially in MOOCs. However, it is difficult for each learner to receive timely and useful feedback for large-scale MOOCs. Using automatic analysis techniques to support learners' learning can be helpful [5]. Bey *et al.* [6] investigated two automatic assessment methods in MOOCs for program evaluation and found that they could benefit learners' learning. Some researchers have included discussion forums in online courses as activities to support better interactions among peers and instructors. On the one hand, forums can record learners' efforts and contributions during the learning process. Research has shown that learners' efforts during the learning process substantially contribute to the learning result, and effort visualization feedback tools are used to help improve team members' performance [7]. On the other hand, posts in forums provide opportunities for students to communicate with each other on specific problems or difficulties. Some researchers provided evidence that learners' posts in forums can reflect their understanding, problems, and needs. Moreover, posts in forums are mostly proposed based on previous learners' ideas, which enable problem-solving and knowledge construction. Therefore, providing linguistic analysis and effective feedback for learners in forums could benefit learners' learning to a considerable extent [8]. From another point of view, some researchers have revealed that the time of feedback can also influence learners' learning [9]. However, it is not feasible to conduct analysis and provide feedback in time manually, as there are usually many posts in MOOCs. Accordingly, some researchers have begun to use automatic tools for online interaction analysis [10]–[12]. These tools can help realize timely feedback, while there is still limited promotion for effective discussion [13], [14]. The reason is that the feedback is not always well-designed and organized, which impedes learners to extract useful information from them. Wang *et al.* [15] investigated the effect of different feedback formats. They found that well-designed and detailed feedback that contained more instructional information could result in better learning perception and lower cognitive load. This made the automation of online feedback even more complex and specific scaffoldings should be provided considering the feedback content and the time and format. Although some researchers have investigated strategies for online learning scaffolding and script design, systematic principles have not yet been established to guarantee effective discourses and feedback, which will be helpful for learners' learning and course completion rates. Based on related research, we argue that effective scaffolding supporting deeper discourse and feedback in online forums should involve the following aspects.

1) *Content-Related Discussions:* Guidance provided by scaffolding should be seamlessly relevant to the learning content.
2) *Timely Feedback:* Feedback should be given in real-time to promote learner motivation and knowledge awareness [16].
3) *Reflect Learning State and Promote Progressive Discourse:* Guidance should support learners using content that reflects their current state and helps them engage better.
4) *Clearly Organized and Navigated:* Posts should be organized according to a specific classification system to help learners easily discover information in massive posts.
5) *Provide Supportive Guidance for Further Learning:* Scaffolding should detect the topic and quality of current discussion, and deliver feedback or prompts related to problems that help learners adjust their learning strategy timely.

This will enable learning engagement, retention rates and performance to be improved.

In MOOCs, instructors and assistants are confronted with a large number of learners which makes it impossible to provide this type of scaffolding manually. Therefore, we developed the principles that support the design of scaffolding and present an automated approach to providing this specific scaffolding in the Learning Cell Knowledge Community [17]. We then investigate its effects using an online course of 955 learners. The research questions in this study focused on the following three aspects:

1) If the scaffolding designed based on the proposed principles can promote learners' learning especially the knowledge construction level?
2) If the scaffolding can increase learners' high-level behavior and engagement?
3) If the scaffolding can increase the course completion rate?

At the end of the experiment, data collected provided evidence that despite an automated approach's limitations, automatically triggered feedback produced positive effects on student discourse and course performance.

II. Literature Review

Section I shows that forum is a widely used method to help learners obtain information and enhance their communication and understanding in online courses. During this process, feedback is essential because learners may deviate from the correct direction, thus providing timely and targeted feedback could help realize more effective learning. However, forums are usually large scale in MOOCs, making it difficult to provide timely, well-designed and detailed feedback manually. Therefore, automated feedback should be involved and specific scaffolding should be developed. Currently, although some studies have engaged in automated feedback in forums, they have not established a systematic framework elaborating the influential factors and integrating them with learners' real needs for effective feedback. This section first reviews the current studies on feedback scaffolding and automatic feedback scaffolding. Based on the problems in the current studies, we investigated factors that promote communication in online forums to help design principles for feedback scaffolding.

A. Scaffolding and Automatic Feedback Scaffolding Promoting Online Discussion

Learners in forums can face many problems that affect their learning, such as cognitive treks, information explosion, lack of feedback, and low-level knowledge construction. To solve these problems, scaffolding should be provided. There are now many studies focusing on facilitating learners' communication and knowledge acquisition in forums. The scaffoldings positively influence learners' learning despite some limitations of scaffolding design.

Forums can promote learners' communication and interaction where correct directions should be ensured for the learning tasks [18]. Some research involved questions, tasks, or teaching assistants to provide facilitation based on specific strategies to guarantee the learning direction [19]. Rienties *et al.* [20] conducted research using the Optimal model which provided tutor guiding and tasks to scaffold learners' discourse. The results showed that scaffolding could reduce learners' off-task discourse. Some researchers have used ICT tools and strategies to reduce learners' cognitive load caused by information explosion in forums. Beers *et al.* [21] used explicit scaffolding to promote learners' discourses which showed that it could help learners engage and contribute more. These scaffoldings helped learners a lot in forums by providing specific feedback or guidance. However, they relied heavily on the tutor and teaching assistants' facilitation which could not work well when the number of learners increased substantially.

Under these circumstances, some research has developed automatic feedback scaffoldings. Automatic feedback scaffolding is designed and embedded in learning software and triggered when learners came up with an operation [22], [23]. In education, automatic feedback is used to provide immediate response and improve learners' participation [24]. Therefore, feedback should be provided promptly, with proper form and with proper content. In practice, machine learning technologies are essential for automatic feedback and a specific framework is required for different types of feedback. Current feedback mainly focuses on emotion regulation and content improvement.

For emotion regulation, there has been much research in the online environment [25], and their methods have been used in education. Provoost [26] used learners' online texts and algorithms to detect their emotions. The results showed that the automated method was moderately consistent with human classification. Lin and Kao [5] conducted an experiment to provide timely feedback on learners' mental state in MOOCs and showed that feedback could facilitate learners' self-awareness of mental efforts and promote their learning. Wen *et al.* [27] investigated sentiment analysis in MOOC forums and found a correlation between learners' sentiment expression and dropout behaviors. Furthermore, Moreno-Marcos *et al.* [28] revealed learners' engagement in MOOC forums based on automatic sentiment analysis, which would help promote learners' learning motivation if proper guidance was provided.

Different analysis schemes were proposed for content improvement [29]. Based on these schemes, researchers could get to know learners from different aspects; thus, different types of scaffoldings could be designed, such as content correction, text analysis, and learning guidance. De Vries *et al.* [30] investigated automatic error detection and provided correction feedback scaffolding in second language learning. Learners using this system had a positive attitude toward learning. Thus, immediate feedback can be precious in scaffolding efficient learning. Akçapınar [31] used text-mining techniques to automatically provide feedback on learners' plagiaristic behavior. The results showed that the method could effectively improve their learning behaviors. Kovanović *et al.* [32] conducted automated content analysis in online discussions and revealed several features that could reflect learners' cognitive presence. They also found that learners' meaningful interaction in forums could improve the learners' social presence and performance [33]. More specifically, in the learning content, Liu *et al.* [34] used multidimensional automatic analysis model to assess learners' content. The dimensions included grammar, spelling, sentence diversity, structure, supporting ideas, coherence, and conclusion. This high-quality model could guide learners' learning and achieve performance that equated to teacher feedback. The results also showed that learners given automatic feedback scaffolding performed significantly better than those in the non-feedback environment. In dialogue guiding, Tsan [35] used an automatic dialogue agent to help young learners with the computer science curriculum, and found that the agent could promote learning. Tegos *et al.* [36] used a conversational dialogue system, "MentorChat," to support online learners' tasks and noted that weaker interventions could help students better complete tasks.

However, researchers have also realized that automatic feedback scaffoldings do not always work well if not designed elaborately according to learners' needs in forums. Howley *et al.* [37] used Bazaar as an automatic feedback agent to investigate the effectiveness of promoting learners' self-efficacy and learning outcomes. They found that if learners were provided with automatic prompts without considering their self-efficacy, they could have negative effects on learners. Mu *et al.* [38] developed a framework for automatic text analysis in online discussions and they made it more adaptive for different learning contexts. They also indicated that well-designed automatic analysis could monitor learners' learning in real-time and provide personalized interventions that would help improve learners' thinking and learning progression.

The literature review revealed different types of feedback scaffoldings. The studies have focused on various aspects, such as timely feedback, emotion-sensing, knowledge guidance, and prompts related to self-efficacy. The results provide evidence that even though these automatic feedback scaffoldings can promote learners' learning at most times, they still cannot sometimes work because they mainly solve one facet of the learners' problem in online learning. However, learning is a systematic process that involves many influencing factors. Only when scaffoldings are designed systematically based on these factors could learners benefit the most; thus, knowing the influencing factors in online learning could be an essential issue.

B. Factors and Principles Promoting the Design of Scaffolding in Online Forums

Although many types of feedback scaffoldings have been designed to facilitate learners' learning, few have elaborated and shared deep discussions in forums [39]. This is due to the learners' complex needs for feedback. In learners' learning practices, the roles of feedback include content correction, idea analysis, providing guidance, identifying what the tutor wants, and giving meaning to improve their learning [40]. A detailed analysis of the functions of feedback scaffoldings should be conducted. In this way, we can construct a systematic principle framework for automatic feedback scaffolding design.

One objective of online forum scaffolding is to promote learners' content interactions. Well-designed interaction and guiding rules are considered some of the most critical factors for achieving high-level cognition in forums. Wang *et al.* [41] investigated the role of interaction in discussion forums and found that learners who interact more tend to achieve more.

Recommendations for scaffolding have been essential to improve learners' learning and cope with a large amount of data online. Information explosion is one factor that could prevent learners from acquiring useful knowledge, and techniques have been recommended to filter adaptive posts or resources for learners. These recommendations help improve learners' reading frequency and summary ability [42]. Moreover, navigation is vital in deciding learning outcomes once online forums become very large [43]. Such navigation, termed social navigation, can be realized by considering learners' ratings or operations [44]. Social navigation is often used in large group learning and visualizes the group cognition of knowledge [45]–[47]. This kind of visualization tends to change learners' attention and could impact learning outcomes [48].

Focusing on knowledge or cognition-related scaffolding, some researchers have demonstrated that forum discussion

TABLE I
PRINCIPLE TABLE FOR SCAFFOLDING DESIGN

Principle Name	Detailed Description	Targeted Problem	Example
(1) Content Related Principle	The scaffolding should ensure the discussion process is in accordance with the topic of the learning content.	Avoid meaningless posts and deviation	If the learner proposes: 'Let's play,' the scaffolding should get him/her back by saying: "You should talk about something related to the content."
(2)Timely Feedback Principle	The scaffolding should give learners feedback in time / at the time learners are faced with problems.	Make the learners feel noticed and reduce a dropout rate	When the learner comes up with a post, the scaffolding should evaluate the post in time and make the learner clear about his/her learning.
(3)Reflect Learning State and Promote Progressive Discourse Principle	The scaffolding should detect the learners' learning state and provide them with information that will facilitate the topic	Promote effective and deep discussion	If the learner comes up with a post, the scaffolding should evaluate it and report the level. After that it should provide information to promote the current discussion
(4) Clearly Organized and Navigated Principle	The scaffolding should have the ability to classify the information generated during the discussion and organize it clearly so that learners can easily find useful things.	Information Explosion	When the learner proposes a post, the scaffolding will assess it and tag it with a specific classification, such as learning level 'L1–L6'.
(5) Provide Supportive Guidance for Further Learning Principle	The role of scaffolding should be to provide external support or guidance for promoting further learning.	Learners' loss during the learning process	If the learner reaches a level where they could interact relatively deeply with others, the scaffolding should provide suggestions for reflection and summary.

could play positive roles in knowledge construction and effectively support knowledge acquisition [49], [50]. Butchart *et al.* [51] found that automated feedback could be helpful for learners' critical thinking cultivation, which will ultimately contribute to their knowledge improvement. Kellogg *et al.* [52] researched learners' essay feedback and found that continuous feedback could help learners' learning compared to intermittent feedback.

To foster knowledge construction in collaborative learning, Weinberger *et al.* [53] investigated socio-cognitive structuring tools that use interactive, content structuring functions to guide learner discussion. Both structuring functions helped promote knowledge construction, whereas only interactive structuring enabled learning outcomes. To ensure the validity and reliability of the content, researchers have developed rules for scaffolding. In designing the rules, they found that more explicit scaffolding could improve engagement [54].

Some researchers have focused on interventions that aim to enrich and promote student interaction [55]. Rule design is an essential element in these interventions. With the rapid development of instant communication techniques, researchers have focused on real-time or automatic scaffolding based on machine learning [56].

These studies indicated that learners face several key problems in large-scale online learning, such as content-related guidance, lack of timely feedback, information explosion, and shallow discussion and interactions. These problems result in lower learning outcomes and higher dropout rates. From another perspective, these problems raise several factors that could guide the design of scaffolding, including content-related discourse, clear navigation or visualization, progressive interaction, instant communication, and adaptive guidance. These factors should be considered when designing online scaffolding. Starting with these factors, we proposed principles (see Table I) that should be considered in the design of scaffolding by integrating learners' requirements of online forums.

1) *Content-Related Principle:* During online discussions, learners often provide irrelevant posts that cause the discourse to deviate from the topic. This reduces the effectiveness of online learning. Evidence suggests that participation in content-related interactions is a predictor of better performance [57]. Scaffolding should maintain or extend the discourse topic to ensure the quality of online learning.
2) *Timely Feedback Principle:* In large-scale online learning, many learners demand the instructor's time, and it is unlikely that timely feedback can be provided to everyone. This may make learners feel despondent and increase their dropout rates. Instant feedback could provide a clear motivation for improving learners' practices [58]. Long delays in receiving feedback make learners feel unnoticed and reduce their throughput [59].
3) *Reflect Learning State and Promote Progressive Discourse Principle:* Content provided by scaffolding should reflect learners' understanding and promote their knowledge progression. Generative discourse could help learners reflect on their understanding and foster knowledge development by comparing and integrating other learners' posts [60], [61]. With this type of scaffolding, learners can achieve deeper learning.
4) *Clearly Organized and Navigated Principle:* This is to ensure that useful posts can be easily located. In large-scale courses, forums are flooded by the learners' posts. Being well-organized makes it easy to navigate the discussion and improves learning efficiency [62]–[64]. Posts proposed by learners should be well-organized within specific categories, such as knowledge level, so that learners can effectively find information at the appropriate level.
5) *Provide Supportive Guidance for Further Learning Principle:* To help learners conduct effective learning based on the discourse in large-scale online course forums, scaffolding should also provide supportive

guidance for further learning. For example, if a learner is currently at a low knowledge level related to low engagement, scaffolding should offer suggestions on how to participate in the activity and think more deeply about the learning content. Such feedback is designed based on script theory, providing an external script or helping improve learners' development of an internal script [65].

In this article, we investigated the proposed principles in scaffolding design and hoped to promote three aspects of MOOCs. The first is to enable constant feedback by providing timely analysis and intervention. The second is to provide content-related prompts and learning guides to support learners' behaviors and engagements, and promote their knowledge construction level. The third is to increase the course completion rate through deeper interactions and clearer navigation. The next section introduces the design of automatic feedback scaffolding in online learning using these principles.

III. Method

This article aimed to help solve problems in online learning forums using the proposed principles for online scaffolding and automatic feedback scaffolding for real use. An automatic feedback model that can realize the principles is quite important during this process. The section first introduces an educational and technical model to support the design of automatic feedback scaffolding. Then, the functions of the scaffolding are presented. Finally, we will provide full details of this article into scaffolding in an online course.

A. Automatic Feedback Model Supporting Scaffolding

We investigated several models to realize these principles. The educational model supports learners with strategies based on learning theories. The technical model enables supporting learners' timely interactions.

1) Educational Model Supporting Scaffolding: An educational model needed to be developed to support scaffolding design to achieve principles (3) and (5). Here, we chose a widely used knowledge construction model [66], [67]. The reason for choosing this model is that it evaluates learners' various knowledge construction features based on their posts. For example, one's post is just to share information or negotiate with the other. According to these classifications, we could measure learners' engagement and knowledge construction, which reflect their contribution or deep understanding of the content. Moreover, the model relies little on the knowledge domain. Thus, it can be used and transferred among domains well. The schema of the model classifies the knowledge construction process into the following five dimensions (P1–P5):

1) sharing and comparing information;
2) discovery and exploration of dissonance or inconsistency;
3) negotiation of meaning and co-construction of knowledge;
4) testing and modification of the proposed synthesis and co-construction;
5) agreement statement(s) and applications of newly constructed meanings.

TABLE II
KNOWLEDGE CONSTRUCTION SCHEMA

Code	Dimension	Behavior types and examples
P0	Irrelevant interactions	Irrelevant information. Very good. You have done a good job. You are an idol for me. I have got a lot from it.
P1	Sharing/comparing of information	Creating a learning cell, adding learning activities, uploading learning material, and releasing reading work and a concept map.
P2	Discovery and exploration of dissonance or inconsistency among participants	Browsing, collecting, and giving feedback on learning cell created by the peer; coming up with confusion during learning. Can "qualitative research" be translated into "质性研究" or "定性研究"?
P3	Negotiation of meaning/co-construction of knowledge	Cooperative editing learning cell, modifying video and content, adjusting content structure, comment. Discussion with the peer on topics and give suggestion on problems.
P4	Testing and modification of proposed synthesis or co-construction	Remark, comment, annotation, pointing out the problem. I cannot hear clearly of the back of the video. I think "educational design research" translated into "教学设计研究" will lead to misunderstanding. The micro-course does not include learning activity.
P5	Agreement statement(s)/application of newly constructed meaning	Reflection Writing reflective journal entries. I think a teacher cannot be replaced by a pedagogical agent. I agree that both internal validity and external validity are important for a study.

Using this schema, we could measure learning by evaluating whether learners have a deep understanding and engagement in a specific area. If the learners do not achieve a satisfactory level, guidance may be needed. Moreover, we added another dimension: "P0: irrelevant information," before P1 to detect meaningless information which may affect the efficiency of the course. The coding schema is shown in Table II. The six dimensions represent the gradually increasing knowledge construction levels of posts. The behaviors related to each dimension are also shown in the table.

2) Training the Technical Model for Scaffolding: A technical model was needed to realize the content-related principle (1), timely feedback principle (2), and clearly organized and navigated principle (4). In addition, the scripts in the technical model provide learners with the automatic and adaptive guidance defined in principles (3) and (5). The technical model makes automated feedback feasible. This section describes the model training process based on the knowledge construction model. A total of 1350 high-quality posts from 3 online courses in different domains (learning science, educational communication and technology research, and photography) were chosen as the training data to ensure that the model was applicable to other domains. We chose three courses to avoid

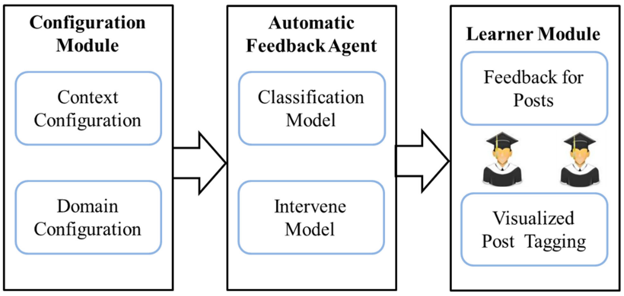


Fig. 1. Architecture of the scaffolding.

the influence of domain knowledge to the greatest extent in the classification. Two doctoral students expertized in these topics conducted data coding using the knowledge construction schema. The two coders discussed the standard to code these posts before coding for one time and agreed. After coding, the Cohen's kappa value of the two coders was 0.735, which proves that their coding was consistent. We divided the 1350 samples into training data (924 items) and validation data (426 items).

The machine learning tool, "LightSIDE," was used for model training. Developed by Carnegie Mellon University, it is a workbench that integrates machine learning functions and algorithms. Researchers can use visual interfaces without programming. The following three steps are required to obtain a model:

1) feature extraction;
2) model training;
3) error analysis.

Feature extraction can divide all input posts into isolated features representing the current post from one point of view. Our feature extraction used the word segment interface from Stanford and 4517 features were obtained. Model building and evaluation could support learners' model training based on the algorithms embedded in LightSIDE. We used logistic regression, and error analysis to adjust the model's performance. After the operation, we obtained a model with a 0.746 accuracy and a kappa value of 0.659. The result was acceptable for making predictions in a real-learning context. We investigated several features for model training, such as word, part of speech, and length during this process. They were used to distinguish between different types of knowledge construction levels. As this article is to provide an acceptable prediction model for automatic feedback scaffolding, we used a logistic regression model and did not compare the different models.

B. Function Design of the Scaffolding Based on LightSIDE

We designed and developed an automatic feedback scaffolding tool to make the principles and model applicable in online learning. The architecture of the automatic feedback scaffolding involves three modules: a configuration module, an automatic feedback agent, and a student module (see Fig. 1).

1) Configuration Module: The configuration module consists of context and domain configurations. This is the basis for realizing principles (3) and (5) because it defines basic knowledge for specific learning areas. The context configuration mainly defines the situations or interfaces where automatic feedback scaffolding can be triggered. Learners' needs determine the context where the feedback is presented. The manager only needs to configure the URL request of a specific context. This research was triggered in forums and aimed to promote learner reflection. This module also provides the function of the domain configuration which allows experts to define domain knowledge. The manager should provide a dataset supported by a coding schema in the specific domain in determining domain knowledge. Subsequently, the feedback model can be updated.

2) Automatic Feedback Agent: The automatic feedback agent is the most critical part of the scaffolding. This supports the principles and provides functions to learners. This module contains classification and intervene models. The classification model is based on the LightSIDE server launched in the Learning Cell Knowledge Community. When learners participate in discussion activities, they must submit posts to show their understanding. The classification model receives posts through the interface provided by LightSIDE and then uses the trained model for classification. After the posts are classified into P0–P5, the automatic feedback scaffolding can decide the learners' knowledge construction level and guide further learning with the intervene model. The interface provided by LightSIDE is available online.[1] These two models can provide learners with timely feedback, content-related judgments, learning state reflections, and guidance.

The intervene model defines the rules for automatic feedback. We wanted learners to regulate their learning process, and thus we did not provide them with specific resources. We only provided them with guidance in terms of strategy. Many researchers have shown how direct resources can trigger relatively shallow learning instead of deep learning. Transforming guidance from "knowledge telling" to "knowledge construction" may be more effective [68]–[70]. These rules are listed in Table III.

As P0 means the post is off-topic, rules should guide the learner to take a more active part in learning. P1 indicates that the learner is shallowly proposing some concepts and lacks knowledge integration and reflection. Rethinking the concepts and participating in the activities will help learners understand more. P2 indicates that learners are confronted with some problems. The first step is to return to the learning content and discuss it with peers. If it does not help, the teacher will provide targeted feedback. P3 means that the learner participates in discussions and negotiations and has a relatively good understanding. They need to reflect more deeply with the help of the learning content and other learners' points of view. P4 means the learner is proposing constructive suggestions for the platform or course based on a relatively good understanding of the course. They require deeper reflections. P5 indicates that the learner has deeply reflected on the learning content. They are doing well on this topic and can move to the next

[1][Online]. Available: http://www.etc.edu.cn/ko/id/comments/tab?content=content

TABLE III
RULES FOR AUTOMATIC FEEDBACK

Classification	Rule (example)
P0	You are not focusing on the learning content. Maybe you failed to master the core concept in the learning. Look back to the learning content and take an active part in the learning activities with peers.
P1	You have got some concepts from the learning content. If you can integrate them with your own knowledge, that will help you a lot. You can focus more on the relations in the content and take part in the learning activities.
P2	You have proposed good questions. That's good, look through the learning content again and discuss with your peers, you may get the answer. If you still don't understand, the teacher will give you detailed feedback.
P3	Your discussion reflects you have a relatively good understanding of concepts in the learning content. However, if you can conduct some reflection based on the discussion as well as the learning activities, it would help you more.
P4	You come up with good suggestions for the course and the platforms. The course will evolve with your suggestions. You can also come back to the learning content and activities for more learning.
P5	You have a really deep reflection on the learning content. That's great. Continue to work hard and you can connect what you learn with the next topic.

Fig. 2. Feedback for post.

topic. They need to continue their performance and make connections between different topics.

3) Learner Module: The learner module is the interface with which learners interact. It provides feedback for the posts. When the learner submits a post, the automatic feedback agent will develop specific guidance and show it to the learner through the interface (see Fig. 2).

Learners might struggle to find useful information in the discussion forum when solving a problem. Visualized post-tagging provided all posts with specific tags. P0 indicated "irrelevant or shallow understanding"; P1 as "listing simple concepts"; P2 as "Proposing questions"; P3 as "discussion and negotiation"; P4 as "constructive suggestions for course improvement"; and, P5 as "deep reflection" (see Fig. 3). These tags are key elements for realizing principle (4).

Fig. 3. Tagged posts.

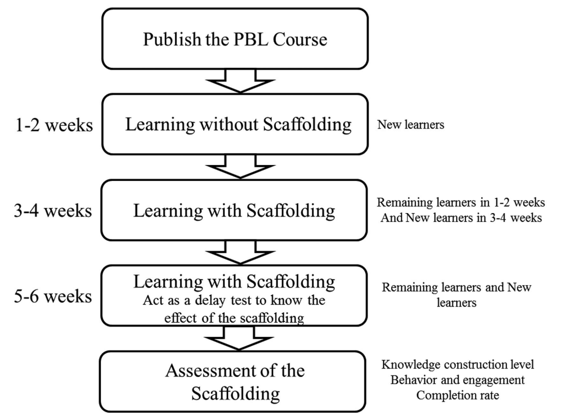


Fig. 4. Process of the course.

C. Research Design Based on the Proposed Scaffolding

1) Participants: The sample consisted of 955 learners who joined the MOOC named "Project-based Learning under Blended Learning Environment" voluntarily (see Fig. 4). The MOOC was set up by a professor from a university in Beijing and the course aimed to promote learners' competency in designing project-based learning courses. Learners in this course include pre-service teachers and in-service teachers in K-12 education and higher education. Some researchers, undergraduate and graduate students were also required to take this course. The learners had not taken part in similar courses. The learners come from different places in China, rural and urban areas. Under these circumstances, the learners are of different backgrounds, and they could represent the population in the educational area. During data analysis, we excluded all learners whose duration in the course was less than 15 min, which is less than one lesson.

2) Experimental Process: "Project-based Learning under Blended Learning Environment" was set up on April 8, 2019. Instructors uploaded 19 learning cells. Each learning cell is a subsection of the course, and contains learning activities and materials to support learning. The content of the course includes the following:

1) the role of digital technologies in PBL;
2) how to choose the topic and set the target in PBL;
3) the design of learning outcomes and schedules;
4) information collection and learning activities;
5) the completion of the products in PBL and the method for evaluating performance.

Learners could join the course at any time, and the suggested learning time was approximately five weeks. Most learners completed the online learning content in four weeks

TABLE IV
ICAP FRAMEWORK

Category	Interactive	Constructive	Active	Passive
Characteristic	Dialoguing	Generating	Manipulating	Receiving
Definition	Generating additional inferences and information vis dialoguing with peers	Generating new inferences or information beyond what is presented	Manipulating learning materials to focus attention	Merely paying attention to receive the learning material
Knowledge change processes	Co-inferring Co-Creating	Inferring Transferring	Applying	Storing Recalling
Learning outcome	Deepest understanding	Deep understanding	Shallow understanding	Minimal understanding
Examples	Comment; Post; Participate（submit work, draw concept map）	Comment; Post; participate（submit work, draw concept map）	Download (download useful resources) Mykoevaluate (check my evaluation result); profile	View (Listen to a Micro lecture, read an article or web page, watch the video); Koevaluate (check the evaluation module)

and designed their products in the last week. To evaluate the role of automatic scaffolding, we did not include the scaffolding in the first two weeks of the course. During the two weeks, all learning took place in a traditional environment without timely feedback or post tagging. We made scaffolding available from the third week (April 23, 2019). After the scaffolding was included, we observed the online learning process. We extended the observation period to more than six weeks so that the learners joined after the scaffolding involved could complete the online lessons. When the learners completed the six-week learning, the six-week course can be divided into the following three stages (see Fig. 4).

1) The first two weeks represented the effect of the learners' performance without automatic feedback scaffolding.
2) Weeks 3–4 represented some new learners' performance when they first used the scaffolding and old learners' performance from 1–2 weeks.
3) Weeks 5–6 represented the delay or persistent effect of the scaffolding in promoting learning.

The reason why we divided the stages with "two weeks" was that the scaffolding was not included in the first two weeks. Therefore, aimed to conduct a better comparison, including learning totally without scaffolding, learning with scaffolding for some time, and learning always with scaffolding. Following this, we collected the course data to assess the effect of scaffolding.

3) Instruments: RQ2 focused on the relationship between learners' knowledge transformation and behavior. Learners with more high-level learning behaviors have a greater probability of achieving deeper learning [71], [72]. Therefore, learning behavior and engagement were measured first. As the scaffolding was developed based on the Learning Cell Knowledge Community, the community had a systematic behavior classification [71]. The classification is similar to the Interactive, Constructive, Active, and Passive (ICAP) framework developed by Chi and Wylie [72], which could relate learning behaviors to deep learning. As this framework has been widely used in different studies [73], [74] and could reflect the idea of our community, it was feasible to use it for behavior analysis in our study. The details of the framework are shown in Table IV, and we defined some examples based on our study. We did not distinguish between the interactive and constructive levels and combined them when conducting behavior analysis because they were all at the deep understanding level. We defined behaviors at three levels: 1) passive; 2) active; and 3) constructive. The framework was used to determine the distribution of behaviors at different depths. Learning behaviors were coded to determine learning engagement to answer the second research question.

4) Data Preparation and Statistical Methods: Data Preparation: The data used in this article included posts, learners' behaviors, and learners' scores. The posts were tagged with their classification "P0–P5." They were collected when the learners submitted their understanding. We collected behaviors defined by the Learning Cell Knowledge Community.

In this article, eight typical behaviors are used: "Comment," "Post," "Participate," "Download," "mykoevaluate (this indicates the learner is checking his own evaluate result)," "Profile," "View," and "koevaluation (this indicates the learner is checking the course's evaluation schema)." All the behaviors were collected based on the xAPI format which indicates "who do something at a specific time." The platform computed learners' scores when they completed specific tasks in the course. The scores were collected during different stages of the course.

Statistical Methods. To answer the research questions, we mainly investigated the frequency features of the collected data, such as frequencies of learners' behaviors, frequencies of learners' knowledge construction level at different stages, and learners' duration and scores in the course. These data types were then analyzed to understand the changes triggered by scaffolding. Generalized Sequential Querier [75] is a computer program for analyzing sequential observational data. It is often used for lagged sequential analysis to assess learners' knowledge transformation and behavior pattern. This method can distinguish learners' learning patterns by computing their behavior sequences. If the transformation *z*-score between two behaviors is higher than 1.96, it means that the transformation pattern is significant and this change is more likely to occur during the learning process. We also tagged the transition lines to represent this transformation pattern as significant. We found evidence of how learners' learning can be influenced by these results.

5) Research Questions: The following three research questions were asked.

RQ1: If the scaffolding can improve learners' knowledge construction level?

RQ2: If the scaffolding can increase learners' higher-level behaviors or engagement? Are these behaviors or engagements related to learners' knowledge transformation?

RQ3: If the scaffolding can promote the course completion rate?

IV. Results and Discussion

Here, we provide evidence of the effectiveness of our scaffolding design. The three aspects of the results will be discussed: the role of scaffolding on knowledge construction level, behaviors, and course completion rate.

A. Role of Scaffolding in Promoting learners' Knowledge Construction Level

We analyzed the role of scaffolding on learners' knowledge construction levels by comparing posts at different levels at different stages. The transformation of posts is presented to demonstrate the positive effect of scaffolding.

1) Learners' Knowledge Construction Levels Before and After Scaffolding: To answer RQ1, we collected data from the three stages and compared the frequency of each classification. In Stage 1, 666 posts were generated, whereas 847 and 878 were generated in Stages 2 and 3, respectively. The percentage of posts at each level is shown in Fig. 5, demonstrating that P0 was higher than all other classifications in Stage 1. However, the percentage of P0 decreased significantly from 55.86% to 31.76% in Stage 2. In turn, the P5 percentage increased. This means that most learners had relatively shallow understanding during the first stage of the course. Most posts were irrelevant to the learning content, and learners merely listed concepts. Learners seldom undertook deep reflections. In Stage 2, scaffolding was embedded into the Learning Cell Knowledge Community. With timely feedback and adaptive guidance, learners can gradually achieve a deeper understanding.

Stage 3 revealed that the percentage of P0 was even lower (14.05%) and the percentage of P5 was higher (39.86%), suggesting consistent improvement. Post quality gradually transformed from low to high.

To demonstrate the role of scaffolding in the transformation, we collected data from new learners who joined the course after scaffolding was introduced in the Learning Cell Knowledge Community. These learners' results are represented as "New User" in Fig. 5. Compared with new users

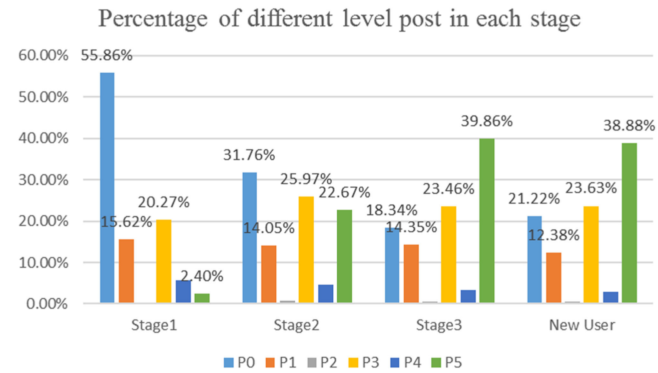


Fig. 5. Percentage of different level posts in each stage.

(learners without scaffolding) in Stage 1, new users' P0 percentage was significantly lower than that in Stage 1 (21.22% versus 55.86%). However, the P5 percentage was significantly different (38.88% versus 2.40%). The scaffolding's guidance and feedback helped improve the quality of learners' posts and promoted a change in knowledge construction from lower to higher levels.

Scaffolding promoted the transformation of learners' knowledge construction levels gradually from lower to higher (P0→P5). In traditional courses without guidance, learners set out to obtain a qualification, but limited feedback means they are often ignored. Learners perform tasks but may not receive timely evaluations and hardly improve their knowledge construction level [76]. We designed guidance and feedback according to the scaffolding principles, viz. content-related, timely feedback, knowledge reflection, and guidance to specifically address the problems noted. As a result, learners received timely feedback on the quality of their posts. They learned about their performance and what they should do to achieve better results. This was an unobtrusive and positive intervention to help learners reflect on their understanding of the learning content. Research suggests that this type of intervention helps improve knowledge cognition [36], [77]. Therefore, knowledge construction levels might be even higher in the second and third stages of this experiment.

2) Further Testing of Knowledge Construction-Level Transformation: Further testing was conducted using lagged sequence analysis to better understand how knowledge construction levels changed during the experiment. The transformation is shown with the transition lines in Fig. 6.

From the figure, we can see that the transition lines changed after the scaffolding was introduced. Before using the scaffolding, learners tended to propose posts at the same level, suggesting they did not know if there were problems with their learning. Some learners typically explained their feelings even if they were not relevant to the content (P0 → P0). Some of them could only express their ideas and could not improve them as the course proceeded, thus, their behaviors remained unchanged (P3 → P3, P5 → P5). However, after the scaffolding was introduced, rich transformations were observed. New transformations between different knowledge construction levels appeared, such as P3 → P1 and P2 → P5. This means

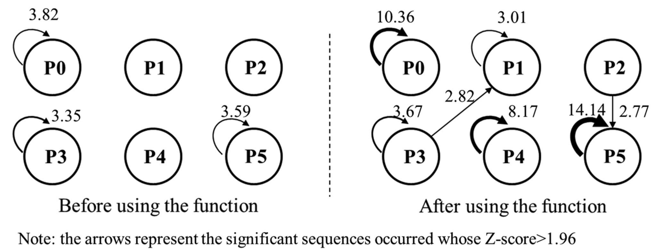


Fig. 6. Comparison of the result of lagged sequence analysis.

that the guidance of automatic feedback acted as a trigger to help learners rethink their understanding. The rethinking process promoted their understanding from shallow to deep [78]. As for scaffolding design, we classified posts with visualized tags to help learners find useful information more easily. This could also promote learners' understanding by interacting with others. During this process, learners' understanding fluctuated so that richer transformations occurred. Although there was no significant transformation from P0 to other classifications, the reduced frequency of P0 indicates that scaffolding helped learners develop. This was also important for the promotion of the course. Changes in the learners' knowledge verified the usefulness of the scaffolding, and the transformation of learners' knowledge proved the scientific value of the principles underpinning the scaffolding design, especially timely feedback, clear organization, and guidance.

B. Role of Scaffolding in Influencing Learners' Higher Level Behavior and Engagement

The role of scaffolding on learners' behavior was measured by comparing the changes in different behavior levels based on the ICAP framework. Following this, we analyzed the relationship between changes in behaviors and knowledge construction.

1) Changes in Learners' Higher Level Behavior Using Scaffolding: We analyzed learners' different behavior types based on the revised ICAP framework. Statistics relating to behaviors before and after scaffolding are presented in Table V. We can see that before the scaffolding was included in the experiment, the percentage of high-level behaviors (constructive) was relatively lower than once the scaffolding was introduced (8.41% versus 14.90%). Moreover, the passive level for learners without scaffolding was much higher than with scaffolding (60.38% versus 51.08%). Scaffolding positively promoted meaningful behaviors and engagement.

With automatic feedback scaffolding, learners could understand their problems immediately, based on principles (2) and (3), after proposing a post. The guidance provided by the scaffolding, based on principle (5), enabled further engagement with learning thereby increasing meaningful and higher level behaviors. This is in accordance with other research using agents for scaffolding [79], [80] and with the result of Section IV-A that learners tended to have higher knowledge construction levels. The increased higher level behaviors could be one reason why knowledge construction levels improved.

2) Relationship Between Increased Meaningful Learning Behaviors and Learners' Knowledge Construction-Level Transformation: We compared the behavior of learners with different knowledge construction-level transformations to verify the relationship between learner behavior and knowledge construction-level transformation. In the course, we defined behaviors that reflect learners' engagement for analysis: comment (create meaningful posts or comments for specific learning content), koevaluate (check the evaluation module of specific learning content), mykoevaluate (check one's evaluation result of specific learning content), participate (take part in a learning activity such as submit work or draw concept maps), post (create posts in a discussion forum), profile (check personal learning history and profile), resource download (download course-related resources), and view (view the learning content). We conducted a statistical analysis of the frequency of each behavior classified by the transformation. As these behaviors reflect different learning engagement levels during the learning process [71], frequency accumulation could not highlight the roles of different behaviors. Therefore, we weighted the frequency of each behavior to obtain a scaled engagement value. For constructive behaviors, we multiplied the value by 2. For activating behaviors, we divided the frequency by 2, and for passive behaviors, we divided the frequency by 5. The results are presented in Table VI. For positive transformations such as 0t1 (representing changes from P0 to P1), 0t3, 0t5, 1t3, 1t5, and 3t5, the scaled behavior engagements are mostly higher than 300, which is a relatively high learning engagement compared with other transformations. However, negative transformations such as 1t0, 3t0, 4t0, 5t0, and 5t3, have very low engagement (lower than 250). For other types of transformations that maintain post quality, such as 0t0, 1t1, 3t3, and 5t5, we see relatively low engagement. However, for high levels (5t5), the engagement is high.

We observed the learning sequences for the different types of learners. The learning sequences tended to form a loop for learners with positive transformation. For learners with negative transformation, learning sequences tended to be simple: they viewed the content, commented, and, then, checked their evaluation. Learning was driven almost entirely by the evaluation module. The results are shown in Fig. 7 and help explain which types of learners could achieve better results. As learners with positive transformation have rich learning behavior and high learning engagement, they promote their knowledge construction level.

This result demonstrates the hypothesis that learners' knowledge construction level transformation may be caused by a change in learning behaviors, especially high-level learning behaviors. Differences in the behavioral mode between positive and negative transformation learners could also be an essential factor that promotes learners' knowledge construction levels. We can see a "Learn and form the personal understanding (learn)—Comment and interact with others (practice)—Evaluate and reflect (reflect)—Learn and improve by taking part in further activities (restructure)" circle which reflects motivated learning. Similar circles have been shown to effectively promote learning by many researchers [81]–[83]. After several iterations of the circle, the learners had higher knowledge construction levels. This result is

TABLE V
DIFFERENCE OF LEARNERS' LEARNING BEHAVIORS BEFORE AND AFTER USING THE SCAFFOLDING

Category	Behavior	Without Scaffolding (percent%)	Total	With Scaffolding	Total
Constructive	Comment	872 (2.34%)	8.41%	2095(3.13%)	14.90%
	Post	1108(2.98%)		2219(3.31%)	
	participate	1149(3.09%)		5674(8.46%)	
Active	mykoevaluate	3949(10.62%)	31.21%	12717(18.97%)	34.02%
	Profile	653(1.76%)		115(0.17%)	
	Download	7004(18.83%)		9975(14.88%)	
Passive	koevaluate	3950(10.62%)	60.38%	11474(17.12%)	51.08%
	View	18506(49.76%)		22764(33.96%)	
Sum		37191		67033	

TABLE VI
STATISTICS OF EACH BEHAVIOR

Transformation	Comment	Post	Participate	Download	mykoevaluate	Profile	View	koevaluate	Scaled result
1t1	8.5	12.5	23	59.5	67.5	1	137	53.5	**248**
1t3	12.2	14	47.8	36.8	116	1.4	160	82.6	**372**
1t5	19.9	19.6	50.6	190	90.1	1.5	181	89.3	**449**
1t0	8.5	6.5	24	47.7	49.3	1.33	103	48.2	**212**
3t1	15.3	16.2	23	74.8	77.8	1.33	149	76.8	**313**
3t2	2	2	0	21	0	1	68	0	**25**
3t3	8.3	7.8	7.5	59	35	0.8	78	31	**140**
3t4	20	23	65	71	253	1	292	245	**833**
3t5	16.1	19.6	39.9	49.9	96.4	1.05	153	76.7	**352**
3t0	8.8	13.5	20.4	31.3	50.3	1.1	112	50.5	**219**
4t3	13.7	25.7	83	44	192	1.33	253	163	**674**
4t5	24	32	53.2	53.4	98	1.2	175.2	114.8	**486.4**
4t0	6	4	7	17	27	1	59	28	**105**
5t1	22	27	29	42	142	1	230	165	**550**
5t3	6	6	4.5	9	12	2.5	101	14	**76**
5t5	26.14	19.29	54.71	54	150.9	1.143	179.1	121	**494.5**
5t0	8	10	11	7	73	1	96	105	**298**
0t1	15.2	13.2	37.6	29.4	93.1	1.11	140	88.7	**349**
0t3	14.5	16.6	35.7	57.6	83	0.95	148	80	**342**
0t4	17.4	23.6	34.4	49.8	87.8	0.8	143	64	**319**
0t5	17.7	17.2	38.1	80	106	1.42	170	100	**408**
0t0	9	7.06	11	21.9	44.6	1	81.3	42.8	**160**

in accordance with the aims of principles (3) and (5), which reflect learners' current states and suggest useful guidance for improving behaviors and promoting the transformation of knowledge.

C. Role of Scaffolding in Promoting Course Completion Rates

The course completion rate is one of the most critical measurements for online courses and can also reflect the influence of automatic feedback scaffolding. The course was composed of 19 lessons (learning cells) in this article, and each had its evaluation module. After each lesson, we determined whether they were completed by comparing the results with the qualifying score for the lesson. Ten lessons were available for learners in the previous two weeks. After the introduction of scaffolding, nine other lessons were published. We collected the learning data for the ten lessons over one to two weeks to measure learning without scaffolding. When scaffolding was included, we collected the new learners' learning state as the comparison. These new learners conducted all their learning with the support of scaffolding.

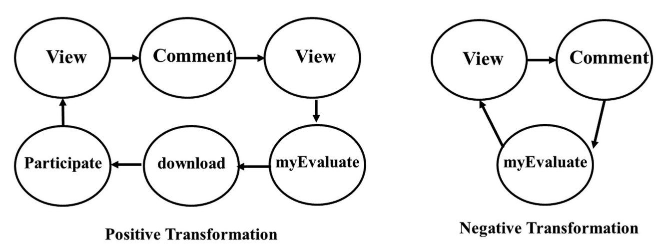


Fig. 7. Comparison of learning behavior for different transformation.

Table VII shows the difference in completion rate between those learning with and without scaffolding. Learners without scaffolding had 1562 lessons and only 423 were completed (completion rate: 27.1%). However, for learners with scaffolding, 1292 lessons were learned and 747 completed (completion rate: 57.8%). The rate was significantly higher when scaffolding was used. To make the results more convincing, we compared the last year's course completion rate. As the last year's course was the first round course and the topics were totally new to the learners, it attracted more attention. After learning, the course completion rate reached 47.52%. In

TABLE VII
STATISTICS OF THE COMPLETION RATE OF THE COURSE

	Totally Without the scaffolding	With the scaffolding all the time
Number of Learners	181	68
Total Number of Learnt Learning Cells	1563	1292
Total Number of Learning Cells which completed learning	423	747
Complete rate	27.1%	57.8%
Average score	25.66	57.18
Duration	5600s	4319s
Complete rate of the whole course after the course ended	81/181	40/68

TABLE VIII
RELATION BETWEEN KNOWLEDGE CONSTRUCTION LEVEL TRANSFORMATION AND COURSE COMPLETION RATE

	Negative	Keep	Positive	5t5	0t0
Number of learners	30	29	128	7	16
Completed learners	12	15	91	7	5
Completion rate	40%	51.72%	71.09%	100%	31.25%
Average score	37.442	49.228	65.302	84.687	36.011

addition, we can see that the course completion rate with scaffolding was also higher than in the last year's course. This means that scaffolding could help improve course completion. The average score obtained by learners indicated that those with automatic feedback scaffolding achieved better scores (57.18 versus 25.66) than learners without scaffolding. For lesson duration, learners with scaffolding used less time (4319 s versus 5600 s) and achieved higher scores. This means that scaffolding can promote learning efficiency. The completion rate of the whole course (all 19 lessons) was higher for learners with scaffolding.

Scaffolding promoted completion of the online course. The following two reasons may have contributed to this result.

1) In the design of scaffolding, we provided timely feedback to learners. This not only gave learners specific guidance but also made them feel noticed. If learners strayed off-topic, it guided them back. Learners engaged more effectively with learning content and were more likely to complete the course. This is in accordance with other research on MOOCs [84], [85].
2) Scaffolding tagged all posts, which helped new learners find information more easily. It solved the problem of learners being overwhelmed by information and unable to achieve effective interaction, ultimately helping them complete the course [86]. From these two phenomena, we can see that principles (1)–(5) ensured a positive learning retention rate.

Were there differences in course completion rates for learners undergoing different knowledge construction-level transformations? Table VIII compares the course completion rate data and the knowledge construction-level transformation data. Transformation trends were coded by a doctoral student by analyzing the changes in post quality throughout the course. These were divided into "Negative" (from higher to lower level), "Keep" (no significant change), and "Positive" (change from lower to higher level). Two transformations in the "Keep" classification are important and they are extracted: 0t0—which means learner discussions were always irrelevant and 5t5—which means learners maintained a consistently high level. Negative learner completion rates were 40%, which was lower than that of the "Keep" (51.72%) and positive groups (71.09%). The average score of these learners was also lower (37.442) than that of the learners in "Keep" (49.228) and positive (65.302) groups. Thus, we can conclude that for learners who achieved "Negative," "Keep," and "Positive" transformations, their course completion rates gradually increased and their performance was in accordance with this trend. For 0t0 learners, the completion rate was 31.25%, and the score was 36.011. These learners participated minimally in the course and were not highly engaged. The completion rate and scores of 5t5 learners were the highest (100% and 84.687), which was not surprising because they worked consistently hard and engaged deeply to maintain their knowledge construction level.

Overall, the data analysis showed that learners with negative transformation have lower learning engagement and less meaningful learning behavior. Because of less meaningful learning behavior, they could not understand the learning content deeply, and this is the reason for the negative transformation at the knowledge construction level. Learners with positive transformations actively participated in the learning process and engaged deeply. They gradually improved their understanding. Ultimately, these transformations affected the course completion rate. Positive transformations tend to enhance course completion rates and scores. Learners with negative transformations tend to achieve lower completion rates and scores. Meaningful activities and comments can affect learners' understanding and, ultimately, course completion [87]. This has important implications for course instructors in managing and guiding online learning.

The automatic feedback scaffolding proposed in this article promotes knowledge construction by improving high-level learning behaviors and engagement. With better engagement, learners can achieve a positive knowledge construction transformation. This transformation then results in better course completion rates. The results support the automation of scaffolding designed to solve specific educational challenges.

V. CONCLUSION AND FUTURE WORK

Forums are essential components of online learning but many learners cannot realize effective knowledge acquisition and support their learning without well-designed scaffolding, especially in large-scale courses. Many researchers developed useful scaffoldings in traditional or small-scale learning environments, and they have provided tips for the design of these scaffoldings [88]–[90]. They designed scaffoldings for different learning contexts to support learners in better handling the learning process [91], [92]. However, effective and systematic principles for scaffolding design in online learning have not been proposed or widely used. Moreover, large-scale courses introduce some new features, such as a large number of learners and interactions that make the discussion in a mass and ultimately affect learning outcomes. Aiming at solving the problems in large-scale courses and focusing on factors that promote highly efficient forum discourse, we proposed an automatic feedback scaffolding and its design principles. Principles for guiding online scaffolding design were formulated and used as the basis for automatic feedback scaffolding to support learners' knowledge reflection, feedback, information discovery, and learning progression in online forums. It provides systematic guidelines considering the features of large-scale online courses when designing scaffoldings in online forums, which is the main contribution based on current studies [93], [94]. Feedback rules in the scaffolding were defined based on a knowledge construction model and the functions present in the interface were designed based on the problems in online forums, such as timely feedback, knowledge organization, and supportive guidance. In real use, the machine learning model could detect the quality of learners' posts and the rules could provide automatic feedback and tagging according to post quality. We used the scaffolding in an online course with 955 learners to determine its effectiveness. Three aspects were investigated: whether the scaffolding could promote the following:

1) learners' knowledge construction level in online learning;
2) higher level learning behavior or engagement;
3) the course completion rate.

To answer research question (1), we analyzed learners' posts and found that high-level posts (P5) significantly increased from 2.4% to 39.86%, whereas low-level posts decreased from 55.86% to 18.34% over the three stages. This indicates that the scaffolding could promote learners' knowledge construction levels. This confirms the results of other researchers that instant guidance in online learning can provide learners with timely evaluation and promote the development of cognition [77], [95]. Posts were substantially transformed for learners under the guidance of scaffolding. Learners were able to rethink or reflect inspired by timely feedback or tagged posts [78], [96], [97]. This suggests that it is necessary to involve specific scaffolding to help learners understand the learning content and reduce their load. To answer research question (2), we further measured the learners' behaviors and engagements under the revised ICAP deep learning framework. Paul and Stephen [40] found that automatic feedback could help provide useful information about errors, and this helps learners improve their behavior. Similarly, in this article, we found that scaffolding induces more high-level learning behavior and promotes greater learning engagement. To understand the relationship between knowledge construction level transformation and learners' meaningful behaviors, we compared the learning behaviors for different transformations. We found that positive transformations were related to more high-level (constructive) behaviors and fewer low-level (passive) behaviors. Negative transformations were related to more low-level behaviors. Positive transformations tended to form a circle of learning behavior, which included learning, practicing, reflecting, and restructuring. More meaningful behaviors and the generation of a learning circle could explain why these learners achieved positive transformation. Compared with the current studies, this research question investigated more in-depth why specific scaffolding could help improve learners' learning. The results suggest that learners' knowledge transformation is related seamlessly to their learning behavior, especially higher level behaviors. Moreover, the results of the research question reveal that the learning mode "tell & practice" [98] cannot always work in online learning. Teachers' guidance should focus on improving learners' behavioral engagement instead of giving them simple instructions. Therefore, in designing online learning scaffoldings, researchers should know more about learners' interactions and their relationship with knowledge acquisition before involving rules in scaffolding. Finally, with the functions provided by the scaffolding, the course completion is also promoted, as demonstrated by research on MOOCs [85], [99], which answered research question (3).

This article started by collating online scaffolding design principles targeted at problems in online forums and developing automatic feedback scaffolding. The scaffolding was applied in a real context, and satisfactory results were achieved. The main contributions of this article are as follows.

1) It provides evidence that the specific design for automated scaffolding realized in real-time educational contexts can solve existing problems.
2) Automatic feedback involving guidance rules and visualized tags may guide learners from lower knowledge construction levels to higher levels, and ultimately promote better learning outcomes.
3) This article integrates the knowledge construction level with learners' meaningful behaviors, which could reflect deep learning in the course. By integrating these data, this article revealed why and how the scaffolding could promote learning and ultimately improve the course completion rate.
4) This article proposes principles for effective online scaffolding design. The influence of the scaffolding in the real learning context demonstrates the value of the principles. The most important value of this article is that it not only develops a method aiming at existing problems in large-scale course forums; it also summarizes the reasons for the problems and proposes some well-designed principles guiding the design of similar scaffoldings. By automating these principles using specific technologies, this research verified the effect of the automatic feedback scaffolding as well as the principles. In addition, the results of this article revealed the

relationship between knowledge transformation and learners' learning behaviors. This made the learning process more explicit by integrating different data that supplemented related research and benefited online learning design [100]. These aspects would be helpful for researchers and instructors to engage in better educational practices. A full understanding of potential problems and elaborate design is of core importance in online course implementation.

This article provided detailed evidence that machine learning methods can be effectively used in designing automatic feedback scaffolding to support efficient learning in online learning forums. With well-designed principles and automatic feedback, we can help promote higher level learning. This could help enrich the theoretical and practical aspects of automatic feedback in online learning. This research could be helpful for general practice in education as it provides design principles supporting automatic feedback in online courses by summarizing the features of large-scale online courses. Using the principles and the practices presented in this article, researchers and instructors can update their knowledge in scaffolding design and improve their design; they will know more about problems in online courses and how to conduct elaborate design based on learners' online interactions [101]. However, there are still limitations to this study. The principles and rubrics designed for scaffolding are relatively basic and can be refined to be more adaptive for all learners. In the future, we plan to integrate personalized recommendations that involve both resources and related users in the feedback, instead of only plain text guidance. During this process, the accuracy of automatic classification should also be improved.

REFERENCES

[1] W. Xing, X. Chen, J. Stein, and M. Marcinkowski, "Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization," *Comput. Hum. Behav.*, vol. 58, pp. 119–129, Jan. 2016, doi: 10.1016/j.chb.2015.12.007.

[2] N. Milikić, D. Gašević, and J. Jovanović, "Measuring effects of technology-enabled mirroring scaffolds on self-regulated learning," *IEEE Trans. Learn. Technol.*, vol. 13, no. 1, pp. 150–163, Jan.–Mar. 2020, doi: 10.1109/TLT.2018.2885743.

[3] A. Littlejohn, N. Hood, C. Milligan, and P. Mustain, "Learning in MOOCs: Motivations and self-regulated learning in MOOCs," *Int. Higher Educ.*, vol. 29, pp. 40–48, Apr. 2016, doi: 10.1016/j.iheduc.2015.12.003.

[4] M. Papi, A. Rios, H. Pelt, and E. Ozdemir, "Feedback-seeking behavior in language learning: Basic components and motivational antecedents," *Modern Lang. J.*, vol. 103, pp. 205–226, Sep. 2018, doi: 10.1111/modl.12538.

[5] F. R. Lin and C. M. Kao, "Mental effort detection using EEG data in E-learning contexts," *Comput. Educ.,* vol. 122, no. 7, pp. 63–79, Jul. 2018, doi: 10.1016/j.compedu.2018.03.020.

[6] A. Bey, P. Jermann, and P. Dillenbourg, "A comparison between two automatic assessment approaches for programming: An empirical study on MOOCs," *Educ. Technol. Soc.*, vol. 21, no. 3, pp. 259–272, 2018. [Online]. Available: https://www.jstor.org/stable/26388406

[7] G. Ella, A. W. Woolley, G. Pranav, and J. K. Young, "Visualized automatic feedback in virtual teams," *Frontiers Psychol.*, vol. 10, Oct. 2019, Art. no. 814, doi: 10.3389/fpsyg.2019.00814.

[8] M. Lyons, N. Aksayli, and G. Brewer, "Mental distress and language use: Linguistic analysis of discussion forum posts," *Comput. Hum. Behav.*, vol. 87, pp. 207–211, Oct. 2018, doi: 10.1016/j.chb.2018.05.035.

[9] M. Paul, C. Bellebaum, M. Ghio, B. Suchan, and O. T. Wolf, "Stress effects on learning and feedback-related neural activity depend on feedback delay," *Psychophysiology*, vol. 57, no. 2, Feb. 2020, Art. no. e13471, doi: 10.1111/psyp.13471.

[10] W. A. Donohue, Y. Liang, and D. Druckman, "Validating LIWC dictionaries: The Oslo I Accords," *J. Lang. Social Psychol.*, vol. 33, no. 3, pp. 282–301, May 2013, doi: 10.1177/0261927X13512485.

[11] A. Humphreys, R. Wang, E. Fischer, and L. Price, "Automated text analysis for consumer research," *J. Consum. Res.*, vol. 44, no. 6, pp. 1274–1306, Apr. 2018, doi: 10.1093/jcr/ucx104.

[12] R. Moore, K. Oliver, and C. Wang, "Setting the pace: Examining cognitive processing in MOOC discussion forums with automatic text analysis," *Interact. Learn. Environ.*, vol. 27, no. 5/6, pp. 655–669, Apr. 2019, doi: 10.1080/10494820.2019.1610453.

[13] A. S. Sunar, S. White, N. A. Abdullah, and H. C. Davis, "How learners' interactions sustain engagement: A MOOC case study," *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 475–487, Oct.–Dec. 2017, doi: 10.1109/TLT.2016.2633268.

[14] Y. Wang and R. Baker, "Content or platform: Why do students complete MOOCs," *MERLOT J. Online Learn. Teach.*, vol. 11, no. 1, pp. 17–30, Mar. 2015.

[15] Z. Wang, S. Gong, S. Xu, and X. Hu, "Elaborated feedback and learning: Examining cognitive and motivational influences," *Comput. Educ.*, vol. 136, pp. 130–140, Jul. 2019, doi: 10.1016/j.compedu.2019.04.003.

[16] J. Schneider, D. Boerner, P. van Rosmalen, and M. Specht, "Can you help me with my pitch? Studying a tool for real-time automated feedback," *IEEE Trans. Learn. Technol.*, vol. 9, no. 4, pp. 318–327, Oct.–Dec 2016, doi: 10.1109/TLT.2016.2627043.

[17] S. Yu, X. Yang, G. Cheng, and M. Wang, "From learning object to learning cell: A resource organization model for ubiquitous learning," *Educ. Technol. Soc.*, vol. 18, no. 2, pp. 206–224, 2015. [Online]. Available: https://www.jstor.org/stable/jeductechsoci.18.2.206

[18] P. Skehan, "A framework for the implementation of task-based instruction," *Appl. Linguistics,* vol. 17, no. 1, pp. 38–62, Mar. 1996, doi: 10.1093/applin/17.1.38.

[19] K. R. Guldberg, "Tutor roles in facilitating reflection on practice through online discussion," *Educ. Technol. Soc.*, vol. 10, no. 1, pp. 61–72, 2007. [Online]. Available: https://www.jstor.org/stable/jeductechsoci.10.1.61

[20] B. Rienties, B. Giesbers, D. Tempelaar, S. Lygo-Baker, M. Segers, and W. Gijselaers, "The role of scaffolding and motivation in CSCL," *Comput. Educ.*, vol. 59, no. 3, pp. 893–906, Nov. 2012, doi: 10.1016/j.compedu.2012.04.010.

[21] P. Beers, H. Boshuizen, P.A. Kirschner, and W.H. Gijselaers, "The analysis of negotiation of common ground in CSCL," *Learn. Instruct.*, vol. 17, no. 4, pp. 427–435, 2007, doi: 10.1016/j.learninstruc.2007.04.002.

[22] N. Fujita and C. Teplovs, "Software-based scaffolding: Supporting the development of knowledge building discourse in online courses," in *Proc. 9th Int. Conf. Learn. Sci.*, vol. 1, pp. 1048–1054, Jan. 2010.

[23] L. H. Wong, I. Boticki, J. Sun, and C. K. Looi, "Improving the scaffolds of a mobile-assisted Chinese character forming game via a design-based research cycle," *Comput. Hum. Behav.*, vol. 27, no. 5, pp. 1783–1793, Sep. 2011, doi: 10.1016/j.chb.2011.03.005.

[24] J. Meyer, K. Sanders, J. Hill, N. Koehler, G. Fyfe, and S. Fyfe, "Online assessment feedback as an instrument of reflective learning practice in human biology," *J. Comp. Hum. Biol.*, vol. 60, no. 3, p. 251, May 2009, doi: 10.1016/j.jchb.2009.02.021.

[25] E. Martinez-Camara, M. T. Martin-Valdivia, L. A. Urena-Lopez, and A. R. Montejo-Raez, "Sentiment analysis in twitter," *Natural Lang. Eng.*, vol. 20, no. 1, pp. 1–28, Jan. 2014, doi: 10.1017/S1351324912000332.

[26] S. Provoost, J. Ruwaard, W. V. Breda, H. Riper, and T. Bosse, "Validating automated sentiment analysis of online cognitive behavioral therapy patient texts: An exploratory study," *Frontiers Psychol.*, vol. 10, pp. 1–12, May 2019, doi: 10.3389/fpsyg.2019.01065.

[27] M. Wen, D. Yang, and C.P. Rosé, "Sentiment analysis in MOOC discussion forums: What does it tell us?," in *Proc. 7th Int. Conf. Educ. Data Mining*, Jan. 2014, pp. 130–137.

[28] P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Merino, I. Estévez-Ayres, and C. D. Kloos, "Sentiment analysis in MOOCs: A case study," in *Proc. IEEE Global Eng. Educ. Conf.*, Apr. 2018, pp. 1489–1496, doi: 10.1109/EDUCON.2018.8363409.

[29] B. D. Wever, T. Schellens, M. Valcke, and H. V. Keer, "Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review," *Comput. Educ.*, vol. 46, no. 1, pp. 6–28, Jan. 2006, doi: 10.1016/j.compedu.2005.04.005.

[30] B. P. De Vries, C. Cucchiarini, S. Bodnar, H. Strik, and R. van Hout, "Spoken grammar practice and feedback in an ASR-based CALL system," *Comput. Assist. Lang. Learn.*, vol. 28, no. 6, pp. 550–576, Mar. 2014, doi: 10.1080/09588221.2014.889713.

[31] G. Akçapınar, "How automated feedback through text mining changes plagiaristic behavior in online assignments," *Comput. Educ.*, vol. 87, pp. 123–130, Sep. 2015, doi: 10.1016/j.compedu.2015.04.007.

[32] V. Kovanović *et al.*, "Towards automated content analysis of discussion transcripts: A cognitive presence case," in *Proc. 6th Int. Conf. Learn. Anal. Knowl.*, Apr. 2016, pp. 15–24, doi: 10.1145/2883851.2883950.

[33] S. Joksimović, D. Gašević, V. Kovanović, B. E. Riecke, and M. Hatala, "Social presence in online discussions as a process predictor of academic performance," *J. Comput. Assist. Learn.*, vol. 31, no. 6, pp. 638–654, Jun. 2015, doi: 10.1111/jcal.12107.

[34] M. Liu, Y. Li, W. Xu, and L. Liu, "Automated essay feedback generation and its impact in the revision," *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 502–513, Oct.–Dec. 2017, doi: 10.1109/TLT.2016.2612659.

[35] J. Tsan, "Supporting elementary school computer science learning with interactive spoken dialogue agents," in *Proc. 11th Int. Conf. Int. Comput. Educ. Res.*, Jul. 2015, pp. 279–280, doi: 10.1145/2787622.2787750.

[36] S. Tegos, S. Demetriadis, and T. Tsiatsos, "A configurable conversational agent to trigger students' productive dialogue: A pilot study in the call domain," *Int. J. Artif. Intell. Educ.*, vol. 24, no. 1, pp. 62–91, 2014, doi: 10.1007/s40593-013-0007-3.

[37] I. Howley, D. Adamson, G. Dyke, E. Mayfield, J. Beuth, and P. Rose, "Group composition and intelligent dialogue tutors for impacting students' academic self-efficacy," in *Proc. Int. Conf. Intell. Tutoring Syst.*, Jul. 2012, pp. 551–556, doi: 10.1007/978-3-642-30950-2_71.

[38] J. Mu, K. Stegmann, E. Mayfield, C. Rosé, and F. Fischer, "The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions," *Int. J. Comput.-Supported. Collab. Learn.*, vol. 7, no. 2, pp. 285–305, May 2012, doi: 10.1007/s11412-012-9147-y.

[39] N. Verdú and J. Sanuy, "The role of scaffolding in CSCL in general and in specific environments," *J. Comput. Assist. Learn.*, vol. 30, no. 4, pp. 337–348, Dec. 2013, doi: 10.1111/jcal.12047.

[40] O. Paul and M. Stephen, "Feedback alignment: Effective and ineffective links between tutors' and students' understanding of coursework feedback," *Assessment. Eval. Higher. Educ.*, vol. 36, no. 2, pp. 125–136, Feb. 2011, doi: 10.1080/02602930903201651.

[41] X. Wang, D. Yang, M. Wen, K. R. Koedinger, and C. P. Rosé, "Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains," in *Proc. 8th Int. Conf. Educ. Data. Mining*, Jun. 2015, pp. 226–233.

[42] P. Y. A. Wang and H. C. B. Yang, "Using collaborative filtering to support college students' use of online forum for English learning," *Comput. Educ.*, vol. 59, no. 2, pp. 628–637, Sep. 2012, doi: 10.1016/j.compedu.2012.02.007.

[43] K. Dave, M. Wattenberg, and M. Muller, "Flash forums and forumReader: Navigating a new kind of large-scale online discussion," in *Proc. ACM Conf. Comput. Supported. Cooperative Work*, Nov. 2004, pp. 232–241, doi: 10.1145/1031607.1031644.

[44] L. Muchnik, S. Aral, and S. J. Taylor, "Social influence bias: A randomized experiment," *Science*, vol. 341, no. 6146, pp. 647–651, Aug. 2013, doi: 10.1126/science.1240466.

[45] D. Bodemer and J. Dehler, "Group awareness in CSCL environments," *Comput. Hum. Behav.*, vol. 27, no. 3, pp. 1043–1045, May 2011, doi: 10.1016/j.chb.2010.07.014.

[46] J. Dehler, D. Bodemer, B. Jürgen, and F. W. Hesse, "Guiding knowledge communication in cscl via group knowledge awareness," *Comput. Hum. Behav.*, vol. 27, no. 3, pp. 1068–1078, May 2011, doi: 10.1016/j.chb.2010.05.018.

[47] G. Stahl, *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA, USA: MIT Press, 2006, doi: 10.7551/mitpress/3372.001.0001.

[48] B. Jürgen, C. Schwind, A. Rudat, and D. Bodemer, "Selective reading of large online forum discussions: The impact of rating visualizations on navigation and learning," *Comput. Hum. Behav.*, vol. 44, pp. 191–201, Mar. 2015, doi: 10.1016/j.chb.2014.11.043.

[49] E. G. Cohe, "Restructuring the classroom: Conditions for productive small groups," *Rev. Educ. Res.*, vol. 64, pp. 1–35, Mar. 1994, doi: 10.3102/00346543064001001.

[50] C. P. Rosé, D. Bhembe, S. Siler, R. Srivastava, and K. VanLehn, "Exploring the effectiveness of knowledge construction dialogues," in *Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies*, U. Hoppe, F. Verdejo, and J. Kay, Eds. Amsterdam, The Netherlands: IOS Press, 2003, pp. 497–499.

[51] S. Butchart, J. Bigelow, G. Oppy, D. Forster, and I. Gold, "Improving critical thinking using web based argument mapping exercises with automated feedback," *Australas. J. Educ. Technol.*, vol. 25, no. 2, pp. 268–291, May 2009, doi: 10.14742/ajet.1154.

[52] R. T. Kellogg, A. P. Whiteford, and T. Quinlan, "Does automated feedback help students learn to write?," *J. Educ. Comput. Res.*, vol. 42, no. 2, pp. 173–196, Feb. 2010, doi: 10.2190/EC.42.2.c.

[53] A. Weinberger, M. Reiserer, B. Ertl, F. Fischer, and H. Mandl, "Facilitating collaborative knowledge construction in computer-mediated learning environments with cooperation scripts," in *Barriers and Biases in Computer-Mediated Knowledge Communication, Computer-Supported Collaborative Learning Series*, vol. 5, R. Bromme, F.W. Hesse, and H. Spada, Eds. Boston, MA, USA: Springer, 2005, doi: 10.1007/0-387-24319-4_2.

[54] P. J. Beers, H. P. Boshuizen, P. A. Kirschner, and W. H. Gijselaers, "The analysis of negotiation of common ground in CSCL," *Learn. Instruct.*, vol. 17, no. 4, pp. 427–435, Aug. 2007, doi: 10.1016/j.learninstruc.2007.04.002.

[55] D. Yang, D. Adamson, and C. P. Rosé, "Question recommendation with constraints for massive open online courses," in *Proc. 8th ACM. Conf. Recommender Syst.*, Oct. 2014, pp. 49–56, doi: 10.1145/2645710.2645748.

[56] M. Klusener and A. Fortenbacher, "Predicting students' success based on forum activities in MOOCs," in *Proc. IEEE 8th Int. Conf. Intell. Data Acquisition Adv. Comput. Syst. Technol. Appl.*, Sep. 2015, pp. 925–928, doi: 10.1109/IDAACS.2015.7341439.

[57] A. F. Wise and Y. Cui, "Learning communities in the crowd: Characteristics of content related interactions and social relationships in MOOC discussion forums," *Comput. Educ.*, vol. 122, pp. 221–242, Jul. 2018, doi: 10.1016/j.compedu.2018.03.021.

[58] G. Athanasopoulos and R. J. Hyndman, "The value of feedback in forecasting competitions," *Int. J. Forecasting*, vol. 27, no. 3, pp. 845–849, Jul. 2011, doi: 10.1016/j.ijforecast.2011.03.002.

[59] K. S. Hone and G. R. E. Said, "Exploring the factors affecting MOOC retention: A survey study," *Comput. Educ.*, vol. 98, pp. 157–168, Jul. 2016, doi: 10.1016/j.compedu.2016.03.016.

[60] J. C. Richardson, A. A. Koehler, E. D. Besser, S. Caskurlu, J. Lim, and C. M. Mueller, "Conceptualizing and investigating instructor presence in online learning environments," *Int. Rev. Res. Open Distrib. Learn.*, vol. 16, no. 3, pp. 256–297, Jun. 2015, doi: 10.19173/irrodl.v16i3.2123.

[61] A. F. Wise, Y. Cui, W. Jin, and J. Vytasek, "Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling," *Internet Higher Educ.*, vol. 32, pp. 11–28, Jan. 2017, doi: 10.1016/j.iheduc.2016.08.001.

[62] C. G. Brinton, M. Chiang, S. Jain, H. Lam, Z. Liu, and F. M. F. Wong, "Learning about social learning in MOOCs: From statistical analysis to generative model," *IEEE Trans. Learn. Technol.*, vol. 7, no. 4, pp. 346–359, Oct.–Dec. 2014, doi: 10.1109/TLT.2014.2337900.

[63] J. Kim, "Influence of group size on students' participation in online discussion forums," *Comput. Educ.*, vol. 62, pp. 123–129, Mar. 2013, doi: 10.1016/j.compedu.2012.10.025.

[64] A. F. Wise, F. Marbouti, Y. T. Hsiao, and S. Hausknecht, "A survey of factors contributing to learners' 'listening' behaviors in asynchronous online discussions," *J. Educ. Comput. Res.*, vol. 47, no. 4, pp. 461–480, Jan. 2013, doi: 10.2190/EC.47.4.f.

[65] F. Fischer, I. Kollar, K. Stegmann, and C. Wecker, "Toward a script theory of guidance in computer-supported collaborative learning," *Educ. Psychol.*, vol. 48, no. 1, pp. 56–66, Jan. 2013, doi: 10.1080/00461520.2012.748005.

[66] C. N. Gunawardena, C. A. Lowe, and T. Anderson, "Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing," *J. Educ. Comput. Res.*, vol. 17, no. 4, pp. 397–431, Dec. 1997, doi: 10.2190/7MQV-X9UJ-C7Q3-NRAG.

[67] H. T. Hou and S. Y. Wu, "Analyzing the social knowledge construction behavioral patterns of an online synchronous collaborative discussion instructional activity using an instant messaging tool: A case study," *Comput. Educ.*, vol. 57, no. 2, pp. 1459–1468, Sep. 2011, doi: 10.1016/j.compedu.2011.02.012.

[68] R. D. Roscoe and M. T. H. Chi, "Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions," *Rev. Educ. Res.*, vol. 77, no. 4, pp. 534–574, Dec. 2017, doi: 10.3102/0034654307309920.

[69] R. D. Roscoe, "Self-monitoring and knowledge-building in learning by teaching," *Instructional Sci.*, vol. 42, no. 3, pp. 327–351, Jun. 2013, doi: 10.1007/s11251-013-9283-4.

[70] H. Martin, M. Marion, and R. Berger, "Cross-age tutoring: How to promote tutees' active knowledge-building," *Educ. Psychol.,* vol. 38, no. 7, pp. 915–926, Feb. 2018, doi: 10.1080/01443410.2018.1444734.

[71] S. Yu, J. Duan, and J. Cui, "Double helix deep learning model based on learning cell," in *Blended Learning: Educational Innovation for Personalized Learning*. (Lecture Notes in Computer Science), vol. 11546, S. Cheung, L. K. Lee, I. Simonova, T. Kozel, and L. F. Kwok, Eds. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-21562-0_3.

[72] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," *Educ. Psychol.*, vol. 49, no. 4, pp. 219–243, Oct. 2014, doi: 10.1080/00461520.2014.965823.

[73] X. Wang, M. Wen, and C. P. Rosé, "Towards triggering higher-order thinking behaviors in MOOCs," in *Proc. 6th Int. Conf. Learn. Anal. Knowl.*, 2016, pp. 398–407, doi: 10.1145/2883851.2883964.

[74] R. Lam and K. Muldner, "Manipulating cognitive engagement in preparation-to-collaborate tasks and the effects on learning," *Learn. Instruct.*, vol. 52, pp. 90–101, Dec. 2017, doi: 10.1016/j.learninstruc.2017.05.002.

[75] R. Bakeman and V. Quera, *Sequential Analysis and Observational Methods for the Behavioral Sciences*. Cambridge, New York, USA: Cambridge Univ. Press, 2011, doi: 10.1017/CBO9781139017343.

[76] L. Xu, F. Wang, and B. Yu, "Social network analysis of MOOC learners' knowledge building," in *Mobile and Ubiquitous Learning*, S. Yu, M. Ally, and A. Tsinakos, Eds. Singapore: Springer, 2018, pp. 363–377, doi: 10.1007/978-981-10-6144-8_21.

[77] F. Lin and C. K. Chan, "Promoting elementary students' epistemology of science through computer-supported knowledge-building discourse and epistemic reflection," *Int. J. Sci. Educ.*, vol. 40, no. 6, pp. 668–687, Feb. 2018, doi: 10.1080/09500693.2018.1435923.

[78] D. Verpoorten, W. Westera, and M. Specht, "Using reflection triggers while learning in an online course," *Brit. J. Educ. Technol.*, vol. 43, no. 6, pp. 1030–1040, Nov. 2012, doi: 10.1111/j.1467-8535.2011.01257.x.

[79] T. W. Bickmore, L. M. Pfeifer, and D. Schulman, "Relational agents improve engagement and learning in science museum visitors," in *Proc. Intell. Virtual Agents Int. Conf.*, Sep. 2011, pp. 55–67, doi: 10.1007/978-3-642-23974-8_7.

[80] M. C. Duffy and R. Azevedo, "Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system," *Comput. Hum. Behav.*, vol. 52, no. 14, pp. 338–348, Nov. 2015, doi: 10.1016/j.chb.2015.05.041.

[81] T. A. Daher and K. A. Kiewra, "An investigation of soar study strategies for learning from multiple online resources," *Contemporary Educ. Psychol.*, vol. 46, pp. 10–21, Jul. 2016, doi: 10.1016/j.cedpsych.2015.12.004.

[82] Q. Wang, G. Ding, and S. Yu, "Crowdsourcing mode-based learning activity flow approach to promote subject ontology generation and evolution in learning," *Interact. Learn. Environ.*, vol. 27, no. 7, pp. 965–983, Aug. 2019, doi: 10.1080/10494820.2018.1509875.

[83] G. J. Hwang and S. Y. Wang, "Single loop or double loop learning: English vocabulary learning performance and behavior of students in situated computer games with different guiding strategies," *Comput. Educ.*, vol. 102, pp. 188–201, Nov. 2016, doi: 10.1016/j.compedu.2016.07.005.

[84] T. Sancho-Vinuesa, N. Escudero-Viladoms, and R. Masià, "Continuous activity with immediate feedback: A good strategy to guarantee student engagement with the course," *Open Learn.*, vol. 28, no. 1, pp. 51–66, Mar. 2013. doi: 10.1080/02680513.2013.776479.

[85] O. Almatrafi, A. Johri, and H. Rangwala, "Needle in a haystack: Identifying learner posts that require urgent response in mooc discussion forums," *Comput. Educ.*, vol. 118, pp. 1–9, Mar. 2018, doi: 10.1016/j.compedu.2017.11.002.

[86] A. F. Wise, Y. Cui, and J. Vytasek, "Bringing order to chaos in MOOC discussion forums with content-related thread identification," *Proc. 6th Int. Conf. Learn. Anal. Knowl.*, Apr. 2016, pp. 188–197, doi: 10.1145/2883851.2883916.

[87] B. K. Pursel, L. Zhang, K. W. Jablokow, G. W. Choi, and D. Velegol, "Understanding MOOC students: Motivations and behaviours indicative of MOOC completion," *J. Comput. Assist. Learn.,* vol. 32, no. 3, pp. 202–217, Mar. 2016, doi: 10.1111/jcal.12131.

[88] S. Ak, "The role of technology-based scaffolding in problem-based online asynchronous discussion," *Brit. J. Educ. Technol.*, vol. 47, no. 4, pp. 680–693, Mar. 2015, doi: 10.1111/bjet.12254.

[89] W. Fernando, "Show me your true colours: Scaffolding formative academic literacy assessment through an online learning platform," *Assessing Writing*, vol. 36, pp. 63–76, Apr. 2018, doi: 10.1016/j.asw.2018.03.005.

[90] O. Levrini, M. Levin, P. Fantini, and G. Tasquier, "Orchestration of classroom discussions that foster appropriation," *Sci. Educ.*, vol. 103, no. 1, pp. 206–235, Nov. 2018, doi: 10.1002/sce.21475.

[91] T. De Jong, "Moving towards engaged learning in stem domains; there is no simple answer, but clearly a road ahead," *J. Comput. Assist. Learn.*, vol. 35, no. 2, pp. 153–167, Jan. 2019, doi: 10.1111/jcal.12337.

[92] H. Meij, S. Veldkamp, and H. Leemkuil, "Effects of scripting on dialogues, motivation and learning outcomes in serious games," *Brit. J. Educ. Technol.*, vol. 51, no. 2, pp. 459–472, Jul. 2019, doi: 10.1111/bjet.12851.

[93] M. Liu, W. Zou, Y. Shi, Z. Pan, and C. Li, "What do participants think of today's MOOCs: An updated look at the benefits and challenges of MOOCs designed for working professionals," *J. Comput. Higher Educ.*, vol. 32, pp. 307–329, Aug. 2019, doi: 10.1007/s12528-019-09234-x.

[94] A. Sari, C. Bonk, and M. Zhu, "MOOC instructor designs and challenges: What can be learned from existing MOOCs in Indonesia and Malaysia?," *Asia Pac. Educ. Rev.*, vol. 21, no. 1, pp. 143–166, Mar. 2020, doi: 10.1007/s12564-019-09618-9.

[95] P. H. Wu, G. J. Hwang, M. Milrad, H. R. Ke, and Y. M. Huang, "An innovative concept map approach for improving students' learning performance with an instant feedback mechanism," *Brit. J. Educ. Technol.*, vol. 43, no. 2, pp. 217–232, Mar. 2011, doi: 10.1111/j.1467-8535.2010.01167.x.

[96] R. Carter, Y. Salamonson, L. Ramjan, and E. Halcomb, "Students use of exemplars to support academic writing in higher education: An integrative review," *Nurse Educ. Today*, vol. 65, pp. 87–93, Jun. 2018, doi: 10.1016/j.nedt.2018.02.038.

[97] T. C. Hsu, "Behavioural sequential analysis of using an instant response application to enhance peer interactions in a flipped classroom," *Interact. Learn. Environ.*, vol. 26, no. 1, pp. 91–105, Mar. 2017, doi: 10.1080/10494820.2017.1283332.

[98] L. Schalk, R. Schumacher, A. Barth, and E. Stern, "When problem-solving followed by instruction is superior to the traditional tell-and-practice sequence," *J. Educ. Psychol.*, vol. 110, no. 4, pp. 596–610, May 2018, doi: 10.1037/edu0000234.

[99] W. Wang, L. Guo, L. He, and Y. J. Wu, "Effects of social-interactive engagement on the dropout ratio in online learning: Insights from MOOC," *Behav. Inf. Technol.*, vol. 38, no. 6, pp. 621–636, Jun. 2019, doi: 10.1080/0144929X.2018.1549595.

[100] B. Wu, Y. Hu, and M. Wang, "Scaffolding design thinking in online STEM preservice teacher training," *Brit. J. Educ. Technol.*, vol. 50, no. 5, pp. 2271–2287, Sep. 2019, doi: 10.1111/bjet.12873.

[101] D. Hernandez-Leo, R. Martinez-Maldonado, A. Pardo, J. Munoz-Cristobal, and M. Rodriguez-Triana, "Analytics for learning design: A layered framework and tools," *Brit. J. Educ. Technol.*, vol. 50, no. 1, pp. 139–152, Jul. 2018, doi: 10.1111/bjet.12645.

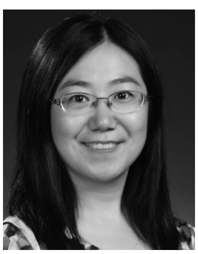
Qi Wang was born in Shandong, China, in 1990. He received the Ph.D. degree in educational technology from Beijing Normal University, Beijing, China, in 2020.

He is currently a Lecturer with Artificial Intelligence and Human Languages Lab, Beijing Foreign Studies University, Beijing, China. His research interests include technology-enhanced learning, AI in education, learning analytics, and automatic design and generation of online courses.

Carolyn P. Rose (Senior Member, IEEE) was born in San Jose, CA, USA, in 1970. She received the B.S. degree in information and computer science from the University of California at Irvine, Irvine, CA, in 1992, and the M.S. degree in computational linguistics and the Ph.D. degree in information and language technologies from Carnegie Mellon University, Pittsburgh, PA, USA, in 1994 and 1997, respectively.

She is currently a Professor of Language Technologies and Human-Computer Interaction with the School of Computer Science, Carnegie Mellon University and the Interim Director of the Language Technologies Institute. Her research interests include computational modeling of discourse to enable scientific understanding the social and pragmatic nature of conversational interaction of all forms, and using this understanding to build intelligent computational systems for improving collaborative interactions. Her research group's highly interdisciplinary work, published in more than 270 peer-reviewed publications, is represented in the top venues of five fields, namely language technologies, learning sciences, cognitive science, educational technology, and human-computer interaction, with awards in four of these fields.

Dr. Rosé is a Past President and Inaugural Fellow of the International Society of the Learning Sciences, Founding Chair of the International Alliance to Advance Learning in the Digital Era, and the Co-Editor-In-Chief of the *International Journal of Computer-Supported Collaborative Learning*. She also a 2020–2021 AAAS Leshner Leadership Institute Fellow for Public Engagement with Science, with a focus on public engagement with artificial intelligence.



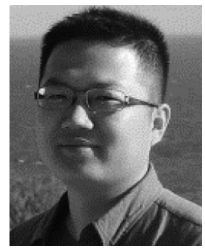
Ning Ma was born in Hebei, China, in 1977. She received the Ph.D. degree in education from Beijing Normal University, Beijing, China, in 2006.

Her major study focused on technology enhanced learning and technology enhanced teacher development. She is currently an Associate Professor with the Faculty of Education, Beijing Normal University. She has been working in the technology enhanced learning for more than 20 years. She participated and conducted many research projects, some of which are very influential in China, such as the Project of Theoretical and Practical Innovative Experimental Research on Leaping Development of Fundamental Education, which has more than 400 experimental schools in China, Theoretical and Practical Research on Propelling the High Quality and Balanced Development of Regional Compulsory Education, which cares about how the K-12 schools in rural areas can get high quality educations, and Cloud Classrooms of Mobile Learning supported by the Joint Laboratory of the Ministry of Education and China Mobile Ltd., which aims to develop 100 K-12 schools where in some classes each student can learn with a tablet computer or other mobile devices. Her research interests include one-to-one digital learning, mobile learning, and technology enhanced in-service teacher development. Till now, she has authored or coauthored more than 80 articles and nearly 10 books in these research areas.

Prof. Ma is currently the Deputy Director of the Mobile Learning—the Joint Laboratory of the Ministry of Education and China Mobile Ltd., and a Core Fellow of the Advanced Innovation Center for Future Education, Beijing Normal University. She was a Fellow of the Project of Promoting the K-12 Teachers' Abilities of Applying ICT in Education, which was sponsored and conducted by the Chinese Ministry of Education that aims to train the over 10 million K-12 teachers of China.

Shiyan Jiang received the Ph.D. degree in technology-enhanced STEM education from the University of Miami, Coral Gables, FL, USA, in 2018.

She was a Postdoc with Carnegie Mellon University, Pittsburgh, PA, USA, and worked in designing an innovative writing feedback system. She is currently an Assistant Professor of Learning Design and Technology with North Carolina State University, Raleigh, NC, USA. She designs and studies technology-enhanced data science learning environments to facilitate the development of STEM identities.



Haogang Bao received the Ph.D. degree in educational technology from Faculty of Education, Beijing Normal University, Beijing, China, in 2021.

He is currently an Assistant Researcher with the National Institute of Education Sciences, Beijing, China. His research interests include learning analytics computer-supported collaborative learning, and artificial intelligence in education.

Yanyan Li received the Ph.D. degree in computer software and theory from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

She is currently a Professor with the Faculty of Education, Beijing Normal University, Beijing, China. Her research interests include learning analytics, computer supported collaborative learning, STEM, and AI in education. She is the Principal Investigator of several projects funded by the National Natural Science Foundation of China, National Social Science foundation of China, and Key Project of Beijing Educational Science "Thirteenth Five-Year Plan." She has authored or coauthored three academic books and more than 100 articles both in journals and international conferences covering interdisciplinary areas, e.g., computer science, learning science, and educational studies.

Dr. Li is the current executive editor of *Journal of Computers in Education* and was an editorial board member of the *International Journal of Distance Education Technologies* and *Journal of Smart Learning Environments*. She was the invited keynote speaker of the 19th Artificial Intelligence in Education Conference, London, U.K., held in June 2018, and the 13th International Conference on Education Technology and Computers, Wuhan, China, held in October 2021. Four of her articles have been awarded as the best paper in international conferences.