Pre-print of full paper presented at CONVERSATIONS 2019 - an international workshop on chatbot research, November19-20, Amsterdam, the Netherlands. The final authenticated version of the paper is available online at https://doi.org/10.1007/978-3-030-39540-7_18

Small Talk Conversations and the Long-Term Use of Chatbots in Educational Settings – Experiences from a Field Study

Sebastian Hobert[1] and Florian Berens[1]

[1] Universität of Goettingen, Goettingen, Germany
{shobert,florian.berens}@uni-goettingen.de

Abstract. In this paper, we analyze the use of small talk conversations based on a dialogue analysis of a long-term field study in which university students regularly interacted with a chatbot during a 3-month period of time in an educational setting. In particular, we analyze (1) how often the students engage with small talk topics during the field study, and (2) whether a larger amount of small talk conversations correlates with the students' engagement in learning activities within our chatbot-based learning system, i.e., if engaging in small talk conversations correlates to a more intensive use of the chatbot during our field test. Our results suggest that small talk conversations might play an important role in the design of our chatbot as students who chat about small talk topics also frequently chat about learning-related topics. Nevertheless, the overall impact of small talk capabilities of chatbots should not be overestimated.

Keywords: Chatbot, Small Talk Conversation, Conversational Agent, Pedagogical Conversational Agent.

1 Introduction

The ability to communicate using natural language in a human-like way seems to be an important design feature of chatbots and virtual assistants, as it can be seen in many examples from practice. A typical capability of such state of the art conversational agents is to enable the chatbots to engage in small talk conversations [1]. For example, chatbots are often able to talk about informal topics that are usually not important for a computer program. For instance, such informal conversations cover greeting users, talking about the chatbot's well-being or telling jokes. In this paper, we define informal conversations that are not important for the chatbot's overall purpose as small talk messages. Even though engaging in small talk does not seem to provide any direct beneficial value for most conversational agents, it may be argued to hold indirect benefits [2]. Small talk is a standard part of communication among humans and makes the conversation flow. Thus, it may be argued that including small talk might be beneficial for chatbots as well. However, there is a lack of knowledge about the specific impact of small talk conversations. To fully understand its effects and to include it in the design of chatbots, there is a need to understand the users' engagement in small talk chats.

In particular, there is a lack of knowledge about chatbots that are used in long-term settings. Currently, many chatbot-based systems are only used to interact with users for a short period of time, e.g., as personal assistants for customer support. In these cases, small talk capabilities are used to demand the users' attention and to start a conversation, e.g., by actively talking to the users (e.g., "Hello, how can I help you?"). In these cases, the interaction with the chatbot often only takes a short period of time. Long-term adoption of chatbots is usually not needed in these cases. However, in other settings, like in education, it seems appropriate to provide students with a chatbot that is not only available for a short period of time but can support them during an entire learning period, e.g., a full lecture term at the university.

To address these long-term settings, this study provides insights into such a field test in which a chatbot interacted with users for several months. Mainly, our study focuses on an educational setting in which we introduced a chatbot-based learning system. To get detailed insights into the long-term impact of including small talk capabilities in the design of our chatbot, we address the following research questions in the remainder of this paper:

RQ1: How often do students engage in small talk conversations during long term use of a chatbot in an educational setting?

RQ2: How does the usage frequency of the chatbot's small talk capabilities correlate with the students' engagement in learning activities?

To answer these research questions, the remainder of this paper is structured as follows: First, we outline related research on chatbots and its' capabilities to act human-like by engaging in small talk conversations in the next section. Subsequently, we describe our research design in Section 3. In Section 4, we present the results of our analysis and discuss our findings in Section 5. Finally, we summarize the results in the conclusion section.

2 Related Research

Chatbots can be defined as information systems with a natural language-based user interface. Due to the interaction with the users in a "conversational-style" [3] using natural language, the interaction of chatbots with users is similar to communication between humans. From a technical perspective, chatbots are usually designed to interact autonomously by relying on methods known from machine learning, artificial intelligence, and natural language processing.

By using chatbots in educational settings, students should be supported during learning processes. According to prior research, it is to be expected that chatbots could provide "significant positive impact on learning success and student satisfaction" [4]. For instance, [5] developed a chatbot that is able to reply to posts of students in a forum of a computer science course. In another project, the conversational agent *MentorChat*

was introduced, which should support students in collaborative learning tasks [6]. Further exemplary use cases are described in recent literature reviews in more detail, see, e.g., [4] and [7].

A common pattern known from various chatbots or other virtual assistants is their capabilities to engage in small talk conversations [1]. For instance, many chatbots are able to respond to messages like "Tell me a joke" or "How are you?". By implementing such capabilities, which are not directly beneficial for the intended purpose of a chatbot, developers want to enable the chatbot to act human-like [8] in a socially-accepted way. In doing so, the adoption of users to interact with a chatbot should be fostered, and in some cases, the chatbot should even hide that it is not a human being. However, most available chatbots used in practice are not designed for educational practices, and the users' interaction is different compared to the use in university courses. Whereas many chatbots are designed for corporate purposes (e.g., customer support [9]) and are only used for a short period of time, chatbots that are introduced in university courses can support the students during the whole lecture period like in [5]. Consequently, the adoption of the students is more important as the usage time is longer. However, the implications of small talk capabilities of chatbots for the long-term adoption of users in educational settings have to the best of our knowledge not yet sufficiently be researched. Thus, we will focus on analyzing the small talk usage of students in a long-term field study in this paper.

3 Research Design

In the following, we outline the research design that we applied to answer our small talk-related research questions. We conducted our study as part of a design-oriented research project based on the Design Science Research Approach [10–12]. As the main result of this design-oriented research project, we conceptualized and implemented a chatbot-based learning system, which is the basis for this investigation on the impact of small talk capabilities (see next section). The chatbot-based learning system is designed as a progressive web-application and provides a user interface that is similar to common instant messenger apps for smartphones. The chatbot-based learning system provides students the possibility to ask open-ended questions concerning the content of a university course. The chatbot is able to answer those questions immediately based on a database consisting of more than 450 learning objects (i.e., definitions of basic terms). Additionally, the chatbot-based learning system provides students access to formative exercise.

Using this implementation of our chatbot-based learning system (see a detailed description in the following section), we conducted a field study starting from April 2019. During this field study, approx. 700 students of an introductory lecture on statistics for social sciences got access to the system and had the possibility to interact with it until the end of the lecture period. To answer our research questions, we analyzed the pseudonymized discourses of the chatbot with the students. To this aim, we used the textual messages from the students and the chatbot as a basis and enhanced it using available metadata (e.g., timestamp of the conversation). To analyze the metadata, we tagged the

text messages with intents computed by our chatbot using the natural language processing library NLP.js [13]. Our trained natural language processing model which is based on approx. 2800 exemplary messages is not accurate in every case. Thus, we manually reviewed those messages that were classified with low accuracy by our chatbot. In doing so, we retrained and improved the intent recognition of future messages. During the manual reviewing process, we adjusted the recognized intents of approx. 1.8 % of all messages sent by the users. Finally, we excluded all messages that were not classified as small talk and resulted in our case base of small talk interactions.

Fig. 1 summarizes the steps we conducted to process the data.

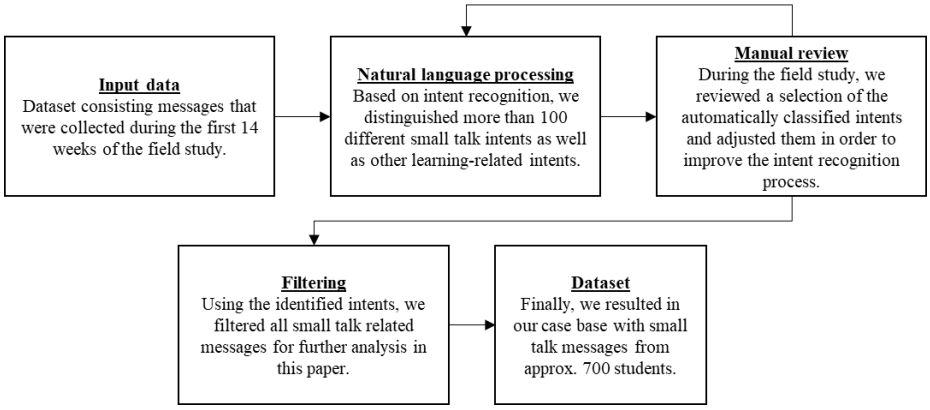

Fig. 1. Research steps for the processing of the data to derive our final dataset

Using the final dataset, we analyzed the small talk interactions of the learners descriptively and on a timely basis to answer research question 1. To respond to research question 2, we tested whether a student's small talk interactions correlates with his/her overall engagement in the learning app during the three-month field test statistically using Spearman's rank correlation coefficient.

4 Chatbot Overview

To conduct the long-term field study, we used our chatbot-based learning system that we develop in early 2019. The system is developed as a messenger-like chatbot system and provides students participating in an introductory lecture on statistics education a natural language-based user interface.

To ensure that every student is able to use the system independently of a specific device or operating system, we implemented it as a progressive web-app. From a technical perspective, we used HTML5, CSS, and JavaScript for the implementation of the students' frontend. The natural language understanding of the students' written input is done using a node.js backend component. Additionally, the lecturer as well as several student assistants had access to an administration control panel, which could be used

for reviewing questioning and answering dialogues (i.e., we conducted quality assurance tasks to improve the intent recognition) that were marked by the natural language processing component with a high probability of errors (i.e., the chatbot was unsure whether the students' messages were understood correctly).

Fig. 2 provides an overview of the technical architecture of the chatbot-based learning system.

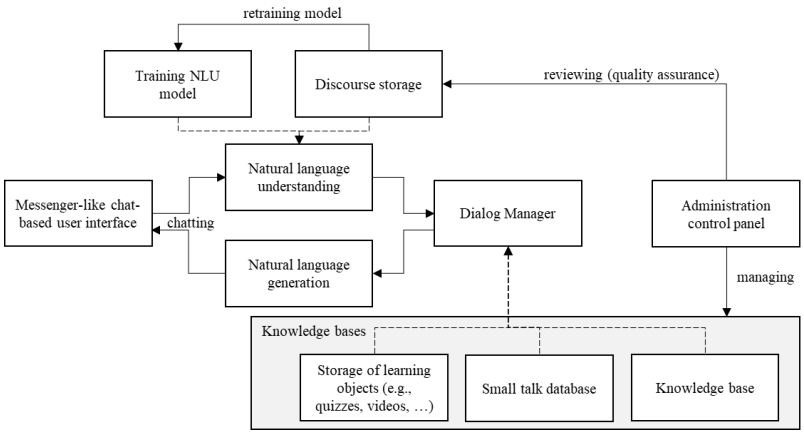

Fig. 2. Simplified overview of the technical architecture

From a students' perspective, our chatbot-based learning system can be seen as the only online resource required in addition to the face-to-face lectures and tutorial sessions. The system provides all relevant additional online materials (like supplementary formative quizzes, video recordings of the lectures, slides of the lectures, etc.) to the students. Besides these learning-related functionalities (see Fig. 3), the chatbot-based learning system is, for instance, capable of answering questions concerning organizational issues and conducting small talk interactions with students.

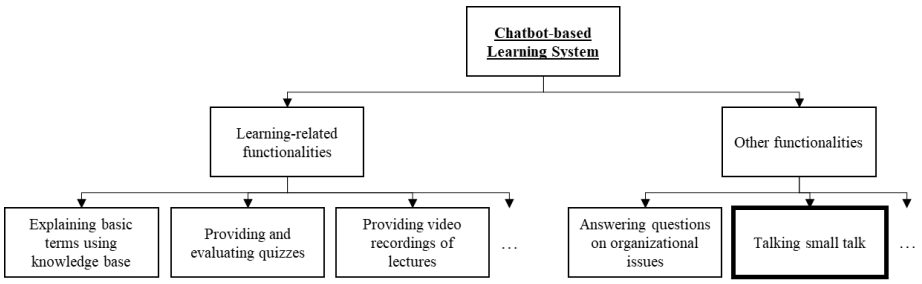

Fig. 3. Overview over core functionalities of the chatbot-based learning system

All learning-related functionalities, as well as the other functionalities, are integrated into the messenger-like chat-based interface. An overview of exemplary messages asked by learners is displayed in Table 1.

Table 1. Exemplary messages sent by learners

Exemplary message	Description
"Show me the video recordings."	Requesting video recordings
"Show me the lecture slides."	Requesting pdf files of lecture slides
"What is the meaning of the standard deviation?"	Asking for a definition of a term mentioned in the lecture
"What is the date of the exam?"	Asking for organizational aspects
"Hi, how are you?"	Talking small talk (i.e., messages that are not related to the learning process)
"What is your name?"	

In the remainder of this paper, we will mainly focus on the functionalities related to the small talk capabilities of the chatbot-based learning system and its influence on the students' usage of the learning-related functionalities.

5 Results

In the following Subsection 5.1, we respond to the first research question by analyzing the number of small talk messages written by users. To this aim, we conducted an analysis based on the recognized intents by the natural language understanding component of our chatbot-based learning system. Subsequently, we focus on correlations of the students' engagement in small talk conversations with learning activities in response to the second research question in Subsection 5.2.

5.1 Engagement in small talk conversations

The first time a student opens the chatbot-based learning system, the system posts a friendly welcome message. This welcome message can be seen on the one hand as an introductory example of how the overall system works. On the other hand, it is used to foster the student to participate in a conversation with the chatbot. Afterward, the student can proceed with the conversation by replying with small talk messages or focusing on learning-related aspects by asking further questions. As the chatbot-based learning system is designed as an open-topic natural language processing system, the user is not restricted on specific topics. In the following, we will only focus on small talk-related messages written by the students. To get an impression about the students' interaction with the chatbot-based learning system when engaging in small talk conversations, Fig. 4 outlines an exemplary dialog excerpt in which a student first chatted about small talk topics with the chatbot before switching over to learning-related contents (i.e., quizzes). The overall dialog continues with further learning-related contents. Thus, in this case, the small talk conversation might be useful to motivate the student to learn.

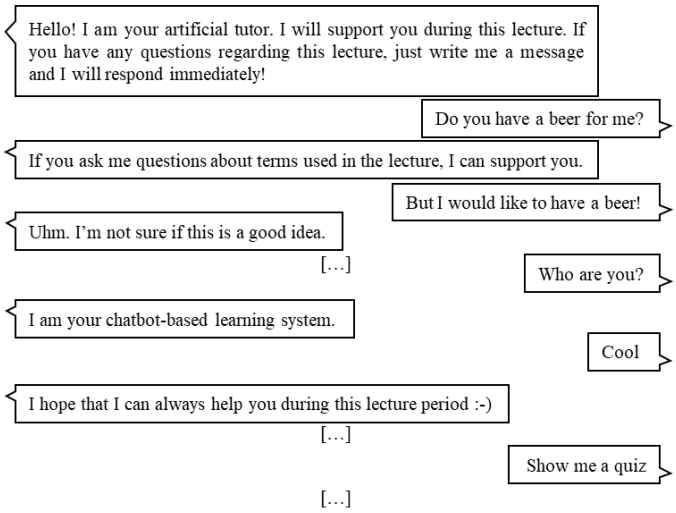

Fig. 4. Exemplary dialog excerpt of one student chatting with the chatbot

As shown in Fig. 5, the amount of small talk messages posted by the students reaches its all-time high at the start of the lecture period[1]. In this first week of the field study, approx. 30 % of all messages sent to the chatbot were labeled as small talk (i.e., approx. 1100 small talk messages). In the subsequent week, still approx. 900 messages are related to small talk. However, the amount decreases to approx. 10 % in this second week, as the number of learning-related messages increased rapidly. After these first two weeks in which the students engaged a lot in small talk conversations, the number of small talk messages decreased substantially and never reached more than approx. 450 messages.

The analysis of the time history of the small talk use thus indicates that small talk conversations are particularly relevant in the beginning when students are interested in discovering the chatbot-based learning system. Independently of the overall system use, the number of small talk messages never reached a high level.

[1] As the lecture started and ended not on the first day of the week, the weeks 1 and 14 encompass only four to six days and are thus shorter than the remaining weeks. This can explain the lower number of written messages by the students in these two weeks.

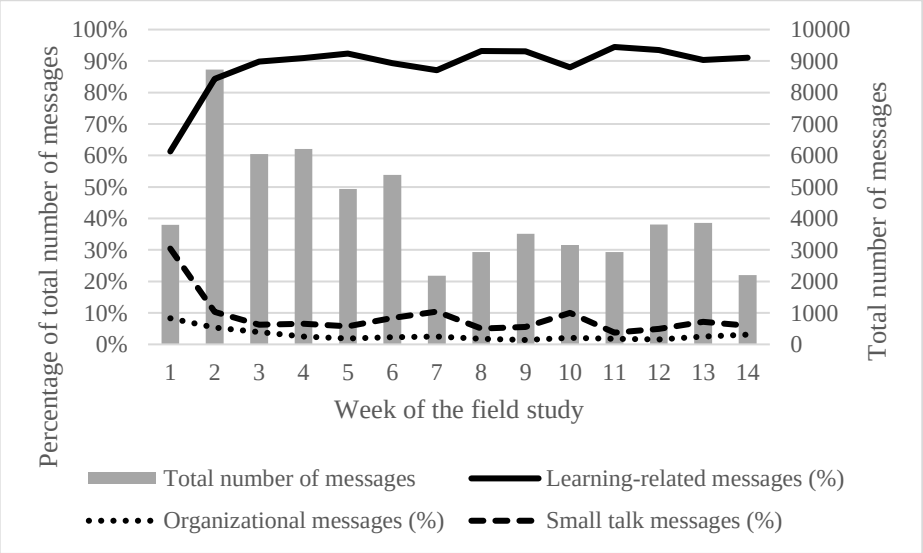

Fig. 5. Overview of the time history of the chat usage

The decreasing interest of the students to engage in small talk conversations cannot only be seen in the total amount of small talk messages but also in the distribution of the written small talk questions (see Fig. 6). Whereas in the first weeks of the field study, almost all active small talk intents that can be recognized by the chatbot were triggered (more than 100 different intents). In the following weeks, the number of triggered small talk intents decreased rapidly.

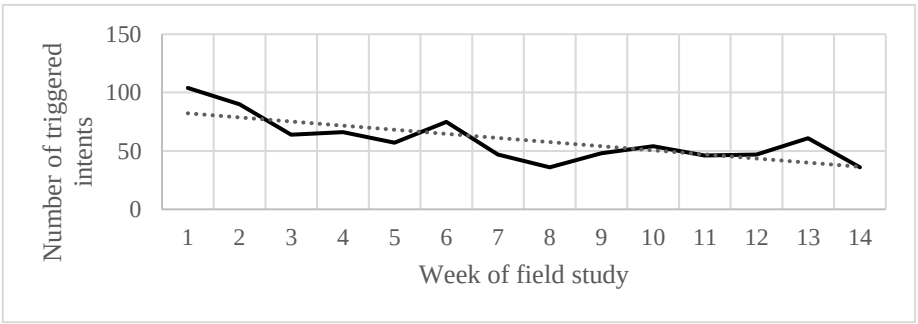

Fig. 6. Overview of the time history of the number of triggered intents

The decreasing number of triggered small talk intents might be explained similarly to the decreasing total number of small talk engagement as seen above. Talking with the chatbot about small talk topics might become less interesting while actually using the system for learning purposes might become more important as the final exam is approaching.

Interesting regarding the small talk usage is that there is a large number of students who are not engaging in small talk conversations at all. Approx. 50 % of all users only sent 0 to 5 messages that were recognized by the chatbot as small talk content. In contrast to that, the remaining approx. 50 % wrote at least 6 or more small talk-related messages. In some cases, even much more than 40 messages as displayed in Fig. 7.

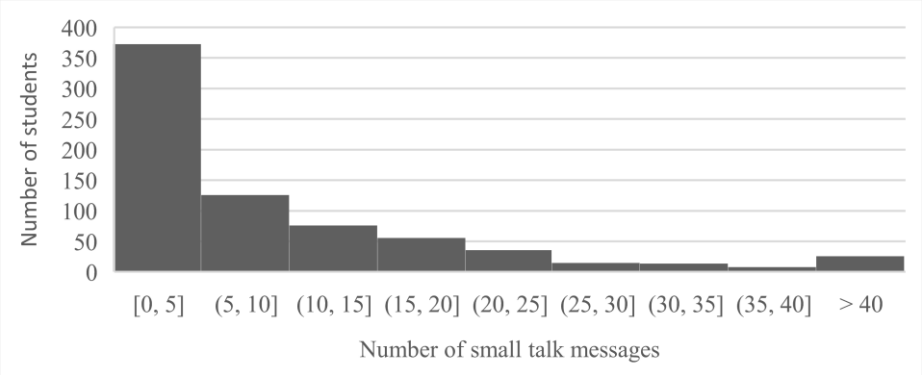

Fig. 7. Grouping of the number of students based on the number of small talk messages

Based on this observation, in the following subsection, we will analyze whether the engagement in small talk conversations correlates with the engagement in learning activities.

5.2 Correlation of small talk engagement and learning activities

As indicated by the high usage of the small talk capabilities of our chat-based learning system with more than 2000 small talk messages sent by the students in the first two weeks of the field study (see Fig. 5), it should be analyzed whether these small talk capabilities are beneficial for the long-term adoption. If students only chat with the chatbot about small talk topics but not about learning-related topics, the overall aim of the system to support the students wouldn't be reached.

First, we analyzed whether there is a correlation between the number of small talk messages written by the students in the first two weeks of the field study and their total number of chat-based learning interactions (e.g., solving a quiz, asking domain-specific questions about lecture content, download learning material) during the field study (see Fig. 8 for a visualization of the dataset). To this aim, we calculated the Spearman's rank correlation coefficient. Based on this, the statement that the number of small talk messages written in the first two weeks of the field study correlates with the total number of learning interactions in the whole period of 14 weeks can be supported ($r_s = 0.348$, $p = 0.000$, $n = 700$). We can confirm a moderate correlation between both aspects.

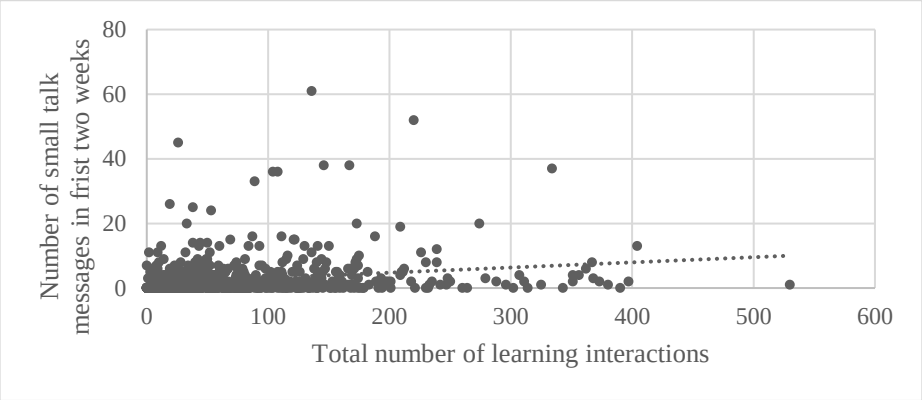

Fig. 8. Comparison of small talk usage in the first two weeks with the total number of learning interactions during the whole period of 14 weeks

Second, we also analyzed whether this correlation also exists when comparing not only the number of small talk messages at the beginning of the field study (i.e., the adoption phase) but also in the overall period of 14 weeks (see Fig. 9). According to the Spearman's rank correlation coefficient, these two variables correlate even stronger ($r_s = 0.557$, $p = 0.000$; $N = 700$) compared to the small talk usage in the first two weeks. This indicates that students who engage in learning-related activities in our chatbot-based learning system more often also talk about small talk topics in the long run.

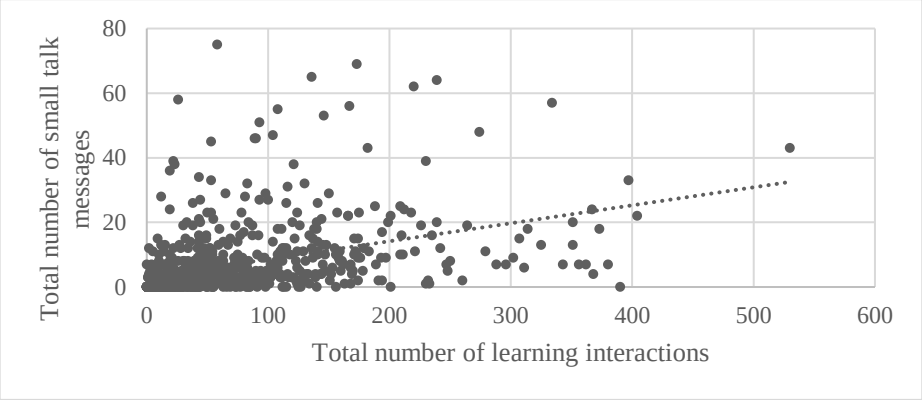

Fig. 9. Comparison of the total number of small talk usage with the total number of learning interactions during the whole period of 14 weeks per student

6 Discussion

In the following, we will discuss the implications and limitations of our study.

6.1 Implications

The results of our analysis of the conversational data of our field study suggest that small talk capabilities might facilitate the adoption of chatbot-based systems in educational settings. In the analysis of the first research question, we showed that many students chatted with the chatbot, particularly at the beginning of the field study. Due to this high amount of more than 2,000 small talk messages sent by students in the first two weeks, we propose that small talk capabilities are a design feature that should be considered. Since the students themselves do not gain any direct advantage or learning increase from small talk conversations, the pure amount of small talk messages they produced nevertheless has to be considered as an indication of the importance students attribute to small talk. This is in line with the results of prior studies and the common practice of chatbots used in other use cases. For instance, [14] showed that building long-term human-computer relationships might be beneficial. We assume from the insights from our study that small talk might be able to foster such a relationship in our field study between the chatbot and the students. On the side of the developers, a non-implementation of small talk would have led to unsatisfying responses (like "Sorry, I couldn't understand you properly."). Thus, the lack of small talk capabilities could possibly have a negative effect on the students' enjoyment and finally on their adoption. To analyze this in more detail, it would be possible to introduce the same chatbot but with disabled small talk capabilities in a similar field setting in the future. Thus, it could be analyzed if a lack of small talk capabilities has a negative impact.

Furthermore, we have evidence based on Spearman's rank correlation coefficient that the amount of small talk messages and the overall amount of learning activities correlates positively within our field study. An interesting aspect of our analysis is that the correlation is even stronger when the number of small talk messages within the considered time span is expended from two weeks to the whole period of 14 weeks of our field study. This might be explained by an increased overall engagement in chatting with the learning system, which also resulted in additional small talk messages that were sent by the actively participating students.

Nevertheless, the effect of implementing small talk capabilities to foster the users' engagement and adoption of a chatbot should be analyzed in further research studies. It should also not be overestimated as the correlation only implies a moderate effect size. However, we assume that not providing small talk capabilities could have a negative impact.

6.2 Limitations

Our analysis is based on a large dataset from our chatbot-based learning system. The dataset contains messages that were manually written by users as well as messages that were automatically created by the chatbots (e.g., answers to questions asked by the users). Due to the large amount of data, classifying each message individually by humans was not possible due to resource constraints. Thus, we used the chatbot's trained intent recognition algorithm for this purpose. Additionally, we tried to manually review as many messages as possible during the field study for quality assurance purposes and to

retrain the algorithm in order to reduce classification errors. Nevertheless, there might still be errors in the classification of the intents.

Our analysis is based on one field study that we conducted in summer 2019 in one university course. We assume that our results are not only valid for this particular lecture, but also for other large-scale lectures on different topics. However, further analyses should be conducted in further fields of study.

Finally, we showed that correlations between the small talk usage and the overall learning engagement in the chatbot-based learning system exist. However, it might be interesting to identify in further studies which additional factors are fostering the students' adoption of chatbot-based learning systems and whether the correlations also imply causalities.

7 Conclusion

In this research study, we analyzed the long-term interaction of approx. 700 students with a chatbot. In particular, we analyzed the students' engagement in small talk conversations that are not directly beneficial for the learning success but are expected to be supportive for the students' adoption of our chatbot-based learning system.

First, we analyzed the overall usage of the chatbot's small talk activities and showed that the amount of small talk messages is particularly high at the beginning of our field study where it reached about 30 % of all messages. Additionally, we showed that the diversity of different small talk topics was higher in the beginning compared to the remaining time period. Second, we significantly showed that there is a positive correlation between the number of small talk messages sent by a student and her/his overall learning activities within our chatbot-based learning system.

To fully understand the importance of small talk interactions in chatbot-based learning systems, further research is necessary. For instance, further in-depth discourse analysis combined with survey results about the perceived adoption of the students might be helpful to analyze the topic in more detail. Additionally, introducing the same chatbot in a similar setting with disabled small talk capabilities would also be interesting to analyze if this results in less engagement. Nevertheless, our findings already show that students actively engage in small talk activities. Thus, we suggest that developers should consider if implementing small talk capabilities is useful in their settings.

References

1. Brandtzaeg, P.B., Følstad, A.: Why People Use Chatbots. In: Kompatsiaris, Y., Cave, J., Satsiou, A., Carle, G., Passani, A., Kontopoulos, E., Diplaris, S., McMillan, D. (eds.) Internet Science. 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings, 10673, pp. 377–392. Springer International Publishing, Cham (2017)
2. Kerlyl, A., Hall, P., Bull, S.: Bringing Chatbots into education: Towards Natural Language Negotiation of Open Learner Models. In: Ellis, R., Allen, T., Tuson, A. (eds.) Applications

and Innovations in Intelligent Systems XIV, 9, pp. 179–192. Springer London, London (2007)
3. Lebeuf, C., Storey, M.-A., Zagalsky, A.: Software Bots. IEEE Softw. 35, 18–23 (2018)
4. Winkler, R., Söllner, M.: Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. In: Academy of Management Annual Meeting (AOM). Chicago, USA (2018)
5. Goel, A., Creeden, B., Kumble, M., Salunke, S., Shetty, A., Wiltgen, B.: Using watson for enhancing human-computer co-creativity. In: AAAI Symposium, pp. 22–29 (2015)
6. Tegos, S., Demetriadis, S., Karakostas, A.: MentorChat: Introducing a Configurable Conversational Agent as a Tool for Adaptive Online Collaboration Support. In: 2011 15th Panhellenic Conference on Informatics, pp. 13–17 (2011)
7. Hobert, S., Meyer von Wolff, R.: Say Hello to Your New Automated Tutor – A Structured Literature Review on Pedagogical Conversational Agents. In: 14th International Conference on Wirtschaftsinformatik, pp. 301–314 (2019)
8. Jain, M., Kumar, P., Kota, R., Patel, S.N.: Evaluating and Informing the Design of Chatbots. In: Koskinen, I., Lim, Y.-k., Cerratto-Pargman, T., Chow, K., Odom, W. (eds.) DIS '18, June 9-13, 2018, Hong Kong, pp. 895–906. ACM Press, New York, New York, USA (2018)
9. Meyer von Wolff, R., Hobert, S., Schumann, M.: How May I Help You? – State of the Art and Open Research Questions for Chatbots at the Digital Workplace. In: Proceedings of the 52th Hawaii International Conference on System Sciences, pp. 95-104 (2019)
10. Hevner, A.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19, 87–92 (2007)
11. Hevner, A., March, S., Park, J., Ram, S.: Design Science in Information Systems Research. Management Information Systems Quarterly 28, 75–105 (2004)
12. Gregor, S., Jones, D.: The Anatomy of a Design Theory. Journal of the Association for Information Systems 8, 312–335 (2007)
13. AXA Shared Services Spain S.A.: axa-group/nlp.js, https://github.com/axa-group/nlp.js
14. Bickmore, T.W., Picard, R.W.: Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact. 12, pp. 293–327 (2005)