EDUCATIONAL CHATBOTS FOR COLLABORATIVE LEARNING: RESULTS OF A DESIGN EXPERIMENT IN A MIDDLE SCHOOL

Michael Burkhard[1], Sabine Seufert[1], Matthias Cetto[2] and Siegfried Handschuh[2]
[1]Institute for Educational Management and Technologies, St.Jakob-Strasse 21, University of St.Gallen, Switzerland
[2]Institute of Computer Science, Rosenbergstrasse 30, 9000 St.Gallen, Switzerland

ABSTRACT

Educational chatbots promise many benefits for teaching and learning. Although chatbot use cases in this research field are rapidly growing, most studies focus on individual users rather than on collaborative group settings. To address this issue, this paper investigates how chatbot-mediated learning can be designed to foster middle school students in team-based assignments. Using an educational design research approach, quality indicators of educational chatbots were derived from the literature, which served as a guideline for the development of the chatbot Tubo (meaning tutoring bot). Tubo is part of a web-based team learning environment in which students can chat with each other and collaboratively work on their group assignments. As a team member and tutor of each group, Tubo guides the students through the learning journey by different scaffolding elements and helps with content-related questions the students have. As part of a first design cycle, the chatbot application was tested with a school class of a technical vocational school in Switzerland. The received feedback suggests that the approach of team-based learning with chatbots has a lot of potential from the students' and teachers' point of view. However, the role distribution of the individual group members may have to be further specified to address the different needs of autonomous as well as more control-oriented students.

KEYWORDS

Educational Chatbots, Smart Personal Assistant Technology, Team-Based Learning, Computer-Supported Collaborative Learning, Scaffolding

1. INTRODUCTION

The use of artificial intelligence (AI) in education is growing rapidly (Okonkwo & Ade-Ibijola, 2021). Among the most popular AI technologies to support teaching and learning activities are chatbots (Okonkwo & Ade-Ibijola, 2021). Chatbots are "computer programs which attempt to simulate conversations of human beings via text or voice interactions" (Winkler & Söllner, 2018). They can interact with multiple students at the same time and be used as *a scaffolding device* to guide students through learning processes (Duffy & Azevedo, 2015; Winkler et al., 2021). Chatbots can further trigger metacognitive thinking processes and increase student motivation (Winkler & Söllner, 2018). Educational chatbots have already been used in various areas and hold multifaceted potential (see e.g., Adamopoulou & Moussiades, 2020; Okonkwo & Ade-Ibijola, 2021; Rapp et al., 2021). However, so far, most studies have focused on individual users rather than on collaborative group settings (see e.g., Okonkwo & Ade-Ibijola, 2021; Rapp et al., 2021). A few examples that address the issue of *chatbot-mediated learning in team-based scenarios*, are David et al. (2019), Kumar (2021) and Winkler et al. (2019). These studies suggest, that chatbots could improve collaboration quality and learning performance in group projects.

However, effective collaboration between group members is challenging and requires guidelines to be effective (Briggs et al., 2013; Nussbaum et al., 2009; Winkler et al., 2019). In addition, it is difficult or even impossible for the teacher to provide guidance and feedback on multiple fronts simultaneously (Jeong et al., 2019). Furthermore, the meta-analysis of Hattie (2009, p. 297) shows that the provision of *formative evaluation* and *feedback* as well as *teacher clarity* are among the most important factors that influence student performance. As a virtual tutor and first-level support, a chatbot could provide additional support for group work and relieve the teacher of some of the workload. Therefore, in light of the identified research desideratum, the following research question should be addressed:

How can we design educational chatbots to foster team-based learning of middle school students?

The objectives of the paper at hand are therefore twofold:

- Elaboration on the quality indicators for developing educational chatbots to support team-based learning in order to provide a theoretical foundation for the design and development of a chatbot within the design experiment;
- design and evaluation of a chatbot that uses scaffolding elements to support student groups in working independently with groupwork assignments.

From a theoretical point of view, the paper at hand contributes to a better understanding of the usefulness of educational chatbots in team-based learning scenarios. From a practical standpoint, the paper aims to contribute to how chatbot-mediated learning can be designed and implemented in the classroom. The quality indicators described may serve as a guideline for other researchers who want to implement similar projects and uncover further potential of the technology in more detail.

To this end, section 2 gives an overview of the theoretical background and elaborates on the quality indicators for developing educational chatbots. Section 3 sheds light on the applied research design and method. Section 4 describes the designed artifact: the chatbot Tubo (meaning tutoring bot) for groupwork assignments. In section 5, the classroom experiment with chatbot Tubo is summarized as part of a formative evaluation. The implications of the experiment are discussed in section 6, while section 7 concludes with some final remarks and gives an outlook on future research areas.

2. THEORETICAL BACKGROUND

2.1 Collaborative Learning & Computer Supported Collaborative Learning

Collaborative learning (CL) is a pedagogical approach where groups of learners work together to solve a problem, complete a task or create a product (Laal & Ghodsi, 2012). CL can have social, psychological, and academic benefits (Laal & Ghodsi, 2012, p. 487). For CL to be effective, Dillenbourg and Schneider (1995, p. 133) consider it important that the groups should be rather small, because in large groups some members tend to be absent. In this context, it is the task of the teacher (or potentially a chatbot) to monitor the CL activities an “take care that no learner is left out of the interaction” (Dillenbourg & Schneider, 1995, p. 134). Under these conditions, different underlying mechanisms can make CL effective (Dillenbourg & Schneider, 1995, pp. 136-141). Such mechanisms may involve the discussion of divergent opinions (*disagreement*), the *(self-)explanation* of content to other group members or the justification of one’s own point of view (*mutual regulation*) (Dillenbourg & Schneider, 1995, pp. 136-141).

One way of addressing the difficulties of solving complex collaborative tasks may be through technology (Jeong et al., 2019). *Computer-supported collaborative learning* (CSCL) investigates how information and communication technology can support learning in groups (Ludvigsen & Mørch, 2010). Among other things, CSCL can support learners through *scaffolding* which attempts to provide guidance along the learning journey (see e.g., Miller & Hadwin, 2015; Splichal et al., 2018). The concept of *scaffolding* goes back to the work of Wood et al. (1976), who defined scaffolding as the process that enables a novice to solve a problem that would be otherwise impossible. Scaffolding can occur through a variety of ways (e.g., hints, prompts, feedback, illustrations, interactive features) (Duffy & Azevedo, 2015). In computer-based environments, pedagogical agents such as chatbots can promote effective learning by providing instructional scaffolds depending on learners' behavior and progress (Duffy & Azevedo, 2015).

2.2 Chatbot-Mediated Learning

Chatbots can be defined as “computer programs which attempt to simulate conversations of human beings via text or voice interactions” (Winkler & Söllner, 2018). They can interact with multiple learners at the same time and guide them through learning processes (e.g., through scaffolding) (Duffy & Azevedo, 2015; Winkler et al., 2021).

Regarding CL, a chatbot could help to improve learning within the group. For example, a chatbot could act as a scaffolding device and content expert to provide assistance during the tasks. In this role, the chatbot could structure the work process, answer questions about the content and provide formative feedback.

In education, chatbots have already been used in different areas (see e.g., Adamopoulou & Moussiades, 2020; Okonkwo & Ade-Ibijola, 2021; Rapp et al., 2021). However, so far, most studies have focused on individual users rather than on collaborative group settings. A few examples that address the issue of *chatbot-mediated learning in team-based scenarios*, are David et al. (2019), Kumar (2021) and Winkler et al. (2019). David et al. (2019) used chatbots as teaching assistants in group settings to support the teacher and improve classroom orchestration. Kumar (2021) investigated how chatbots can facilitate team-based projects and concluded that chatbots can improve teamwork and learning performance. Winkler et al. (2019) developed a chatbot with the aim of improving the quality of collaboration among group members in a laboratory experiment. Their results suggest that groups supported by a chatbot can benefit from better task outcomes and a higher quality of collaboration (Winkler et al., 2019).

Table 1. Quality indicators for educational chatbots (PIRU)

Category	Quality attributes	(1)	(2)	(3)	(4)	(5)	(6)
Pedagogy[(1)]	- Precise learning contents that are specific to the subject.	X	X	X	X		
	- Alignment with the learning goals.	X	X	X	X		X
	- Active learning that promotes reflection and metacognition.	X		X			X
	- Personalize learning and feedbacks.	X	X	X	X		
	- Progress management.	X	X	X			
Interactivity	- Appropriate chatbot-interactivity to support the pedagogic concept.		X				
	- Ability to respond to specific questions and maintain themed discussions.		X			X	
	- Context relevant chatbot characteristics (e.g., degree of chatbot proactivity).		X		X	X	X
	- Appropriate elements of interaction (e.g., greetings, humor, empathy).	X	X		X	X	X
Reliablitity[(1)]	- No deception.				X	X	
	- Easy access to a stable platform.	X	X				X
	- Robustness to unexpected input.					X	
	- Protect and respect privacy.	X				X	X
	- Feedback for continuous improvement.	X		X			X
Usability	- General ease of use.	X				X	
	- Make tasks more fun and interesting.					X	
	- Entertain and/or enable participant to enjoy the interaction.		X			X	

Source: Based on (1) Kumar (2021), (2) Smutny & Schreiberova (2020), (3) Gonda et al. (2019), (4) Garcia Brustenga et al. (2018), (5) Radziwill & Benton (2017), (6) Kerly et al. (2007) and own contributions.

For the development of chatbots in education, implementation frameworks are needed (Hwang et al., 2020). Building on the work of Kumar (2021) and based on further design aspects mentioned in the literature by Smutny and Schreiberova (2020), Gonda et al. (2019), Garcia Brustenga et al. (2018), Radziwill and Benton (2017), and Kerly et al. (2007), an overarching *quality indicator framework* was combined (see Table 1).

The framework consists of the categories *Pedagogy*, *Interactivity*, *Reliability*, and *Usability* (PIRU). Smutny and Schreibernova (2020) point out that pedagogical aspects should play a key role in the design of educational chatbots. The PIRU framework therefore puts *Pedagogy* at the heart in order to design pedagogical concepts that are meaningful and aligned with a certain chatbot interaction. The categories of *Interactivity*, *Reliability* and *Usability* can therefore be regarded as side constraints that need to be fulfilled for a well-functioning product but should always be attuned to the relevant underlying concept of *Pedagogy* (in our example collaborative learning). The *quality indicators* derived in Table 1 may serve as a guideline for the development of similar chatbot applications.

3. RESEARCH DESIGN AND METHOD

As a methodological foundation for the design and development of our chatbot, we rely on the educational design research (EDR) approach by McKenney and Reeves (2018). Similar to design-based research, the goal of EDR is to create and test specific models of learning within a real-world context (McKenney & Reeves, 2021, p. 84). EDR (as well as design research in education generally) aims to develop, test and implement innovative practices to reimagine and improve socially constructed forms of teaching and learning (Kelly et al., 2008, p. 3). By finding innovative practical solutions for unsolved problems, novel conditions for learning can be evaluated (Euler, 2014, p. 17).

The EDR approach by McKenney and Reeves (2018) is structured into the three core processes of 1) *analysis and exploration*, 2) *design and construction*, and 3) *evaluation and reflection*. The EDR approach can be used in multiple iterations to refine educational applications. The aim of EDR is to provide both maturing interventions and improved theoretical understanding (McKenney & Reeves, 2018, p. 86). In this way, EDR contributes to theory (as a building block and guideline for designing future interventions) and to practice (by addressing the problem at hand).

In section 2, the research topic was already explored and the main findings summarized (*phase 1: analysis and exploration*). Section 4 provides now more detailed insights into the chatbot development (*phase 2: design and construction*). Section 5 and 6 further elaborate on the results of the evaluation carried out (*phase 3: evaluation and reflection*).

4. ARTIFACT DESCRIPTION: CHATBOT FOR GROUP ASSIGNMENTS

The learning environment and the chatbot Tubo were developed by an interdisciplinary team of educational researchers and computer scientists, while all programming was done in-house. Feedback from the wider developer network was repeatedly sought to refine the tool. The insights from this initial development phase are described in detail in the following article (see Burkhard et al., 2021) and have been incorporated into the current version of our prototype.

Figure 1 gives a conceptual overview of the web-based team learning environment. Each student uses their own laptop or computer for this purpose. The learning application is launched through the web browser. During the registration process, students are divided into groups of three. Besides the students, the chatbot Tubo is also a member of each group. Each group thus has its own virtual tutor who can communicate with the students through a joint group chat.

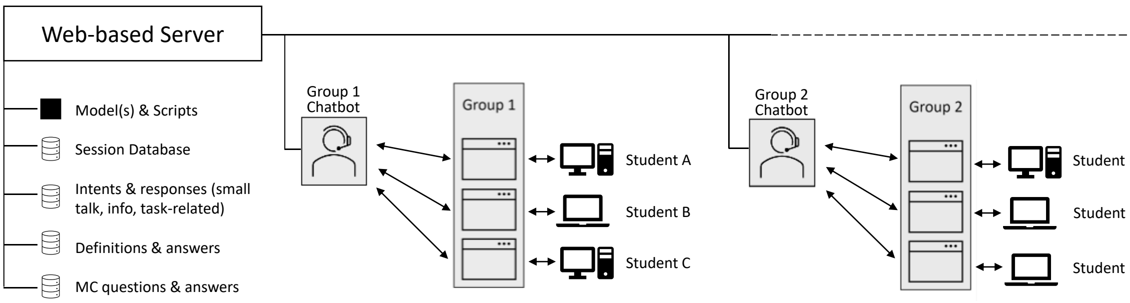

Figure 1. Web-based team learning environment. Source: own representation

Within a group, all students see the same learning environment, which is constantly synchronized between all students in the group. For example, if student B in group 1 writes a message in the group chat, this message is visible to student A, student C and the chatbot Tubo from group 1 (see Figure 1). This also applies to the assignments, which allows students to work together on answering open tasks or multiple-choice tests. For the web-based application to run, several models and scripts have to be executed in the background. In addition to multiple databases that determine the behavior of the chatbot Tubo, a session database is needed in order to manage the current state (including chatbot dialogue, submitted assignments, analytics, user interface) individually and independently for each group.

Figure 2 gives an overview of the user interface for each student. On the right-hand side the students can communicate with each other and with the chatbot Tubo in a group chat. Through the group chat, Tubo introduces the students to the learning environment and provides them with their learning tasks. On the left side (see Figure 2) students see the current group task they have to solve. In the spirit of a storytelling approach, the students help Silvia, the CEO of a pharmaceutical company, and answer her open questions. The students have to work together and type their answers into the text editor (on the bottom left side in Figure 2). If the students have questions, they can contact Tubo directly and the chatbot will try to help them. Among other things, Tubo can provide students with relevant definitions (e.g., in this use-case “what is a monopoly?”) or refer them to further text sources they may need to answer the questions.

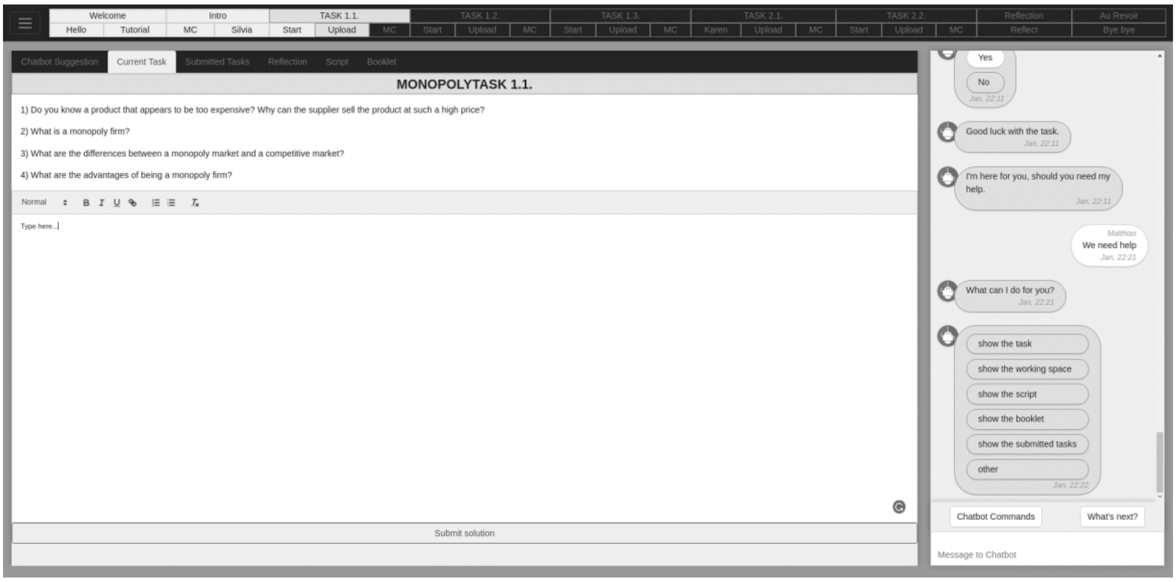

Figure 2. User interface of the designed chatbot webtool application. Source: own representation

While working on the tasks, students have the possibility to further personalize their learning environment (top bar in Figure 2). For this purpose, several tabs can be opened to allow students to do their own research, such as reading the script or reviewing their submitted assignments. The students' individual layout settings remain until a new task needs to be worked on and attention needs to be drawn to new content.

After students have answered and submitted an open-ended work assignment, Tubo asks students multiple-choice questions on key concepts to monitor learning success. To stimulate student reflection on group work, Tubo collects data on students' group behavior. Students can see how they perceived the difficulty of each task and how long it took them to complete each subtask. The data collected also reveals which student made contributions to the team output how often and in what way (e.g., regarding open-ended and multiple-choice questions).

Table 2 summarizes the task areas that the chatbot Tubo performs within the learning environment. To get a more detailed impression of Tubo and its learning environment, the following YouTube video (https://youtu.be/3yLjZYLc2fM) can be watched.

Table 2. Task areas of the educational chatbot Tubo within the learning environment

Task areas	Objectives	Usage
Onboarding-Bot	To greet and warm up the students. To explain to the students the learning environment.	Only at the beginning.
Assignment-Bot	To provide students with learning instructions and clarify the assignments.	For all open tasks.
Quiz-Bot	To ask multiple-choice questions about key concepts to monitor learning success. To provide feedback to the MC-questions.	For all MC-questions.
Orchestration-Bot	To adjust the user interface and display different contents depending on the current task.	If Tubo is asked for help / new task begins.
Progress-Bot	To rate self-perceived progress in students' assignments, identify teamwork issues.	Before/after each task.
FAQ-Bot	To answer frequently asked questions, to provide students with definitions and directing them to relevant textbook sections.	Always.
Smalltalk-Bot	To make jokes and engage in Smalltalk with students; to motivate students for the next tasks.	Always.

Source. Own representation based on Kumar (2021).

5. FORMATIVE EVALUATION OF THE FIRST DESIGN CYCLE

5.1 Classroom Experiment

To iteratively improve the chatbot Tubo, the tool was tested in a first design experiment with a school class of a technical vocational school (in Zurich, Switzerland). The eleven students (9 male, 2 female) were in their third year of training as "event managers" and worked alongside their school education for television-, opera- or similar event-companies. The average age of the students was 18.7 years (minimum 17, maximum 21 years). The chatbot tool Tubo was tested with the school class live on site during four 45-minute lessons (due to covid-19 measures one student had to participate in the lessons online). The lessons were delivered by the students' class teacher A and a second teacher B who had been involved in the development of the chatbot tool and therefore already knew the learning setting well. In addition, a member of our research team was present to document the progress of the lessons.

At the beginning of the first lesson, the students formed four groups (usually three people per group). Each student had their own laptop to work on. After a short thematic introduction by teacher B, each student was asked to access the chatbot tool through the web browser and log into their group. This worked well for most students and the chatbot welcomed the group members. Figure 3 illustrates how the students worked during the lessons with Tubo in small groups.

Figure 3. Classroom experiment: Student groups working collaboratively with the chatbot Tubo

In its role as a virtual tutor, the chatbot Tubo explained the learning process to the students and gave further tips on how the students could best interact with the learning environment. For example, if students had questions, they could ask for help by typing “Tubo” into the group chat. In addition, a “What's next”-button was available whenever students wanted to know how to continue with their learning process (e.g., after having small talk with Tubo).

Over the course of the four lessons, the student groups were guided through the content by the chatbot Tubo and solved all five assignments. Each assignment contained both story elements and open-ended questions that the students could answer directly in the learning environment's text editor. At the end of each assignment a multiple-choice quiz had to be taken which monitored learning success. The multiple-choice quiz had to be answered fully correctly by the students in order to move on to the next assignment. If the answers were wrong, the students received further tips from Tubo. Often the groups had to do a multiple choice quiz several times because they did not get all the answers right at the first time. Some students found this tedious and verbally expressed it. The two teachers observed the behavior in the individual groups and motivated the students if necessary. They also provided answers to content-related questions that Tubo was not able to answer sufficiently from the students' point of view.

5.2 Student Feedback

After the four lessons, teacher B initiated a discussion about the lesson. Besides technical feedback on the learning environment (e.g., login problems, lags), the students expressed positive and negative aspects they had noticed while working with Tubo.

As positive aspects, several students mentioned that they liked the overall concept. One student said that chatting and working with Tubo had something “playful” and “adventurous” and that it was a welcome change compared to the regular school day. Another student mentioned that he could imagine working with Tubo on his own and in home schooling.

The students had different opinions about the way they were guided from one task to the next (*scaffolding elements*). Several students found the guidance and instruction provided by chatbot Tubo useful because they knew what was required of them. However, three students also criticized this relatively unvarying approach, as in their opinion, it became somewhat repetitive and boring over time. In the view of these students, it encouraged over time “to click through without thinking”. These findings are line with the results of Rienties et al. (2012) who conclude that it is difficult to find the right balance of scaffolding to simultaneously support autonomous and more control-oriented learners.

The students’ views also differed regarding the benefits of the group setting. Some students found it useful to work in groups with the tool as they could help each other and discuss the content in depth. Other students, however, found it difficult to divide the subtasks of the assignments within the group in a meaningful way. From the point of view of these students, this led them to take turns answering the learning content instead of working together.

6. DISCUSSION

After the design experiment, a discussion took place with the two teachers involved and the research team. As a guideline for the discussion, the quality indicators for educational chatbots PIRU (*Pedagogy*, *Interactivity*, *Reliability*, *Usability*) were used (see Table 1).

Regarding *Pedagogy*, the two teachers said, that group learning with a chatbot seemed to be an interesting approach that could also have potential in hybrid class situations. Although one student was only able to participate in class online due to covid-19 restrictions, this person could still be integrated into the lessons. However, from the teachers' point of view, working with chatbot Tubo required a relatively high level of commitment on the part of the students. If the students took turns answering the questions or simply tried to “click” through the tool as quickly as possible, the learning effect of the group assignments was partly lost. To have a certain control as a teacher and to motivate the students if necessary, the use in this on-site setting was considered reasonable.

According to teacher B, it would theoretically be possible for each student to work with Tubo alone and not in a group setting. However, this would mean that in-depth discussions about the questions and contents would be missing. Discourse within the group about "right" and "wrong" answers can promote a common understanding and create an atmosphere in which students can learn from each other (see e.g., Dillenbourg & Schneider, 1995). Teacher A agrees with this but thinks that the distribution of team roles within a group should be further specified and improved. Since all team members received the same work assignments, it was difficult for some students to divide the work within the group in a meaningful way. Different roles and subtasks for individual group members could be added in a next iterative design cycle (e.g., through the "jigsaw" method that provides each group member only with partial data to require more collaboration (see e.g., Dillenbourg & Schneider, 1995, p. 135)). In this way chatbot *interactivity* could be better aligned with the *pedagogical* concept of collaborative learning.

Regarding *reliability* and *usability*, the language understanding of the chatbot as well as the user interface could be further optimized in a next design cycle. In order to increase the chatbot *reliability* and better manage students' expectations, it might be helpful to display a larger number of question-and-answer options in the form of buttons. This approach could actively direct the students' attention to content on which Tubo can provide competent information.

In line with the results of Rienties et al. (2012), the design experiment showed that scaffolding elements can be viewed ambivalently. On the one hand, the structure, and the way chatbot Tubo guided them through the learning environment was helpful for more control-oriented students because they always knew what to do. On the other hand, this relatively rigid grid was also perceived as a restriction by some students who preferred autonomy. These students desired more individual freedom in their actions and found it boring to answer the tasks after a while.

7. CONCLUSION & OUTLOOK

This paper investigated how chatbot-mediated learning can be designed to foster team-based learning of middle school students. Based on the *quality indicators* for chatbot development identified in the literature, an extended framework (see Table 1) was developed that served as a guideline for the implementation of our chatbot application. The chatbot Tubo was tested with a middle school class as part of a formative evaluation. The received feedback suggests that the approach of team-based learning with chatbots has a lot of potential from the students' and teachers' point of view. However, the role distribution of the individual group members may have to be further specified to address the different needs of autonomous as well as more control-oriented students (compare to Rienties et al., 2012).

This study is subject to several limitations. First, the tool was only tested with a relatively small school class of eleven students. A larger sample size would be desirable to validate the external validity of the results. Second, the application was only tested over a relatively short period of four lessons. It could be that a *novelty effect* occurred (see e.g., Jeno et al., 2019), which may have initially increased student motivation but decreases over time once the students get used to chatbot Tubo. From this point of view, a more long-term use in the classroom seems to be meaningful. Third, no control group design was applied, and it is therefore difficult to draw a definitive conclusion on the achieved learning outcomes. However, in line with the applied educational design science approach and the aim of obtaining formative feedback as part of a first design cycle, this limitation was accepted on purpose.

From a theoretical point of view, the paper at hand can serve as a starting point for further research and discussion, as it explores the relatively new area of *chatbot-mediated learning in team-based scenarios*. From a practical standpoint, the paper contributes to a better understanding of how chatbot-mediated learning scenarios can be designed and implemented in the actual classroom. The derived framework in section 2 may serve as a guideline for other researchers who want to implement similar projects and uncover further potential of the technology in more detail.

Further research is needed to validate our designed learning scenario in terms of learning outcome. In a next design cycle, this could be achieved with a larger sample and a more quantitative approach. To further develop the chatbot application, learning analytics elements (see e.g., Daud et al., 2017, Mandinach & Schildkamp, 2021) could be implemented to provide the teacher with real-time data of what is happening in each group in order to support teachers' data-based decision making. This would enable the teacher to identify problems within groups more quickly and respond to them in a more targeted way.

REFERENCES

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications, 2*. https://doi.org/10.1016/j.mlwa.2020.100006

Briggs, R. O., Kolfschoten, G. L., de Vreede, G. J., Lukosch, S., & Albrecht, C. C. (2013). Facilitator-in-a-box: process support applications to help practitioners realize the potential of collaboration technology. *Journal of Management Information Systems, 29*(4), 159-194. https://doi.org/10.2753/MIS0742-1222290406

Burkhard, M., Seufert, S., Cetto, M., & Handschuh, S. (2021). The Textbook Learns to Talk: How to Design Chatbot-Mediated Learning to Foster Collaborative High-Order Learning?. In *Innovate Learning Summit*, 12-21. Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/p/220264/

Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., & Alowibdi, J. S. (2017). Predicting student performance using advanced learning analytics. In *Proceedings of the 26th international conference on World Wide Web companion*, 415-421. https://doi.org/10.1145/3041021.3054164

David, B., Chalon, R., Zhang, B., & Yin, C. (2019). Design of a collaborative learning environment integrating emotions and virtual assistants (chatbots). In *2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD),* 51-56. IEEE. https://doi.org/10.1109/CSCWD.2019.8791893

Dillenbourg, P., & Schneider, D. (1995). Mediating the mechanisms which make collaborative learning sometimes effective. *International Journal of Educational Telecommunications, 1*(2), 131-146. https://www.learntechlib.org/primary/p/15155/

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior, 52*, 338-348. https://doi.org/10.1016/j.chb.2015.05.041

Euler, D. (2014). Design-Research – a paradigm under development. *Design-Based Research*, 15-44. Stuttgart: Franz Steiner Verlag

Garcia Brustenga, G., Fuertes-Alpiste, M., & Molas-Castells, N. (2018). *Briefing paper: Chatbots in education*. eLearn Center, Universitat Oberta de Catalunya. https://doi.org/10.7238/elc.chatbots.2018

Gonda, D. E., Luo, J., Wong, Y. L., & Lei, C. U. (2019). Evaluation of developing educational chatbots based on the seven principles for good teaching. *Proceedings of the 2018 IEEE international conference on teaching, assessment, and learning for engineering, TALE 2018, Australia*, 446–453. IEEE. https://doi.org/10.1109/TALE.2018.8615175

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Taylor & Francis.

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Jeno, L. M., Vandvik, V., Eliassen, S., & Grytnes, J. A. (2019). Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. *Computers & Education*, *128*, 398-413. https://doi.org/10.1016/j.compedu.2018.10.008

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of Computer-Supported Collaborative Learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational research review, 28*, 100284. https://doi.org/10.1016/j.edurev.2019.100284

Kelly, A. E., Lesh, R. A., & Baek, J. Y. (2008). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. Routledge. https://www.routledgehandbooks.com/doi/10.4324/9781315759593

Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: *Towards natural language negotiation of open learner models*. Knowledge-Based Systems, 20(2), 177–185. https://doi.org/10.1016/j.knosys.2006.11.014

Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. *International Journal of Educational Technology in Higher Education, 18*(1), 1-28. https://doi.org/10.1186/s41239-021-00302-w

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia-social and behavioral sciences, 31*, 486-490. https://doi.org/10.1016/j.sbspro.2011.12.091

Ludvigsen, S., & Mørch, A. (2010). Computer-supported collaborative learning: Basic concepts, multiple perspectives, and emerging trends. *The international encyclopedia of education, 5*, 290-296. https://doi.org/10.1016/B978-0-08-044894-7.00493-0

Mandinach, E. B., & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation, 69*, 100842. https://doi.org/10.1016/j.stueduc.2020.100842

McKenney, S. & Reeves, T. C. (2018). *Conducting educational design research*. (2nd Ed.) Routledge: Oxon.

McKenney, S., & Reeves, T. C. (2021). Educational design research: portraying, conducting, and enhancing productive scholarship. *Medical Education, 55*(1), 82-92. https://doi.org/10.1111/medu.14280

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior, 52*, 573-588. https://doi.org/10.1016/j.chb.2015.01.050

Nussbaum, M., Alvarez, C., McFarlane, A., Gomez, F., Claro, S., & Radovic, D. (2009). Technology as small group face-to-face Collaborative Scaffolding. *Computers & Education, 52*(1), 147-153. https://doi.org/10.1016/j.compedu.2008.07.005

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence, 2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint*. https://arxiv.org/abs/1704.04579v1

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 102630. https://doi.org/10.1016/j.ijhcs.2021.102630

Rienties, B., Giesbers, B., Tempelaar, D., Lygo-Baker, S., Segers, M., & Gijselaers, W. (2012). The role of scaffolding and motivation in CSCL. *Computers & Education, 59*(3), 893-906. https://doi.org/10.1016/j.compedu.2012.04.010

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers and Education, 151*, 103862. https://doi.org/10.1016/j.compedu.2020.103862

Splichal, J. M., Oshima, J., & Oshima, R. (2018). Regulation of collaboration in project-based learning mediated by CSCL scripting reflection. *Computers & Education, 125*, 132-145. https://doi.org/10.1016/j.compedu.2018.06.003

Winkler, R. & Söllner, M. (2018): Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. In: *Academy of Management Annual Meeting (AOM)*. Chicago, USA. https://www.alexandria.unisg.ch/publications/254848

Winkler, R., Neuweiler, M. L., Bittner, E. A. C., & Söllner, M. (2019). Hey Alexa, Please Help Us Solve This Problem! How Interactions with Smart Personal Assistants Improve Group Performance. In *International Conference on Information Systems (ICIS)*. Munich. https://aisel.aisnet.org/icis2019/general_topics/general_topics/17/

Winkler, R., Söllner, M., & Leimeister, J. M. (2021). Enhancing Problem-Solving Skills with Smart Personal Assistant Technology. *Computers & Education, 165*. 104148. https://doi.org/10.1016/j.compedu.2021.104148

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. J*ournal of Child Psychology and Psychiatry, and Allied Disciplines, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x.