

Using Smart Personal Assistants for Online Learning Activities: What benefits can we expect?

Rainer Winkler¹, Pablo Weingart², Matthias Söllner³

¹ University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland

rainer.winkler@unisg.ch

² University of Kassel, Information Systems, Kassel, Germany

pabloluis@posteo.de

³ University of Kassel, Information Systems and System Engineering, Kassel, Germany

soellner@uni-kassel.de

Abstract. With the increasing popularity of massive open online courses, universities are able to reach a wider audience without restriction and for comparably low costs. However, in these learning environments, educators are hardly able to offer individual support to their learners. According to learning theory, the lack of individual support in online learning environments can considerably limit learning success. In our paper, we argue that new emerging Smart Personal Assistants (SPAs) have the capabilities to address this problem by building up an interaction with the learner similar to a human tutor. Drawing on the concept of the Interactive, Constructive, Active and Passive (ICAP) Framework, we designed an experiment with 76 participants, where one group received text-based learning materials and the other group received additional scaffolds from an interactive, text-based smart personal assistant. Posttest-score comparisons revealed that learners interacting with the SPA were able to achieve higher learning outcomes.

Keywords: smart personal assistant, online learning, intelligent tutoring system, mechanical turk, experiment

1 Introduction

With the increasing popularity of massive open online courses (MOOCs), education providers have the possibility to reach a wider audience for comparatively low costs [1]. However, this positive development does not come without disadvantages. A major disadvantage of MOOCs is the lack of individual support from educators due to financial and organizational restrictions [2, 3]. According to predominant constructivist learning theories, individual support is crucial for learning success [4].

Information systems have often tried to fill this gap. There is a huge body of literature reaching about 30 years back in which IS and educational research investigated the impact of intelligent tutoring systems (ITS) on learning outcomes [5]. When compared to no tutoring, ITS showed that they are able to improve learning outcomes to the same extent as human tutoring [6]. Some studies also showed that

ITS are able to improve learners' problem-solving skills in selected domains [7, 8]. Although there is general consensus about the effectiveness of ITS, they are still not integrated enough into learning environments [9]. One reason for that might be that developing and introducing ITS often requires a high amount of technological knowledge and required time from the educator [10]. Thus, they are mainly used by technology-savvy domains, such as computer science. For example, Hooshyar et al. [7] and Wang et al. [8] introduced an ITS in a computer science course where learners received help to conduct problem-solving activities. Furthermore, ITS applications are often complex to transfer from one context to the other and do not work on different devices and operating systems that learners use every day [10]. Smart personal assistants (SPA) are a promising way to address the mentioned problems [11]. SPAs are computer programs that are able to interact with users on a specific topic in natural language [12]. Natural language means that users can use their normal language similar to human-human communication. SPA can be text-based or voice-based. SPA providers offer rich ecosystems with intuitive and easy-to-use toolkits where educators can create SPAs without having deep programming knowledge. The use and popularity of SPAs has risen sharply in the recent years [13]. In the United States, 59 percent of respondents from the ages of 18-24 stated that they were heavy (at least once per day) users of smart personal assistants [14]. However, the use of SPAs for online learning environments is rare and only few studies exist that concentrate on SPAs in online learning environments. For example, Song et al. [15] designed and developed a SPA to reflect with the learners about their learning experience in online courses. They found out that most of the students (98.5%) used the SPA during the course. Nevertheless, empirical evidence about the effect of SPAs for online learning activities is still scarce [16]. Thus, our study tries to answer the following research question:

RQ: Do learners improve their learning outcomes when interacting with a smart personal assistant in online learning activities?

To answer this research question, we developed and integrated a text-based SPA on Amazon Mechanical Turk (MTurk), where 76 participants had to solve a 20-minute task. One group received only a text as input while the other group was additionally able to interact with a SPA. Both groups had the same amount of time and the amount of information provided was equivalent for both groups. Based on the Interactive, Constructive, Active and Passive (ICAP) Framework, we argue that the interaction with the SPA brings participants into a more interactive learning mode, which is considered as the gold-standard of learning.

2 Theory and Hypotheses Development

2.1 Smart personal assistants in education

In general, SPAs are computer programs that are able to interact in natural language with users on a specific topic or domain [17]. SPAs include all types of software that

enable people to converse with computers. The main task of the SPA is to improve the user's perception, cognition and/or ability to act [18].

The main advantages of using SPAs are the low cost of services and higher user satisfaction due to fast response times [11]. Compared to other e-learning technologies, SPAs stand out for their ability to interact with learners synchronously and to respond to their individual needs, allowing them to control their own learning process [11]. SPAs can be divided into two types: (1) built-in SPAs that use multi-purpose devices and (2) stand-alone SPAs that use dedicated devices. Examples of built-in SPAs include Siri (for Apple products) and Cortana (for Windows-based PCs). Examples of stand-alone SPAs include Alexa (that uses Echo, Echo Dot and Tab dedicated devices) and Google Assistant (that uses Google Home dedicated devices, 19).

SPAs, such as Amazon's Alexa, Google's Assistant, Apple's Siri, and others are running on "SPA-enabled devices" (endpoints) such as Apple's iPhone, iPad and Mac, Amazon's Echo, Google's Home, etc. The main functionality, the "brain" of a SPA, is typically housed as a cloud service that processes voice or text data [20]. This new generation of SPAs has achieved commercial success in contrast to previous implementations, and also shows significant progress in terms of natural language capability and its correctness. The use of SPAs for online learning activities has several benefits. First, it allows educators to develop SPAs for their own learning environments without being dependent on software designers or having to invest a large amount of time. Second, it allows educators to convert the use of SPAs to different contexts in an efficient way. And third, SPAs are becoming daily companions on everyday devices of students in their private and school life. Researchers pretty much agree that SPAs are able to improve learning outcomes. For example, Hooshyar et al. [7] introduced an SPA to aid learners in learning computer programming.

2.2 ICAP framework

The ICAP framework proposed by Chi and Wylie [21] is based on a constructivist view of learning. It explains the process of effective learning by classifying observable learner behaviours into four modes: interactive, constructive, active and passive. It predicts that these modes will be ordered by effectiveness as follows: *interactive* > *constructive* > *active* > *passive*. Educators have long recognized that although learners can learn from receiving information passively, they learn much better by learning actively [22]. Learning actively requires learners to engage cognitively and meaningfully with the tasks they are dealing with. They think about their learning material in depth rather than just passively receiving it [23]. Each mode of the ICAP framework corresponds to different types of behaviours and knowledge-change processes predicting different learning outcomes. Table 1 outlines the different modes (from passive to interactive) showing that for different learning scenarios (listening to a lecture, reading a text, and observing a video), learners show different modes of behavior.

Table 1. Distinction of Learner Behavior according to ICAP Framework adapted from Chi and Wylie [21]

	PASSIVE (Receiving)	ACTIVE (Manipulating)	CONSTRUCTIVE (Generating)	INTERACTIVE (Dialoguing)
Reading a text	Reading entire text passages silently/aloud without doing anything else	Underlining or highlighting; Summarizing by copy- and-delete	Self-explaining; Integrating across texts; Taking notes in one's own words	Asking and answering comprehension questions with a partner

A passive learner behavior describes learners who deal with the presented instructional information without additional physical activity (e.g., reading a text without doing anything else). An active learner behavior includes “doing something physically” (e.g., underlining or highlighting a text). A constructive learner behavior requires “producing outputs that contain ideas that go beyond the presented information” (e.g., taking notes in one’s own words). An interactive behavior requires “dialoguing extensively on the same topic, and not ignoring a partner’s contribution”. This means that both partners’ utterances are primarily constructive, and a sufficient degree of turn-taking occurs (e.g., asking and answering comprehension questions with a partner [21]). For example, Gobert and Clement [24] compared learning gains among learners who studied learning materials by only reading a text (passive), writing summaries of the text (active), and drawing diagrams from the text (constructive). As predicted by the ICAP hypothesis, the results showed that the *constructive* (diagram) group did better on measures of learning outcomes than the *active* (summary) group did, which in turn did better than the *passive* reading group. We argue that SPAs in learning activities might be able to help students engaging deeper with the learning materials compared to non-interactive text materials only.

2.3 Learning outcomes

The learning outcome describes knowledge and skills that a learner has acquired as a result of the learning process [25]). Learning outcomes are used to assess and compare the performance of participants and organizations and can be divided into self-assessment and test-based assessment. In the former, the learner evaluates his own learning outcome. In the latter case, the assessment is carried out by third parties according to a standardized procedure [26]. The literature shows that learners often accurately assess their own learning performance. Although these assessments are distorted to a certain extent, they are still a suitable supportive construct for measuring learning success [26]. To get a holistic view of learning success, both approaches should be investigated in parallel [27].

Properly designed SPAs used in learning activities have the capabilities to build up meaningful interactions with learners bringing them from a passive to an interactive learning mode [21], [27]. Thus, we propose:

H1a: Learners who interact with a SPA during online learning activities achieve higher levels of objective learning outcomes compared to those learners who learn without this interaction.

H1b: Learners who interact with a SPA during online learning activities achieve a higher level of subjective learning outcome compared to those learners who learn without this interaction.

3 Method

3.1 Tutor design

The SPA tutor was developed based on Python and used the MTurk API to automate the experiment. ParlAI [28] was used for higher abstraction and simplification of the MTurk-API in dialogue-oriented applications. Initially, we wanted to provide a more sophisticated tutor assistance by using deep learning models. We used the site NLPPROGRESS.COM to select the question answering algorithm as a possible algorithm which could be used to provide such assistance by answering questions from learners by searching a background information text. The pretests revealed that questions of learners showed a high variance whose answers therefore could not be found in the background information text. Furthermore, the answers were not always correct. Therefore, we assumed that these deep learning models are not yet capable to provide sufficient value in our generic research context and focused on the described simpler case. We used a proactive logic where the tutor was able to ask students comprehension questions. The guidance by the tutor and the answers by the learner were transmitted by chat messages in the MTurk interface as seen in Figure 2. All five questions can be found in the Appendix.

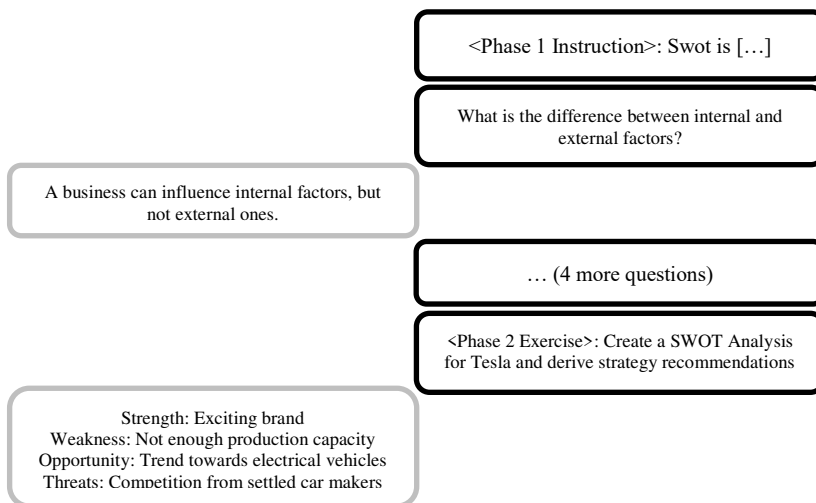


Figure 1. Dialogue overview

The tutor was perceived to be a text-based SPA, and can be classified as Chatbot Assistant according to the archetypes of Knotte et al. [29]. Table 2 summarizes the properties of the tutor. The tutor communicates text-based on both the input and output sides. The direction of interaction is mainly system-to-user because the SPA initiated the interaction with asking questions about the text.

Table 2. Degree of interaction and intelligence

Interaction	
Communication mode	Text-based
Direction of explicit interaction	System-to-User; User-to-System
Query Input	Text-based
Response Output	Text-based
Action	No Service Execution
Intelligence	
Assistance Domain	General Assistance
Accepted Commands	Manual Data Entry
Adaptivity	- / static
Collective Intelligence	-
Embodiment	-

3.2 Sample

To answer our research questions and test our hypothesis, we conducted an online experiment on Amazon MTurk. The tested sample is composed as follows: 78.9% of treatment group 1 (TG) participants (n = 30) and 94.7% of TG2 participants (n = 36) live in the USA. India follows with 10.5 (n = 4) and 5.3% (n = 2) of the participants respectively. The remaining four TG1 participants come from the United Kingdom, Brazil, the Netherlands and Australia. 68.4% of the TG1 participants (n = 26) and 63.2% of the TG2 participants (n = 24) indicated that they are male. The average age of both TG1 and TG2 is in the category “30 to 34 years old”. Most of the learners have at least a high school degree (> 80%). The t-test and the chi-square test were used to check whether the two groups were equal regarding age, gender, nationality and educational background. No differences were found ($p > 0.05$). Furthermore, we used linear regression to check that none of the mentioned characteristics had an effect on our dependent variables. The payment per fulfilled task was \$1.80.

3.3 Experimental procedure

We employed a fully randomized online experiment over a period of two months. Amazon MTurk was chosen as a platform to perform the experiment online. Apart from traditional online surveys, MTurk offers the possibility to create richer tasks that meet the requirements of our experiment [30]. The experiment lasted 20 minutes at

maximum. We chose to use a SWOT analysis task because of different reasons. First, we wanted to ensure that the participants do not have much pre-knowledge of the task, because this would disturb our experiment. Second, our aim was to not only ask participants about reproducing information but also to address higher learning outcomes. In the experiment, we created a SPA that had an interaction with the learner during the instruction phase.

The experiment consisted of two phases: the instruction phase and the exercise phase. In the instruction phase, learners of both groups received a text about the main components of a SWOT analysis and its purpose of use. The SPA group (TG2) was able to have an additional conversation with the SPA. The SPA asked questions related to the elements of a SWOT analysis and additional examples. The time given for the instruction phase was equal in both groups and the information provided by the SPA was equivalent to the information in the text. The outline of the experiment procedure can be seen in Figure 1.

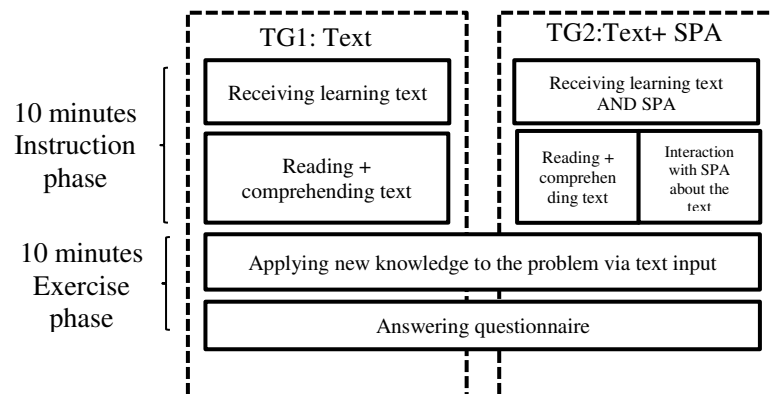


Figure 2. Experiment procedure

In the exercise phase, the participants of both TGs were supposed to apply their newly learned knowledge by applying a SWOT analysis to the company “Tesla”. The task was the following:

Create a SWOT Analysis for Tesla and derive strategy recommendations from it.

Finally, we collected their solutions, and learners of both TGs answered the questionnaire regarding demographics and the subjective learning outcome. We argue that the learner should have a tendency to show an interactive learning behavior with the help of a SPA resulting in increased learning outcomes. Chi and Wylie [21] highlight exemplary learning activities that help students to move into an interactive learning mode. One of the examples they mentioned was the following: “Asking and answering comprehension questions to a partner”. Based on that, our text-based SPA asked the learners five comprehension questions, which were answered by the learners of TG2. They were intended to encourage them to reflect on what they read in the text. The reading of the text and the interaction with the text lasted 10 minutes (same time as in the control group).

After the instruction phase, the learners had to complete the task. Afterwards, the learners answered a questionnaire to collect further constructs to be described and evaluated in the following sections. The questionnaire was released and answered immediately after the task was completed. Part of this questionnaire was also an open question, which should qualitatively examine whether and in which way the interaction with the SPA was helpful and meaningful. This question was only asked to the participants of the interactive TG2.

3.4 Measurement and Analysis

The **objective learning outcome** was measured by evaluating the learners' solutions of the tasks. The evaluation of the solution was conducted by two raters with a commonly discussed evaluation framework. The two raters rated the solutions blindly and independently from each other. Both raters are experienced in rating SWOT analysis and were active in university teaching in the past. The final objective learning outcome score was calculated by the average of the two raters' scores. Each correct example for a category within the SWOT analysis was evaluated with the same number of points.

The **subjective learning outcome** was measured in the post-test questionnaire with a seven point Likert scale ranging from strongly agree (1) to strongly disagree (7). The variable was intended to measure the learning outcome from the learner's point of view and thus reflect the construct described in the theory. Before evaluating the data, the negatively interpreted items were inverted so that a higher value in the data analysis corresponds to a better result. In the case of a negative item, this inversion was not necessary. The items are based on Gupta and Bostrom [31] and, according to Cronbach's Alpha, have an internal consistency of 0.91 and thus exceed the standard of 0.8. The items are shown in the Appendix.

Additionally, we measured the **used time** for the whole exercise to see if there are differences between the groups. The time was programmatically collected via MTurk. Starting point marked the acceptance of the task in the MTurk interface with simultaneous transmission of the learning text, so that the text could be read. The time was stopped when the task was finished by the participant or the assigned time had expired (maximum 20 minutes). The duration of the questionnaire was not considered because the focus was on the learning process itself.

In order to examine to what extent the SPA interaction might have helped learners to understand the SWOT analysis, the TG2 participants were asked the following question.

“Did the questions that were provided support you in learning SWOT analysis and in what way? (1-2 sentences/bullet points)”

Following the method of Ryan and Bernard [32], we identified topics from the answers by using “repetitions” and “similarities & differences” .

For data analysis, we used a linear regression model to see whether the scores are significantly different. We used R for the statistical evaluation. In the further data analysis, it was first examined whether the assumptions of linear regression, i.e. in particular normal distribution, homoscedasticity and linearity, are given for the

models set up. The R library "gvlma" of Peña and Slate [33] was used for this test. If the hypotheses could be accepted, the effect strength was calculated according to Cohen's d [34].

4 Results

Table 2 shows the averages and standard deviations for the data collected for all variables. Brackets next to the variable name indicate the ranges that the values could assume. A higher average corresponds to a better result.

Table 3. Mean and standard deviation for the variables

	TG1		TG2	
	Mean	SD	Mean	SD
Objective learning outcome (0-12)	6,1578	3,7885	7,7631	2,7453
Subjective learning outcome (1-7)	4,0368	0,9697	4,4316	1,0214
Used time in seconds (0-∞)	540	208	545	223

The results of the regressions for hypotheses 1a and 1b are presented in Table 3. The coefficients marked with * declare significance levels as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

We identified a significant difference in objective learning outcome ($t = 2.115$, $p = 0.0378$). Cohens d is 0.4852. Thus, we can support H1a. The corresponding confidence interval is between 0.0214 and 0.949. This is almost a medium effect according to the rule of thumb of Cohen. For the subjective learning outcome (H1b), no statistically significant difference could be detected ($t = 1.728$, $p = 0.0882$). If the weaker standard of $p = 0.1$ is applied, there is a statistically significant difference. Cohen's d would then indicate a small effect with $d = 0.396$ at a confidence interval of -0.0652 and 0.858. Moreover, no significant difference in the time used could be identified ($t = 0.096$, $p = 0.924$).

Table 4. Regression results

Variable	Coefficient	p-value
H1a: Objective learning outcome	1,6053	0,0378**
H1b: Subjective learning outcome	0,3947	0,0882*

A total of 76.3% ($n = 29$) of the participants responded positively to the questions. 10.5% ($n = 4$) of the participants tended to express a negative opinion on the questions. The remaining 13.2% ($n = 5$) of the answers could not be assigned because they did not directly refer to the question. In total, three topics were identified. They are displayed below and describe how the learners were supported by the questions.

Practical relevance and examples describes that the SPA helped them to transfer the acquired knowledge to the task. The learners were able to more easily integrate the learning content with their own experiences and existing knowledge.

Structuring aid refers to the ability of the SPA to help participants structure the learning content. This way, it was easier for them to remember the content when conducting the task. The participants have emphasized that this structuring aid in particular has enabled them to reduce the SWOT analysis to its essential content.

Internalization refers less to the understanding of learning than to the fact that the SPA makes the learning process more memorable and challenges the learner interactively. This gave the learner impulses to internalize the learning content in its essential parts so that it can be retrieved more easily later.

In the exercise, the SPA supports learners by guiding them through the task.

5 Discussion

In our paper, we investigated whether participants are able to increase their learning outcomes when interacting with a SPA. Drawing on the concept of the ICAP framework, we created a SPA and implemented it on Amazon MTurk. The findings revealed that participants interacting with SPAs were able to significantly increase their learning outcomes compared to participants having no interaction. This goes in line with past research in the field of multimedia learning that was able to show that interactivity is a key design principle when designing learning material [35]. For example, Evans and Gibbons were able to find out that students using an interactive learning system outperformed those using a non-interactive system. Moreover, the results of this work confirm the revealed positive effects in existing research on intelligent tutoring systems [36], and SPAs in particular [37]. For example, Winkler et al. [37] found out that SPAs are able to increase group task outcome in a collaborative problem-solving setting. Regarding the subjective learning outcome, no significant differences were found. One reason for that might be that the positive learning effect with the SPA was possibly offset by the effect that the SPA showed learners their knowledge gaps which is reflected in a less well perceived learning effect.

Interestingly, although one would expect that the participants need more time to read and interact with the SPA, they did not need additional time. This let us assume that the learning process of both TG was fundamentally different and that the interaction with the SPA was not just an add-on to the reading of the text. Some participants might have skipped the reading of the text and started to interact with the SPA earlier, which shows that they preferred the conversational style of interaction.

Our work connects to the already shown positive effects of SPAs on learning outcomes. In specific, this study makes several contributions to research. First, we contribute to intelligent tutoring systems and online learning research by showing that new emerging SPAs are a promising solution to address the problem of individual support in online learning environments such as MOOCs. When learners are able to interact with SPAs individually (e.g., ask questions and receive immediate, personalized answers), they might be more motivated, do not quit the course that fast

and might show better results in the end. This would help educational institutions and lecturers to reduce costs while improving the quality of their learning material. Second, we contribute to research around the ICAP Framework by indicating that IT might be able to bring learners into an interactive learning mode, which can be seen as the gold-standard in learning. Past research mainly focused on human partners to achieve learners' interactive behavior. We showed that new emerging kinds of IT can bridge the gap between learners' constructive and interactive learning mode. Third, we showed that Amazon MTurk might be a suitable platform for SPA research to test different aspects of SPAs in online learning environments without being dependent on resource-intensive laboratory and field experiments.

Our work also has several implications for practice. First, we argue that the implementation of SPAs in online learning environments might enrich learners' learning outcomes. SPA providers such as Amazon or Google offer large eco-systems that allow an educator to create their own SPAs with little or no programming knowledge. With this work, an educator might be motivated to create a SPA for their own learning environment. Second, MOOC providers such as Coursera and EdX might be inspired to offer SPA templates to their course providers in order to increase learning outcome and their own business value.

With regard to the type of survey, the data collected and the corresponding analysis showed that MTurk is a suitable platform for evaluating SPAs in the e-learning context. The initial fear of time scarcity could not be confirmed, so that an increase in complexity seems possible. The samples show that a large part of the MTurk workers come from the USA and a smaller one from India. Overall, the structure of the samples is similar to that of recent studies on the demography of MTurk workers [38].

6 Limitations and Future Research

There are a number of limitations to this study that should be noted. First, we used Amazon MTurk to examine the impact of a SPA on learning success. Despite the benefits of quick and reliable data, Amazon MTurk also has its downsides. There is a relatively small sample of MTurk workers who participate in academic studies each year. This means that most participants have been exposed to several experimental manipulations and, thus, might be biased. Moreover, MTurk workers might be less diverse than the general population. The population of MTurk workers is significantly less politically diverse, more highly educated, younger, and less religious compared to the normal population. However, in our study, we are interested in how interactions with SPAs influence online course students. Thus, the population of MTurk might be suitable. Nevertheless, it would be interesting to see if experiments with other samples can show similar effects. Second, the experiment duration was relatively short and under positive conditions. Such studies tend to produce better outcomes due to novelty effects and hyperattention to experiment details [39]. For future research, it would be interesting to see whether SPAs are able to increase learning outcomes over a longer period of time (e.g., over the period of an online course). Third, we used a SPA that was not able to adapt its questions to the answers of the learner. Thus, we

cannot be sure that the SPA was able to bring learner into an interactive learning behavior. Future studies should try to build SPAs that are able to analyze users' answers and can act accordingly. Fourth, we only look at short-term learning outcomes. To identify the true impact of SPAs on learning, we also need insights about their long-term learning effects. Fourth, we tested the use of the SPA for a very narrow context (SWOT analysis task). Future research might test SPA interactions for different kinds of tasks and different kinds of knowledge (factual knowledge, applying knowledge, etc.). Fifth, the development of such a SPA is quite time-consuming and is hardly transferable to different learning contexts. This raises the question about its practical value of SPAs. Future research should try to develop design principles or patterns that help educators to create SPAs in an efficient way.

7 Conclusion

Our work answered the question whether the use of SPAs for online learning activities could increase learning outcomes. We designed and introduced a SPA on the Amazon MTurk platform where participants had to solve a SWOT analysis task with the help of SPAs. One group received text material only whereas the other group received text material and was able to interact with the SPA. The SPA interactions helped the learners to engage with the learning material in a meaningful way, resulting in increased learning outcomes. These findings let us assume that SPA interactions in online learning activities might be able to bring students into an interactive learning behavior, which is considered as gold standard according to the ICAP-Framework [40]. These findings contribute to intelligent tutoring system and online learning research. While further research is required, it appears that the use of new emerging SPAs are a promising enrichment for online learning environments.

References

1. Vardi, M.: Will MOOCs destroy academia? *Communications of the ACM* 55, 5 (2012)
2. Doherty, I., Harbutt, D., Sharma, N.: Designing and Developing a MOOC. *Med.Sci.Educ.* 25, 177–181 (2015)
3. North, S.M., Richardson, R., North, M.M.: To Adapt MOOCs, or Not? That Is No Longer the Question. *Universal Journal of Educational Research* 2, 69–72 (2014)
4. Vygotsky, L.S.: *Mind in society: The development of higher psychological processes*. Harvard university press (1980)
5. Kulik, J.A., Fletcher, J.D.: Effectiveness of Intelligent Tutoring Systems. *Review of educational research* 86, 42–78 (2016)
6. Graesser, A.C., Hu, X., Sottolare, R.: Intelligent tutoring systems. In: *International handbook of the learning sciences*, pp. 246–255. Routledge (2018)
7. Hooshyar, D., Ahmad, R.B., Yousefi, M., Fathi, M., Horng, S.-J., Lim, H.: Applying an online game-based formative assessment in a flowchart-based

- intelligent tutoring system for improving problem-solving skills. *Computers & Education* 94, 18–36 (2016)
8. Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., Ren, G.: A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. *Computers & Education* 81, 102–112 (2015)
 9. Nye, B.D., Pavlik, P.I., Windsor, A., Olney, A.M., Hajeer, M., Hu, X.: SKOPE-IT (Shareable Knowledge Objects as Portable Intelligent Tutors): overlaying natural language tutoring on an adaptive learning system for mathematics. *International journal of STEM education* 5, 12 (2018)
 10. Elaine, K.: Design of a domain-independent, interactive, dialogue-based tutor for use within the GIFT framework. In: *Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3)*, p. 161 (2015)
 11. Winkler, R., Söllner, M.: Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. In:
 12. Chung, H., Iorga, M., Voas, J., Lee, S.: Alexa, can I trust you? *Computer* 50, 100–104 (2017)
 13. eMarketer: Alexa, Say What?! Voice-Enabled Speaker Usage to Grow Nearly 130% This Year, <https://www.emarketer.com/Article/Alexa-Say-What-Voice-Enabled-Speaker-Usage-Grow-Nearly-130-This-Year/1015812>
 14. Statista: US Voice Technology Usage by Age, <https://www.statista.com/statistics/879907/us-voice-technology-usage-age/>
 15. Song, D., Oh, E.Y., Rice, M.: Interacting with a conversational agent system for educational purposes in online courses. In: *2017 10th international conference on human system interactions (HSI)*, pp. 78–82 (2017)
 16. Winkler, R., Söllner, M.: Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. In:
 17. Huang, J., Zhou, M., Yang, D.: Extracting Chatbot Knowledge from Online Discussion Forums. In: *IJCAI*, 7, pp. 423–428 (2007)
 18. Jalaliniya, S., Pederson, T.: Designing wearable personal assistants for surgeons: An egocentric approach. *IEEE Pervasive Computing* 14, 22–31 (2015)
 19. Chung, H., Iorga, M., Voas, J., Lee, S.: Alexa, Can I Trust You? *Computer* 50, 100–104 (2017)
 20. Cheung, J.H., Burns, D.K., Sinclair, R.R., Sliter, M.: Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology* 32, 347–361 (2017)
 21. Chi, M.T.H., Wylie, R.: The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist* 49, 219–243 (2014)
 22. Bonwell, C.C., Eison, J.A.: *Active Learning: Creating Excitement in the Classroom*. ASHE-ERIC Higher Education Report No. 1. ERIC, Washington (1991)
 23. King, A.: From sage on the stage to guide on the side. *College teaching* 41, 30–35 (1993)
 24. Gobert, J.D., Clement, J.J.: Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic

- knowledge in plate tectonics. *Journal of Research in Science Teaching* 36, 39–53 (1999)
25. Otter, S.: Learning outcomes in higher education (1995)
 26. Douglass, J.A., Thomson, G., Zhao, C.-M.: The learning outcomes race: The value of self-reported gains in large research universities. *Higher education* 64, 317–335 (2012)
 27. Caspersen, J., Smeby, J.-C., Olaf Aamodt, P.: Measuring learning outcomes. *European Journal of Education* 52, 1–11 (2017)
 28. Miller, A.H., Feng, W., Fisch, A., Lu, J., Batra, D., Bordes, A., Parikh, D., Weston, J.: Parlai: A dialog research software platform. *CoRR* (2017)
 29. Knotte, R., Janson, A., Söllner, M., Leimeister, J.M.: Classifying Smart Personal Assistants: An Empirical Cluster Analysis. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pp. 2024–2033 (2019)
 30. Chandler, J., Mueller, P., Paolacci, G.: Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior research methods* 46, 112–130 (2014)
 31. Gupta, S., Bostrom, R.: Research note—An investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning. *Information Systems Research* 24, 454–469 (2013)
 32. Ryan, G.W., Bernard, H.R.: Techniques to Identify Themes. *Field Methods* 15, 85–109 (2003)
 33. Peña, E.A., Slate, E.H.: Global validation of linear model assumptions. *Journal of the American Statistical Association* 101, 341–354 (2006)
 34. Cohen, J.: A power primer. *Psychological bulletin* 112, 155–159 (1992)
 35. Evans, C., Gibbons, N.J.: The interactivity effect in multimedia learning. *Computers & Education* 49, 1147–1160 (2007)
 36. Kulik, J.A., Fletcher, J.D.: Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of educational research* 86, 42–78 (2016)
 37. Winkler, R., Söllner, M., Neuweiler, M.L., Rossini, F.C., Leimeister, J.M.: Alexa, Can You Help Us Solve This Problem? How Conversations With Smart Personal Assistant Tutors Increase Task Group Outcomes. In: *CHI'19 Conference on Human Factors in Computing Systems*, pp. 1–6. SIGCHI, Glasgow (2019)
 38. Difallah, D., Filatova, E., Ipeirotis, P.: Demographics and dynamics of mechanical turk workers. In: *Proceedings of the eleventh acm international conference on web search and data mining*, pp. 135–143 (2018)
 39. Cheung, A.C.K., Slavin, R.E.: The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational research review* 9, 88–113 (2013)
 40. Chi, M.T.H., Wylie, R.: The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational psychologist* 49, 219–243 (2014)

Appendix

Introductory text

“SWOT is a strategic planning technique used to help a person or organization identify strengths, weaknesses, opportunities, and threats related to business competition or project planning.

The aim of SWOT is to specify the objectives of the business venture or project and identify the internal and external factors that are favorable and unfavorable to achieving those objectives.

SWOT consists of internal factors and external factors.

The internal factors are the strengths and weaknesses internal to the organization.

External factors are the opportunities and threats presented by the environment external to the organization.

Examples for internal factors are financial performance, human capital, know-how, image, customer loyalty, patents, corporate culture, leadership, human capital satisfaction or innovativeness.

Examples for external factors are competitors as well as technological, cultural, social and political trends and changes.

The difference between internal factors and external factors is that a business can actively influence internal factors, while a business for the most part only react to the external factors which affect the company's environment.

To apply SWOT analysis, you have to name examples for each category for a given business.”

Questions

1. What is the difference between a strength and an opportunity?
2. Please name a Strength for Amazon.
3. Please name a Weakness for Volkswagen (german diesel car manufacturer).
4. Please name an Opportunity for SpaceX (the space company by Elon Musk).
5. Please name a Threat for Facebook.

Subjective learning outcome items

Please rate your level of agreement with the following statements.

1. I have mastered the use of SWOT analysis.
2. I cannot yet use SWOT analysis as well as I would like.
3. I am able to perform SWOT analysis well.
4. It is not yet possible for me to use SWOT analysis at the level I like.
5. I think my ability to use SWOT analysis can be improved substantially.