Please quote as: Winkler, R., Söllner, M., and Leimeister, J. M. (2020) Improving Students? Problem-Solving Skills with Smart Personal Assistants. In 80th Annual Meeting of the Academy of Management, Academy of Management, Vancouver.

Improving Students' Problem-Solving Skills with Smart Personal Assistants

ABSTRACT

Tomorrow's organizations need employees who are able to deal with rapid changes and solve non-routine problems. Gaining problem-solving skills is considered the number-one skill for future employees to succeed professionally. Predominant learning theories agree that the most effective way to gain these skills is for everyone to receive individual support by their own private tutor. For educational institutions such as high schools and universities, this is often not possible due to financial and organizational restrictions. A new emerging class of information technology – specifically Smart Personal Assistants (e.g., Google's Assistant or Amazon's Alexa) – has the potential to address this problem by interacting with students in a manner comparable to human tutors because of its high degree of adaptability, interactivity and accessibility. Even though there exists a growing body of research about the design and use of Smart Personal Assistants for learning, empirical evidence of their ability to help students improve their problem-solving skills is still scarce. Grounded on technology-mediated learning theory, this study uses a mixed-method approach consisting of two field quasi-experiments and one post-experiment focus group discussion at a business high school and a vocational business school with a total of 90 students to measure the effect of using Smart Personal Assistants on acquiring problem-solving skills. The empirical results show that students in the experiment classes acquired significantly more problem-solving skills than those in the control group mainly explained by changes in their learning process. The findings provide empirical evidence for the importance of using new emerging Smart Personal Assistants on general skill development, and specifically on problem-solving skill development. Moreover, our work can guide educational institutions and educators in designing and implementing Smart Personal Assistants in their own learning environments.

Keywords: Smart personal assistant, Intelligent tutoring system, Problem-solving skills, Technology-mediated learning, Education, Mixed-method, Field quasi-experiment.

INTRODUCTION

Tomorrow's organizations increasingly need employees who are able to deal with fast-changing environments and solve non-routine problems (Creative problem solving, 2014). There is clear evidence of this change in the demand for skills in several countries, such as Germany and the United States (David et al., 2006; Spitz-Oener, 2006). Past research shows, that high school business students seldom lack problem-solving skills to solve complex, real-world problems (Bartlett, 2002). According to predominant constructivist learning theories, students need individual interaction with a personal tutor to best learn these skills (Vygotsky, 1978). However, educational institutions such as high schools and universities struggle to offer this kind of individual support to students due to financial and organizational constraints (Rietsche, Duss, Persch, & Soellner, 2018). The growing number of classroom sizes at high schools, large-scale lectures at universities with more than 100 students per lecturer and massive open online courses (MOOCs) with more than 1,000 participants, make individual interaction with a teacher or tutor even more difficult (Oeste, Lehmann, Janson, Söllner, & Leimeister, 2015). The tension between increasing student-educator ratios and the need for individual interaction raises the question of how to offer individual support to students to enable them to gain the necessary problem-solving skills.

Research in the area of technology-mediated learning (TML) has tried to address this challenge by leveraging the potentials of IT. In specific, there is a huge body of literature reaching back over 40 years in which IS research and educational research have investigated the impact of computer tutoring systems on learning outcomes (Kulik & Fletcher, 2016). Despite their proven effectiveness in the past (Graesser, Hu, & Sottilare, 2018), existing computer tutoring implementations can sometimes be considered as rather static with only short, text-based types of answers being possible (Al-Bastami & Naser, 2017; Baker, 2016) or as relying on rather complicated software architectures that require a lot of technical know-how, time and effort (Afzal et al., 2019). This is why paper-based learning materials are still the most widely-used learning aids for acquiring skills outside lectures.

A new emerging class of information systems called Smart Personal Assistants (SPAs, e.g., Google's Assistant or Amazon's Alexa), have the potential to address this gap. An SPA is an application that uses inputs, such as the user's voice, vision (images), and contextual information, to provide assistance by answering questions in natural language, making recommendations, and performing actions (Hauswald et al., 2015). This means that SPAs are able to modify their answers based on users' utterances and/or contextual information, which helps them to build up an almost natural dialogue with their users. Globally, the market of SPAs is predicted to increase from US$3 billion in 2017 to US$15.7 billion in 2021 (Statista, 2019b). 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 SPAs (Statista, 2019a). Recent research indicates that education is a very promising application area for SPAs (Arend, 2018; Goksel-Canbek & Mutlu, 2016). In contrast to existing computer tutoring systems, SPAs might be better able to imitate educators' scaffolding behavior which can be seen as a gold standard (Merrill, Reiser, Ranney, & Trafton, 1992). As common for emerging research fields in IS, the majority of existing empirical research in this area is rather explorative (observations, qualitative interviews with teachers, etc.) and few empirical studies that rigorously measure the effectiveness of SPAs in learning environments exist (Dousay & Hall, 2018; Jean-Charles, 2018). Especially when it comes to longitudinal effects, e.g., whether SPAs can help to train certain skills over the course of a longer period of time, empirical studies are missing. This is crucial, since we know from prior research that the development of skills takes time, and can hardly be achieved, e.g., within a single short experiment (Soderstrom & Bjork, 2015). To contribute to a better understanding of the influence of SPAs on skill development, this paper seeks to answer the following two research questions:

RQ1: Does using a Smart Personal Assistant help students develop their problem-solving skills?

RQ2: Which aspects of Smart Personal Assistants affect the learning process of students?

To answer our research questions, we employ a mixed-method approach, complementing the quantitative results of two field quasi-experiments (RQ1) with the qualitative data of a post-experiment focus group discussion (RQ2). Specifically, we implement SPAs in two different types of business schools in a European country over a period of five weeks each. Our field quasi-experiments are driven by the hypothesis that students learn better with using SPAs compared to traditional, static learning materials. The theoretical foundation for this hypothesis is based upon the ICAP framework Chi and Wylie (2014) and other education research. We argue that SPAs are better able to change students learning process by being adaptive, interactive and accessible. Our study provides strong evidence for our hypothesis. Students interacting with SPAs during their homework assignments showed significantly more problem-solving skills than students working with static learning materials. This study contributes to two different research areas in management education. First, we contribute to TML research by empirically proving how SPAs are able to change the learning process of students resulting in increased skill development. Second, we contribute to technology-enhanced scaffolding by providing evidence that adaptive and interactive scaffolds contribute significantly more to skill development compared to static scaffolds. To provide practical implications, the study exemplarily showed how to implement SPAs in an existing learning environment.

The remainder of this paper is organized as follows. The next section (Section 2) explains the theoretical background and hypothesis development. Section 3 describes the pre- and post-test field quasi-experiments and the focus group discussion in more detail, including the task design, the design and functionality of the SPA, the experimental process and the measurement and analysis of our constructs. Section 4 presents the results for research questions 1 and 2. We finish with the discussion, the limitations of our study, possible avenues for future research, and our conclusion.

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

In this section, we define the term Smart Personal Assistant for learning purposes and use educational research that provides potential explanations about the how SPAs may improve skill development.

Smart Personal Assistants

A *Smart Personal Assistant* (SPA) is an application that uses natural language input such as the user's voice, vision (images), and contextual information, to provide assistance by answering questions in natural language, making recommendations, and performing actions (Hauswald et al., 2015). SPAs have an agent program running on SPA-enabled devices (endpoints), such as Apple's iPhone, iPad, and Mac or Amazon's Echo or Google's Home, etc. They can be considered as smart because their main functionality (the "brain" of an SPA) is typically housed as a cloud service that uses machine learning and natural language processing techniques to handle voice data (converting voice-to-text, performing linguistic context analysis, and providing answers to questions; Chung, Iorga, Voas, & Lee, 2017). SPA providers offer rich ecosystems with intuitive interfaces that allow a large number of users to create their own skills without having deep programming knowledge, thereby increasing SPA providers' own business value. 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; Chung et al., 2017). In our study, we focus on both types of SPAs by enabling access via different devices.

SPAs that are used to improve learning outcomes can be seen as a subcategory of intelligent tutoring systems (ITS). Researchers in the area of IS and education have often explored the effectiveness of ITS (Kulik & Fletcher, 2016). However, most of the currently-implemented ITS are text-based and rely on somewhat complicated software architectures that require a lot of technical know-how, time and effort from educators (generally 200 hours of development time for one hour of teaching/instructions; Hilles & Naser, 2017). Thus, many researchers describe ITS development as inefficient and notoriously costly (Barbhuiya, Karim, & Suraiya, 2011).

Given that the design and use of SPAs is an emerging field in education, most of the studies are explorative and few empirical studies rigorously measuring the effect of SPAs on learning performance exist. Canbek and Mutlu (2016) examined the potential use of SPAs in learning environments by conceptually investigating potential use cases of different SPAs, such as Apple's Siri and Amazon's Alexa. They found out that one major benefit of SPAs is to improve students' listening and speaking skills without needing human tutors. Dousay and Hall (2018) observed how teachers and administrators perceive the implementation of 90 Amazon's Alexa Echo Dots in four school districts in the US with approximately 900 students. Teachers and administrators used Alexa to set reminders for activities, events, student dismissals, student medications, etc. They found out that one of the biggest challenges of implementing SPAs was the training of teachers and administrators. Jean-Charles (2018) focused more on the perception of teachers regarding using SPAs in their learning environments. In specific, he asked pre-service teachers about their perspective on using Google's Assistant in the classroom. One of the biggest concerns of the teachers was to keep the information secure while allowing the SPA to gather the data to perform its functions. Moreover, in agreement with Dousay and Hall (2018), Jean-Charles (2018) found out that most of the teachers did not feel prepared to use this technology in a meaningful way. Arend (2018) implemented an SPA for a specific learning environment. She implemented Apple's Siri during a literacy activity in a third-year college English class. She used a conversation analysis approach and found out that one of the big advantages of interacting with Siri was that Siri triggered students to make their thoughts explicit, which helped them to structure their thinking processes. Summing up, we can conclude that research already provides explorative insights into the potential of SPAs in learning environments, but there is still little focus on rigorously measuring the effect of SPAs on skill development.

Scaffolding Theory

Bruner et al.'s (1976) theory of scaffolding emerged around 1976 as a part of social constructivist theory, and was particularly influenced by the work of Russian psychologist Lev Vygotsky (1978). Vygotsky (1978) argued that we learn best in a social environment, where we construct meaning through interaction with others. His Zone of Proximal Development Theory, which states that we can learn more in the presence of a knowledgeable other person, became the basis for the Theory of Scaffolding (Vygotsky, 1978). Among other things, this theory states that when students strive to acquire new knowledge, they need individualized support falling within their individual zone of proximal development. As they advance and become more independent in their thinking, this support can gradually fade away. Saye and Brush (2002) distinguish between static and adaptive scaffolds. Static scaffolding is defined as remaining the same over time for all students (i.e. one may provide a list of instructions that helps users to perform a learning activity). Assistance is not adjusted for individual students. Alternatively, adaptive scaffolding requires a constant analysis of the student's progress, adjustment as needed, and a reduction of support over time (i.e. one can monitor the progress of the student and provide scaffolds when needed in the learning process, Molenaar, Roda, van Boxtel, & Sleegers, 2012). The main goal of an educator is to offer scaffolds, such as questions and hints, within students' individual zone of proximal development. Adaptive and interactive SPAs might be able to imitate educator's scaffolding dialogs to some extent resulting in a meaningful learning process and increased skill levels.

Technology-mediated Learning

Technology-mediated learning (TML) is defined as an environment in which the student's interactions with learning materials (readings, assignments, exercises, etc.), peers and/or instructors are mediated through advanced information technologies (Alavi & Leidner, 2001). Gupta and Bostrom (2009) propose a holistic, theoretical model that takes two different perspectives on TML. The *deterministic perspective* focuses on the structures and their influence on the learning process. Structures consist of the learning context structures and the learning method structures. One component of the learning method structures is information technology. The *process perspective* focuses on the learning process itself. The learning process is viewed as an appropriation of the structures during learning. How students use the information technology during learning determines their learning outcomes. Structures of a learning method are constituted recursively as students interact with them. In specific, the TML model proposes that individual differences and appropriation support of students are a decisive factor that explains the relationship between structures and learning outcome. The appropriation support and individual differences have a direct and indirect effect on learning outcomes. Scaffolding as appropriation support is usually designed by the instructor and gradually fades as students become more independent, confident, and competent. Past research has shown that scaffolding, as appropriation support, can help ensure a successful appropriation of learning methods (cite). The TML model argues that scaffolds facilitate the learning process via increasing the appropriation of the structures involved. We argue that SPAs as a new type of information system is able to offer adaptive scaffolds similar to human educators. This will lead to a better appropriation of the learning method resulting in higher levels of learning outcome.

ICAP Framework

The ICAP framework proposed by Chi and Wylie (2014) is based on a constructivist view of learning. It explains the process of effective learning by classifying observable student behaviors into four modes: **i**nteractive, **c**onstructive, **a**ctive, **p**assive. It predicts that these modes will be ordered by effectiveness as follows: *interactive* > *constructive* > *active* > *passive***.** Educators have long recognized that although students can learn from receiving information passively, they learn much better by learning actively (Bonwell and Eison, 1991). Learning actively requires students 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 (King, 1993). Each mode of the ICAP framework corresponds to different types of behaviors and knowledge-change processes predicting different learning outcomes. Appendix A outlines the different modes (from passive to interactive), showing that students show different modes of behavior for different learning scenarios (listening to a lecture, reading a text, and observing a video).

According to Chi and Wylie (2014), a *passive* student behavior describes students who deal with the presented instructional information without additional physical activity (e.g., reading a text without doing anything else). An *active* student behavior includes "doing something physically" (e.g., underlining or highlighting a text). A *constructive* student 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". Interactive 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 about a previously read text, Chi and Wylie, 2014). Until now, traditional computer tutoring systems were mostly able to bring students into a constructive learning behavior by asking them one-size-fits-all questions and giving them the opportunity to continue to the next question after answering.

New emerging information systems, namely SPAs, have the capabilities to interact with the students approaching the gold standard of a human educator by adapting their answers to students' utterances and also proactively guiding them through a task. Thus, we argue that SPAs can change students learning behavior from constructive to interactive resulting in increased skill development. Thus, we propose:

Hypothesis. The use of scaffolding-based Smart Personal Assistants will improve students' problem-solving skills.

RESEARCH METHODOLOGY

The objective of our study is to investigate whether scaffolding-based SPAs are able to increase students' problem-solving skills in a management education context. To answer our research questions, we used a mixed-method approach where we designed two field quasi-experiments with a pre- and post-test control group design and one post-experiment focus-group discussion. Mixed-method research can develop insights into new phenomena of interest that cannot be fully understood by using only quantitative or qualitative methods (Venkatesh, Brown, & Bala, 2013). We chose a mixed-method approach because we wanted to obtain complementary views on the quantitative relations between SPA usage and problem-solving skills reflecting

Venkatesh's first type of reasons to conduct a mixed-method approach (type 1 – complementary; Venkatesh et al., 2013).

Background and Setting

We implemented SPAs in a second-year business high school (October/November 2018) and a second-year vocational business school (January/February 2019) in a western European country. Each of the two experiments ran over a period of five weeks. The choice of implementing SPAs in these learning environments has two main benefits. First, in each school, the experiment class and control class were similar in terms of pre-knowledge, school grades, gender, and learning goals. Moreover, in each school, the two classes had the same teacher and the same learning content, making the two classes suitable for a field quasi-experiment approach. Second, choosing two different school types allowed us to widen the scope and strengthen our results. In business high school, school performances are better as students aim to prepare themselves for university. In the vocational business school, school performances are generally a bit lower and students tend to start to work after school. Both schools are located in the capital city of its region. The high school had approximately 1300 students and the vocational business school had 500 students. Each school lasts four years. The relevant subject in both schools was "business and law". The relevant teacher in both schools was very experienced, having taught for over 25 years in different types of schools (elementary school, high school, vocational school, etc.). They reported using technology with their students daily (computer-based teaching, tablet-based teaching, etc.). Both classes in both schools had 3 lessons a week, resulting in 15 lessons over the experiment period of five weeks. During the period of five weeks, both teachers in both schools dealt with the topic "introduction into law". Both classes in both schools had exactly the same six learning goals (LG). These learning goals are depicted in Appendix B. The teachers used the same methods in both classes, consisting of whole-group instruction, followed by individual and partner work. None of the four classes had had law-related subjects before.

Sample

Experiment 1 in Business High School

The sample in experiment 1 consisted of two second-year classes of a high school in a western European country. The experiment class had 9 males and 13 females, with an average age of 16.9 years. The control class had 11 males and 12 females, with an average age of 17.1 years. We conducted ANOVA tests to make sure that the two groups are similar. The tests revealed that there was no significant difference in the background of the students in terms of their gender ($p = 0.650$), school grades in business and law ($p = 0.558$), pre-experience with SPAs ($p = 0.941$), personal innovativeness ($p = 0.191$) and the pre-test results ($p = 0.774$). No one from this class participated in the post-experiment focus group discussion.

Experiment 2 in Vocational Business School

The sample in experiment 2 consisted of two second-year classes of a vocational business school in European country. The average age of the experiment class was 17.4 years, with 12 males and 10 females. The average age of the control class was 17.2 years, with 12 males and 11 females. The vocational students had apprenticeships with mostly large-sized banks and insurances (n=18), followed by large merchandise trade companies (n=10), small and medium-sized companies (n=10) and others (n=7). We conducted ANOVA tests to make sure that the two classes were similar. The tests revealed that there was no significant difference in the background of the students in the two classes in terms of their gender ($p = 0.248$), school grades in business and law ($p = 0.1000$), pre-experience with SPAs ($p = 0.552$), personal innovativeness ($p = 0.332$) and the pre-test results (p = 0.376). The experiment class of this school participated in the post-experiment focus group discussion.

Task Design

Between the end of week 1 and the end of week 4, the students had to do four problem-based, 30-minute homework assignments for learning goals (LGs) 3, 4 and 5. The students from the experiment class used the dynamic, scaffolding-based SPA as individual support for the homework assignments. All the homework assignments were in the same style as the pre- and post-test tasks and fulfilled the requirements of a problem-based task according to Jonassen (2000). Accordingly, problem tasks should be ill-structured, open-ended, realistic, and resonate with the executors' experience. As an example, in Appendix C, we show homework assignment 2 and how it addresses these requirements.

Design and Functionality of the scaffolding-based Smart Personal Assistant

The goal of our study was to develop an SPA that is able to interact with the students during their homework assignments in the most natural way possible. Based on the ICAP-framework, building up a natural dialogue with students is a necessary condition to trigger their most effective interactive learning behavior. Therefore, we used Amazon's Alexa SPA ecosystem (Amazon, 2019). Specifically, we used Alexa's Skill Development Kit 2.0 with nodeJS because this framework seems to offer one of the most developed state-of-the-art capabilities regarding speech recognition and natural language processing. Following the principles of dynamic scaffolding, we included two different logics in our interaction model: proactive and reactive. In the proactive logic, the SPA tutor proactively guided students through the task by using five different problem-solving activities adapted from Kim and Hannafin's (2011) problem-solving phases: *(a) problem* *identification and engagement, (b) problem exploration, (c) problem reconstruction, (d) solution presentation and communication,* and *(e) reflection and negotiation.* The SPA tutor provided its problem-solving steps in an interval of five minutes. In the reactive second logic, the SPA provided feedback on students' utterances and reacted to students' intents between the different problem-solving steps. Specifically, the SPA was able to react to three different types of intents. Type 1 was the students' intent to skip to the next step, Type 2 was the students' intent to ask content-specific questions (concept clarifications, asking for hints, etc.), and Type 3 was to ask the tutor to rephrase the current problem-solving step. The proactive and reactive logic helps students to interact with the SPA, thereby internalizing the problem-solving steps, which should result in a gain of their problem-solving skills. As an example, Figure 1 shows an excerpt of a student dialogue and the corresponding proactive and reactive functions.

INSERT FIGURE 1 ABOUT HERE

Before introducing the SPA in the two experiment classes, we tested the SPA several times within the research team as well as with students. Moreover, we asked the two teachers to review our interaction model and the corresponding interaction dialogues. A key challenge when designing SPA materials for experimental treatments is the issue of informational equivalence. According to Larkind and Simon (1987), two representations are informationally equivalent if all the information from one representation can also be inferred from the other representation and vice versa. This ensures that differences in effect stem from the IT artifact itself and not from the content of a representation. Therefore, we included all the information incorporated in the SPA also on the scaffolds provided for the control class. Students in the control class received static scaffolds on a soft copy (white pdf document). Figure 2 shows the same excerpt as in Figure 1.

INSERT FIGURE 2 ABOUT HERE

Experimental Procedure

Figure 3 depicts the study timeline for experiments 1 and 2.

INSERT FIGURE 3 ABOUT HERE

Several tasks were conducted before the experiment started, i.e., during the pre-experiment state. During a lesson, all students completed a survey to help identify and control for pre-existing class differences (pre-experience with SPAs, personal innovativeness, gender, age; see Appendix D). Also, the classroom teachers administered a 30-minute, 3-subtask pre-knowledge test to all students. The goal of the pre-knowledge test was to identify students' problem-solving skills in the relevant subject. Moreover, the research team conducted a pre-experiment meeting with the teachers to discuss the experiment details with them and to ensure that learning goals, teaching methods and tools stayed the same in the experiment and the control classes over the period of the experiment.

At the beginning of week 1, we introduced the SPA devices (Amazon's Alexa Echo Dot Device) in the experiment classes and helped students to install the accompanying Alexa software on their smartphones, tablets, and laptops. At the end of week one, we ensured that students in the experiment classes had access to Alexa on one or more devices. After week 1, the students received their four homework assignments in the form of a paper script (non-digital). The experiment classes were instructed to use Alexa on their preferred devices (standalone device, smartphone, tablet and laptop) while conducting their homework assignments. The control classes were instructed to use the static scaffolds that were sent to them in form of a pdf document (with exactly the same amount of information as on Alexa). At the end of week 4, all students submitted their homework assignments. One week later, they took their 30-minute post-knowledge test and post-experiment survey. The teacher was instructed to not discuss the homework assignments before the post-knowledge test was conducted to ensure that the final results were not corrupted. In the post-experiment phase, the experiment class of the vocational business school participated in a 45-minute focus group discussion.

Measurement and Analysis

Quantitative Data

For measuring problem-solving skills (dependent variable), we constructed a 3-task pre- and post-experiment test with the help of the teachers of the class. The tasks within the pre-test and post-test had the same structure as the homework assignments, all addressing the requirements of problem-based tasks proposed by Jonassen (2000). The pre-test and post-test had the same total number of points and the tasks addressed learning goals 3, 4 and 5 (see Appendix E). All tasks require the application of Kim and Hannafin's (2011) problem-solving steps. Two teachers compared the pre-test and the post-test in terms of difficulty and found them to be similar. We analyzed pre-test and post-test results with three experienced raters independently and blinded with a pre-defined and commonly discussed rating framework. One of the raters was the class teacher. The rating framework allowed us to evaluate the application of the problem-solving steps proposed by Kim and Hannafin (2011), giving points to each well-applied problem-solving step within a task (see Appendix F). The final scores of the pre- and post-test results arose from an average of the individual appraisals. Moreover, we calculated gain scores as the difference between a students' posttest and pretest scores. We checked for inter-rater agreement with the help of Pearson's correlation for both classes in both schools (interrater agreement = 0.9233, $p < 0.05$). The pre-experiment survey consisted of items regarding the students' pre-experience with SPA, personal innovativeness, apprenticeship company, gender, and age. For pre-experience with SPAs, we asked students how often they use SPAs (e.g., Apple's Siri) in a week. For personal innovativeness, we used the four items from van Raaij and Schepers' (2008) scale. The post-survey included two open-ended questions. Question 1 was about the students' perceived helpfulness of the learning aid (SPA or paper-based learning materials). Question 2 was treatment-specific, asking the experiment classes about their experiences with the SPAs used. The items of the post-survey are depicted in Appendix G.

To analyze the quantitative data, we ran a one-way analysis of covariance (ANCOVA), including the pre-test results as a covariate. ANCOVA helps to analyze variances between the groups and control for covariates. It is therefore suitable for pre- and post-test designs (Dimitrov & Rumrill, 2003). We calculated Cohen's d (1988) to show the effect size. Moreover, we calculated 95% confidence intervals and used the statistic program R as a tool for analysis (Team, 2013).

Qualitative Data

The focus group discussion was based on the answers of the open-ended questions from the post-survey and addressed RQ2. The objective of the focus group discussion was to gain a more in-depth understanding of how SPAs affect students' learning processes. The discussion lasted 45 minutes and had one of the researchers as the facilitator. All students from the experiment class in the vocational business school participated. The focus group discussion was structured as follows. First, the facilitator introduced the goal of the group discussion (to gather students' perceptions of using SPAs while learning). Second, students were asked to divide into workgroups of four to six people and were invited to discuss and negotiate opinions about how they used the SPAs during their homework assignments. To help students structure their discussion, they received three broad subject areas identified from the post-survey responses: pros of using SPAs, cons of using SPAs, and neutral observations while interacting with SPAs. Finally, a plenary discussion moderated by the facilitator encouraged further discussion and the gathering of additional content. We recorded the session, transcribed it and used a thematic analysis to induce topics following the method of Ryan & Bernard (2003). Specifically, we used the keywords-in-context method for this study. With the help of this technique, we identified keywords indicating some aspects of the learning process with SPAs and then systematically searched the corpus of the transcribed text to find all instances of the word or phrase. Each time we came across an instance of the word or phrase, we made a note of it and its immediate context. We identified themes by physically sorting the instances into piles of similar meaning (Ryan & Bernard, 2003). Then, we conducted a respondent validation by getting participants to verify our identified themes.

RESULTS

First, we present the results related to our first research question: To what extent does using a Smart Personal Assistant help students develop their problem-solving skills?

Experiment 1 in Business High School

Table 1 shows a summary of the means of the pre-test and post-test results, including the standard deviations and gain scores.

INSERT TABLE 1 ABOUT HERE

To check that the assumptions for the ANCOVA model are met, we conducted a graphical test for normality, homogeneity of variance, and homogeneity of regression slopes (see Appendix G). The plots indicate that all assumptions for conducting an ANCOVA are met. We ran the ANCOVA with post-test scores as the dependent, the treatment group as the independent variable and the pre-test score as a covariate. Results of the test indicate that, controlling for the pre-test, there is a significant relation between SPA usage and problem-solving skills ($F (2, 42) = 4.514$, r^2 adjusted = 0.5255, $p = 0.0396$, $N = 45$). Cohen's d is 0.5178, indicating a difference between the means of 0.5178 standard deviations. Since this effect is considered as medium (Cohen, 1988), we can conclude that the use of SPAs has a positive effect on acquiring problem-solving skills compared to soft-copy based learning materials. Moreover, a comparison of the gain scores reveals that participants in the SPA group learned significantly more than participants in the paper-based learning materials group ($p = 0.0400$).

Experiment 2 in Vocational Business School

Table 2 shows a summary of the means of the pre-test and post-test results, including the standard deviations and gain scores.

INSERT TABLE 2 ABOUT HERE

The plots given in Appendix H indicate that all assumptions for conducting an ANCOVA are met. We ran the ANCOVA with the post-test scores as the dependent, the treatment group as the independent variable and the pre-test score as a covariate. Results of the ANCOVA indicate that, controlling for the pre-test, there is a highly significant relation between SPA usage and problem-solving skills ($F (2, 42) = 30.573$, r^2 adjusted = 0.42, $p = 1.88e-06$, $N = 45$). Cohen's d is 1.70, indicating a difference between the means of 1.70 standard deviations. Since this effect is considered as high (Cohen, 1988), we can conclude, that the use of SPAs has a positive effect on acquiring problem-solving skills compared to the use of soft-copy based learning materials. Moreover, a comparison of the gain scores reveals that participants from the SPA group learned significantly more than participants from the paper-based learning materials group ($p = 8.17e-06$). More detailed R Statistics for both experiments are depicted in Appendix I.

Focus Group Discussion

Based on our quantitative findings regarding the relationship of SPA usage and problem-solving skills, we investigate our second research question on *how SPAs affect students' learning processes* to gather deeper insights into the shown effect. The three main themes we identified from the focus group data are *interaction, usage behavior,* and *individualization.* They are presented next and further elaborated in section 5 (Discussion).

Interaction. This theme relates to SPAs being able to build up a dialogue with the students. Several students mentioned that using the SPA while completing their assignments felt like having an interaction with a peer or a tutor. Students appreciated that the SPA listened to them and adapted its answers accordingly. Some students mentioned that receiving challenging questions and hints from the SPA helped them to think of the next solution steps. Some other students perceived it as more entertaining compared to "business as usual" paper-based learning materials. For example, one student commented: "I liked her [Alexa]. It felt like I was talking to a teacher. She responded immediately and also asked me challenging questions that helped me with the next steps."

Moreover, several students mentioned that they liked the way how the SPA was helping them. Specifically, they said that they appreciated thinking on their own about the solution first and that they can control when they want to receive help. For example, one student commented: “It was nice that Alexa was waiting until I asked her for help. That helped me to first think on my own and only receive hints when I want.” Moreover, some students also mentioned that they liked saying their solutions out loud. While hearing themselves talk, they came up with new ideas more easily. For example, one student commented: “I liked speaking the answers out loud and not writing them down. When I hear myself, I get new ideas.”

Usage Behavior. This theme relates to the different ways students used the SPAs compared to paper-based learning materials. Most of the students used the SPA on their smartphone rather than their standalone device and mentioned that it was really easy and fun to access Alexa similarly to Apple’s Siri or Google’s Assistant. For example, one student commented: “It was like speaking with Siri. Just like having your personal tutor always in your pocket anytime and anywhere you want.” The easy access to the learning tutor on smartphones and also other devices (e.g., Amazon’s standalone device Amazon Echo Dot) led to a great variety of different learning locations. Some students indicated that they did their tasks in other places than usual. For example, one student commented: “When I was lying on my couch, I talked to Alexa a few times, too.” However, a few students also mentioned some areas where SPAs were not very functional. For example, when students were in public, they seldomly used Alexa because they felt uncomfortable talking to an SPA in public. One student commented: “First, I wanted to conduct the tasks in the train. But then I decided to do something else; speaking with my smartphone in front of others felt weird to me”.

Individualization. The third theme we identified from the focus group data relates to the capacitity of SPAs to recognize students’ individual characteristics. Some students mentioned that they liked that Alexa was able to adapt her answers to their utterances and prompted them individually until they were able to find their own solution. For example, one student commented: “It was helpful that Alexa recognized how she can help me discover the solution.” Moreover, some students mentioned that they liked that the SPA remembered the current status of progress whenever students interrupted the task. However, some other students wished that the SPA could recognize where students have their biggest weaknesses. For example, one student commented: “It would be great if she [Alexa] can remember our mistakes and then concentrate on helping us with that.”

DISCUSSION

Using a mixed-method approach, this study investigates the use of new emerging SPAs as a new learning technology in everyday learning environments to increase students’ problem-solving skills, which is considered as a number-one skill in management education (Ungaretti, Thompson, Miller, & Peterson, 2015). Our two field quasi-experiments are driven by the hypothesis that students achieve better problem-solving skills using SPAs than using soft-copy based learning materials. The findings of both experiments provide strong support for our hypothesis and also offer insights into how SPAs are able to change students’ learning processes. Students who had used SPAs performed significantly better on the post-experiment knowledge test than students who had used soft-copy based learning materials.

This is in accordance with prior research which showed that intelligent tutoring systems can support problem-solving activities (Hooshyar, Ahmad, Yousefi, Fathi, Horng, & Lim, 2016; Wang, Han, Zhan, Xu, Liu, & Ren, 2015). Our results indicate that SPAs might be able to trigger students’ interactive behavior. According to the ICAP-framework (Chi & Wylie, 2014), students’ interactive learning behavior can be considered as the gold standard leading to better learning outcomes. This is also reflected by our qualitative findings indicating that students perceive the interaction with

SPAs as similar to human tutors (theme 1 form the focus group discussion). Interestingly, the students mentioned that they liked how the SPA fostered their understanding until they found the answer on their own. This is a big difference to "business-as-usual" paper-based learning materials, where there is no interaction possible. It seems that they enjoyed having their personal tutor on their side and knowing that they can receive individual support when they cannot get any further on their own. The SPA then supported them in constructing knowledge on their own. This helped students to better internalize the learning materials and gain their problem-solving skills.

What also contributes to the positive relationship between using SPAs and students' problem-solving skills might be that students can change their learning behavior (theme 2). In particular, students can change their learning places and learning times. Similar results can be confirmed by Taylor (2006), who stated that using smartphones as learning assistants changed learning contexts, for instance in terms of ergonomics (user posture, lighting, background noise), social context and demands of users' attention. The different learning places and times might lead to higher amounts of learning time and, finally, better skill-development. On the other hand, students also mentioned that they felt weird using SPAs in public, which is also confirmed by findings from Moorthy and Vu (2015). Their results showed that participants preferred using SPAs in a private location (e.g., their home) due to social acceptability. One other reason for the superiority of the SPA compared to paper-based learning materials might be that SPAs are able to provide individual support. In specific, students highlighted that the SPAs allowed them to learn at their own pace, receiving help whenever they wanted it. Learning at their own pace motivates students and gives them the feeling of working on their own academic progress (Chen, 2008). These effects of receiving help whenever they want can also be confirmed by several other research papers in the area of personalized learning (Ammar, Neji, Alimi, & Gouardères, 2010; Song, Wong, & Looi, 2012). Students did however mentioned that they wished that SPAs could detect individual students' characteristics and in particular track their weaknesses. This confirms the research effort in the area of ITS research trying to capture students' characteristics and adapt ITS accordingly. For example, Ammar et al. (2010) designed an ITS equipped with emotional management capabilities, which makes the capture of student's emotions possible during learning and responds accordingly.

Our work contributes to two different research areas education. First, we contribute to technology-mediated learning research by providing much-needed empirical evidence on the effect of scaffolding-based SPAs on students' long-term skill development compared to soft-copy based learning materials, especially in the field of management education. To the best of our knowledge, empirical evidence in this area is still missing. Moreover, as SPA providers offer intuitive interfaces to build skills, we showed that SPAs are a promising way for future implementations by bridging the gap between the technological know-how of software designers and the pedagogical/didactical know-how of educators. Second, we contribute to the ICAP Framework by showing that a new kind of information technology can bring students into a more interactive learning mode, resulting in increased levels of skill development. In regard to its practical implications, this study exemplarily showed educators how to build and integrate SPAs in an existing "business as usual" learning environment.

LIMITATIONS AND FUTURE RESEARCH

There are a number of limitations to this study that should be noted. First, we used a field quasi-experiment design to examine the impact of SPAs on students' problem-solving skills. With this treatment design, pre-treatment group differences between the experimental group and the control group may confound post-treatment outcomes. We tried to address this point by collecting pre-experiment data and using ANOVAs to check if experimental and control groups are similar.

Moreover, both classes in each school had the same teacher with the same learning goals and using the same teaching methods. Nevertheless, it would be interesting to see if future experimental research is able to confirm and extend our results. Second, the sample size of our two field quasi-experiments (n=45, 2 classes per school) can be considered as rather small for using an ANCOVA. Though, our statistics showed a medium effect (in the business high school) and a large effect (in the vocational business school) between these two groups, which indicates that there is a relationship between SPA usage and the increase of problem-solving skills. Furthermore, we applied a mixed-method approach to partially compensate for the rather small sample size (Venkatesh et al., 2013). It would be interesting to see if studies with larger sample sizes are able to confirm our effect. Third, we conducted the SPA experiment over a relatively brief period under positive conditions. Such studies tend to produce better outcomes due to novelty effects and hyperattention to experiment details (Cheung & Slavin, 2013). For the current study, novelty effects may be small, given the large percentage of participants reporting the high usage of SPAs on their smartphones (approximately 40% used SPAs every day on their smartphones; e.g., Siri). Finally, this study also raises some ethical questions about data security and the potential benefit of one group and not the other. Students put their information on Amazon's ecosystem, which might be a threat to their personal security. While it is hard to control student behavior, we educated students about the experiment and the corresponding risks at the beginning of the study. For future research, it would be interesting to see how SPAs can be implemented in-class to improve students' skill development. Moreover, future research should focus on individual differences in students' characteristics that might influence learning processes with SPAs.

CONCLUSION

Our study answered the question whether the use of SPAs could increase students' long-term problem-solving skills in a management education context as well as identified students' perceptions about the use of SPAs. We designed and introduced SPAs over a period of five weeks and evaluated students' long-term problem-solving skills with the help of a pre- and post-test control group design. One group used paper-based learning materials, whereas the other group used SPAs instead. The findings of this study provide strong support for the proposition that SPAs have the potential to be an effective tool for improving long-term skill development in managemend education. These findings contribute to both technology-mediated learning and education research. While further research is required, it appears that education is a promising area within SPAs can improve long-term skill development.

References

Afzal, S., Dhamecha, T., Mukhi, N., Sindhgatta, R., Marvaniya, S., Ventura, M., & Yarbro, J. 2019. Development and Deployment of a Large-Scale Dialog-based Intelligent Tutoring System. In Human Language Technologies (Ed.), ***Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics***, vol. 2: 114–121.

Alavi, M., & Leidner, D. E. 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. ***MIS quarterly***, 25(1): 107.

Al-Bastami, B. G., & Naser, S. S. A. 2017. Design and Development of an Intelligent Tutoring System for C# Language, 4(10): 8795–8809.

Amazon. 2019. ***New Amazon Alexa Skill Kit 2.0***. Retrieved from https://developer.amazon.com/blogs/alexa/post/1a4e8b01-663d-4680-8efd-c28e96e31655/now-available-version-2-of-the-ask-software-development-kit-for-java. Accessed 2018, November 20.

Ammar, M. B., Neji, M., Alimi, A. M., & Gouardères, G. 2010. The affective tutoring system. ***Expert Systems with Applications***, 37(4): 3013–3023.

Arend, B. 2018. Investigating Siri as a virtual assistant in a learning context, ***Proceedings of 12th annual International Technology, Education and Development Conference***: 1–10.

Baker, R. S. 2016. Stupid tutoring systems, intelligent humans. ***International Journal of Artificial Intelligence in Education***, 26(2): 600–614.

Barbhuiya, R., Karim, K. M., & Suraiya, J. 2011. Design specifications for a generic Intelligent Tutoring System, ***Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE)***. Las Vegas, USA: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Bartlett, J. E. 2002. Analysis of Motivational Orientation and Learning Strategies of High School Business Students. ***56***, 4: 18–23.

Brush, T. A., & Saye, J. W. 2002. A summary of research exploring hard and soft scaffolding for teachers and students using a multimedia supported learning environment. ***The Journal of Interactive Online Learning***, 1(2): 1–12.

Chen, C.-M. 2008. Intelligent web-based learning system with personalized learning path guidance. ***Computers & Education***, 51(2): 787–814.

Cheung, A. C. K., & Slavin, R. E. 2013. The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. ***Educational Research Review***, 9(1): 88–113.

Chi, M. T. H., & Wylie, R. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. ***Educational Psychologist***, 49(4): 219–243.

Chung, H., Iorga, M., Voas, J., & Lee, S. 2017. Alexa, Can I Trust You? ***Computer***, 50(9): 100–104.

Cohen, J. 1988. ***Statistical power analysis for the behavioral sciences 2nd ed.:*** Erlbaum Associates, Hillsdale.

David, H., Katz, L. F., & Kearney, M. S. 2006. The polarization of the US labor market. ***American economic review***, 96(2): 189–194.

Dimitrov, D. M., & Rumrill, P. D. 2003. Pretest-posttest designs and measurement of change. ***Work***, 20(2): 159–165.

Dousay, T. A., & Hall, C. 2018. Alexa, tell me about using a virtual assistant in the classroom, ***Proceedings of EdMedia: World Conference on Educational Media and Technology***. Amsterdam, Netherlands: Association for the Advancement of Computing in Education (AACE).

Easwara Moorthy, A., & Vu, K.-P. L. 2015. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. ***International Journal of Human-Computer Interaction***, 31(4): 307–335.

Goksel-Canbek, N., & Mutlu, M. E. 2016. On the track of Artificial Intelligence: Learning with Intelligent Personal Assistants. ***International Journal of Human Sciences***, 13(1): 592.

Graesser, A. C., Hu, X., & Sottilare, R. 2018. Intelligent tutoring systems, ***International handbook of the learning sciences***: 246–255: Routledge.

Gupta, S., & Bostrom, R. P. 2009. Technology-mediated learning: A comprehensive theoretical model. ***Journal of the Association for Information Systems***, 10(9): 686.

Hauswald, J., Laurenzano, M. A., Zhang, Y., Li, C., Rovinski, A., Khurana, A., Dreslinski, R. G., Mudge, T., Petrucci, V., Tang, L., & Mars, J. 2015. Sirius: An open end-to-end voice and vision personal assistant and its implications for future warehouse scale computers. In ACM SIGPLAN Notices (Ed.), ***ACM SIGPLAN Notices***: 223–228 (50th ed.): ACM.

Hilles, M. M., & Naser, S. S. A. 2017. Knowledge-based Intelligent Tutoring System for Teaching Mongo Database, 4(10): 8783–8794.

Hooshyar, D., Ahmad, R. B., Yousefi, M., Fathi, M., Horng, S.-J., & Lim, H. 2016. Applying an online game-based formative assessment in a flowchart-based intelligent tutoring system for improving problem-solving skills. ***Computers & Education***, 94: 18–36.

Jean-Charles, A. 2018. Internet of Things in Education: Artificial Intelligence Voice Assistant in the Classroom. In Proceedings of Society for Information Technology & Teacher Education International Conference (Ed.), ***Proceedings of Society for Information Technology & Teacher Education International Conference***: Association for the Advancement of Computing in Education (AACE).

Jonassen, D. H. 2000. Toward a design theory of problem solving. ***Educational Technology Research and Development***, 48(4): 63–85.

Kim, M. C., & Hannafin, M. J. 2011. Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. ***Computers & Education***, 56(2): 403–417.

Kulik, J. A., & Fletcher, J. D. 2016. Effectiveness of Intelligent Tutoring Systems. ***Review of educational research***, 86(1): 42–78.

Larkin, J., Simon, H. 1987. Why a Diagram is (Sometimes) Worth Ten Thousand Words. ***Cognitive science***, 11(1): 65–100.

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. 1992. Effective Tutoring Techniques: A Comparison of Human Tutors and Intelligent Tutoring Systems. ***Journal of the Learning Sciences***, 2(3): 277–305.

Molenaar, I., Roda, C., van Boxtel, C., & Sleegers, P. 2012. Dynamic scaffolding of socially regulated learning in a computer-based learning environment. ***Computers & Education***, 59(2): 515–523.

OECD. 2014. ***Creative problem solving***: ***Students' skills in tackling real-life problems***. Paris: OECD.

Oeste, S., Lehmann, K., Janson, A., Söllner, M., & Leimeister, J. M. 2015. Redesigning University Large Scale Lectures: How To Activate The Learner, ***75th Academy of Management Annual Meeting***. Vancouver, British Columbia.

Rietsche, R., Duss, K., Persch, J. M., & Soellner, M. 2018. Design and Evaluation of an IT-based Formative Feedback Tool to Foster Student Performance, ***39th International Conference on Information Systems (ICIS)***: 1–17. San Francisco, CA, USA.

Ryan, G. W., & Bernard, H. R. 2003. Techniques to identify themes. ***Field methods***, 15(1): 85–109.

Soderstrom, N. C., & Bjork, R. A. 2015. Learning versus performance: An integrative review. ***Perspectives on Psychological Science***, 10(2): 176–199.

Song, Y., Wong, L.-H., & Looi, C.-K. 2012. Fostering personalized learning in science inquiry supported by mobile technologies. ***Educational Technology Research and Development***, 60(4): 679–701.

Spitz-Oener, A. 2006. Technical change, job tasks, and rising educational demands: Looking outside the wage structure. ***Journal of labor economics***, 24(2): 235–270.

Statista. 2019a. ***US Voice Technology Usage by Age***. Retrieved from https://www.statista.com/statistics/879907/us-voice-technology-usage-age/. Accessed 2019, May 1.

Statista. 2019b. ***Virtual Digital Assistant Market Worldwide***. Retrieved from https://www.statista.com/statistics/589079/worldwide-virtual-digital-assistants-consumer-market/. Accessed 2019, May 1.

Taylor, J. 2006. Evaluating mobile learning: What are appropriate methods for evaluating learning in mobile enviroments. ***Big issues in mobile learning***, 1(3): 25–27.

Team, R. C. 2013. ***R: A language and environment for statistical computing***. Retrieved from http://www.R-project.org/. Accessed 2019, May 1.

Ungaretti, T., Thompson, K. R., Miller, A., & Peterson, T. O. 2015. Problem-Based Learning: Lessons From Medical Education and Challenges for Management Education. ***Academy of Management Learning & Education***, 14(2): 173–186.

van Raaij, E. M., & Schepers, J. J. L. 2008. The acceptance and use of a virtual learning environment in China. ***Computers & Education***, 50(3): 838–852.

Venkatesh, V., Brown, S. A., & Bala, H. 2013. Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. ***MIS quarterly***: 21–54.

Vygotsky, L. S. 1978. ***Mind in society***: ***The development of higher mental process:*** Cambridge, MA: Harvard University Press.

Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., & Ren, G. 2015. A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. ***Computers & Education***, 81: 102–112.

Wood, D., Bruner, J. S., & Ross, G. 1976. The role of tutoring in problem solving. ***Journal of child psychology and psychiatry***, 17(2): 89–100.

ATTACHMENTA

FIGURE 1

Excerpt of an exemplary dialogue between the student and SPA and corresponding proactive and reactive logic.

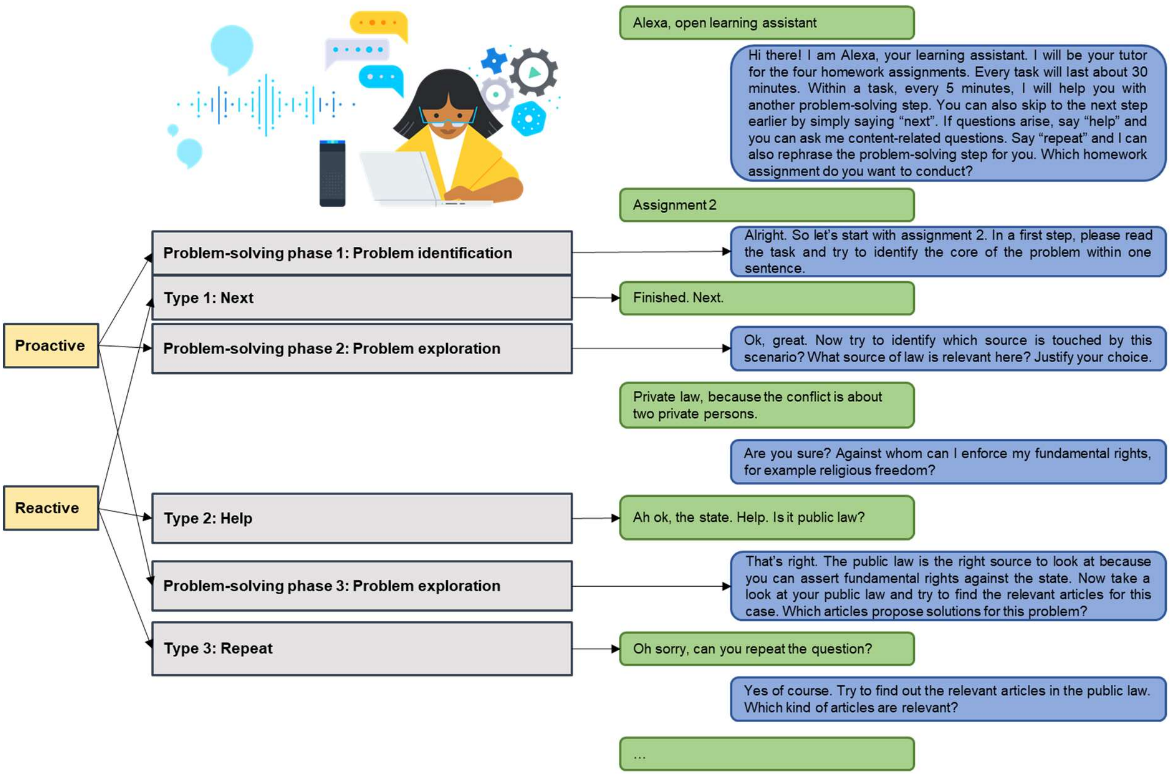


FIGURE 2

Excerpt of static scaffolds for homework assignment 2 provided for control group.

Assignment 2:

Step 1: Let’s start with assignment 2. In a first step, please read the task and try to identify the core of the problem within one sentence.

Step 2: Try to identify which source is touched by this scenario? What source of law is relevant here? Justify your choice.

Step 3: Take a look at your public law and try to find the relevant articles for this case. Which articles propose solutions for this problem?

…

Hints:

….

Step 3: The public law is the right source to look at because you can assert fundamental rights against the state.

FIGURE 3

SPA experiment implementation timeline

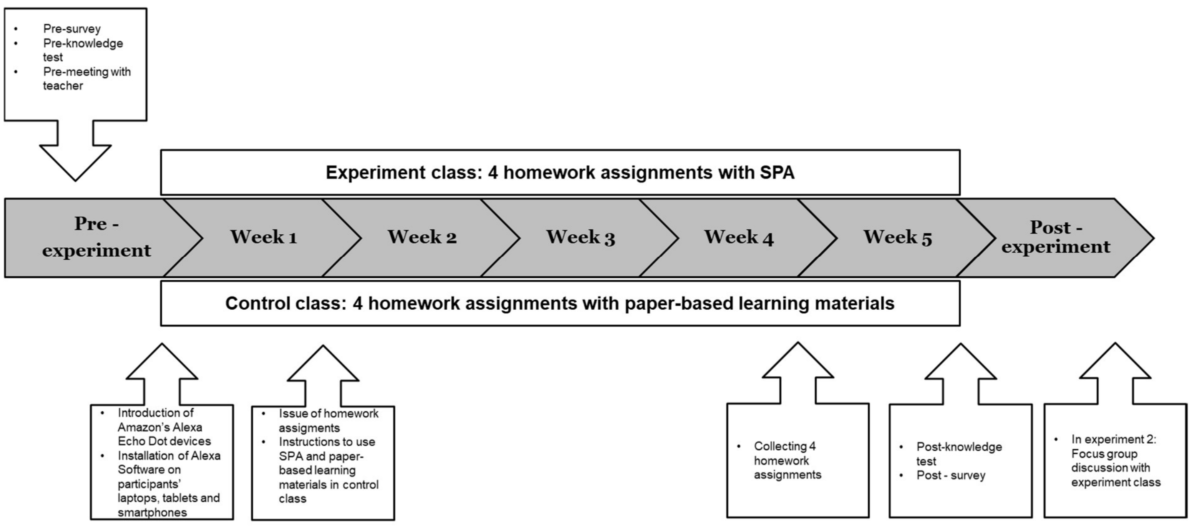


TABLE 1

Summary of Means

	Experiment Class (SPA)			Control Class (paper-based learning materials)		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Pre-test	10.46	1.67	23	10.27	2.44	22

Post-test	23.15	4.33	23	20.63	5.60	22
Gain scores	12.69	3.74	23	10.36	3.51	22

TABLE 2

Summary of Means

	Experiment Class (SPA)			Control Class (paper-based learning materials)		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Pre-test	11.7	1.63	22	10.67	2.16	23
Post-test	27.95	3.27	22	21.88	3.67	23
Gain scores	16.25	3.50	22	11.21	3.74	23

ATTACHMENT B

Appendix A. Distinction of student behavior according to ICAP framework **adapted from Chi and Wylie (2014).**

	PASSIVE (Receiving)	**ACTIVE (Manipulating)**	**CONSTRUCTIVE (Generating)**	**INTERACTIVE (Dialoguing)**
Listening to a lecture	Listening without doing anything else but oriented toward instruction	Repeating or rehearsing; Copying solution steps; Taking verbatim notes	Reflecting out loud; Drawing concept maps; Asking questions	Defending and arguing a position in dyads or small group
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
Observing a video	Watching the video without doing anything else	Manipulating the tape by pausing, playing, fast-forward, rewind	Explaining concepts in the video; Comparing and contrasting to prior knowledge or other materials	Debating with a peer about the justifications; Discussing similarities & differences

Appendix B: Lessons and learning goals

Lesson (15 lessons in total per class)	Learning Goals
Lesson 1	LG1: Students should understand the necessity of law in everyday life.
Lesson 2 to 4	LG2: Students should explain differences between morality, custom and law.
Lesson 5 to 7	LG3: Students should solve problems related to the freedom of opinion.
Lesson 8 to 10	LG4: Students should solve problems related to the freedom of religion.
Lesson 11 to 13	LG5: Students should solve problems related to the property guarantee.
Lesson 14 to 15	LG6: Students should analyze in which cases fundamental right have their limitations.

Appendix C: Homework assignment 2 and charateristics of a problem-based task

Homework Assignment 2	
Task: Imagine the following scenario: The use of smartphones during the lesson is forbidden in your school. Nevertheless, Thomas K. (a classmate of yours) uses his smartphone during the lesson to text his mother that he will be late that day. The teacher collects the smartphone and tells Thomas K. that she will keep the smartphone until the end of the week. How would you solve this emerging problem with the help of the law?	
Characteristics of a problem-task (Jonassen, 2000)	**How we addressed them**
Ill-structured	The task can be considered as ill-structured, because not every piece of information is given. Moreover, it gives no advice on how to solve the task.
Open-ended	The task leaves room for interpretation. There is not a single right answer. The students have to interpret the legal articles in a correct way.
Realistic	The task relates to a topic that is currently highly discussed in the relevant country. It can therefore be considered as realistic.
Resonate with the executors' experience	Since every participant owns a smartphone and has experienced similar situations in their own class, they can put themselves in Thomas K.'s position.

Appendix D: Pre-survey items

Pre-survey		
Variable	**Item**	**Scale**

Pre-experience with SPAs	1. Have you ever used a Smart Personal Assistant (e.g. Amazon's Alexa, Google's Assistant, Apple's Siri? 2. If yes, how often do you use a Smart Personal Assistant per week on average?	Yes/No open
Personal innovativeness	1. If I heard about a new information technology, I would look for ways to experiment with it. 2. Among my peers, I am usually the first to try out new information technologies. 3. In general, I am hesitant to try out new information technologies (reverse-scored). 4. I like to experiment with new information technologies.	1 to 7 (7 = highest) 1 to 7 1 to 7 1 to 7
Demographics	1. Age 2. Gender 3. Apprenticeship company (for vocational business school only)	open open open

Appendix E: Pretest- and post-test tasks

Learning goal	Pre-test	Post-test
LG3: solve problems related to the freedom of opinion	Imagine that in your country, political parties issue posters that are obviously against foreigners. Some foreigners complain against that. How would you solve this problem?	Imagine that in your country, Michael P, a good friend of yours, issues flyers that are obviously against one of the left-wing politicians in the country. Some people complain against that. How would you solve this problem?
LG4: solve problems related to the freedom of religion	Imagine that in your country, women are not allowed to wear burqas in public. Adem and Merve, two Islamic women, do not care about this rule and go out with their burqas. They are caught by the police and have to pay a fee now. They are complaining about it. How would you solve this problem?	Imagine that Muslims in your country have built a mosque in your neighborhood. After a while, a neighbor gets excited and criticizes the construction of the mosque. How would you solve the problem?
LG5: solve problems related to property guarantee	Imagine the state you are living in wants to build a road on your land. If you do not allow this, your land will be expropriated. You are filling a complaint against the expropriation of your land. How would you solve this problem?	Imagine that the city you are living in want to create a one-week food festival and, thus, needs two of your land plots, against which you are filing a complaint. How would you solve this problem?

Appendix F: General rating framework per task

Problem-solving step	Key Questions (1 Point per question)	Points
1. Problem identification and engagement	What is the main problem that is touched here? Who are the involved parties?	2
2. Problem exploration	Which area of the law is affected? What are the interests of the individual parties?	2

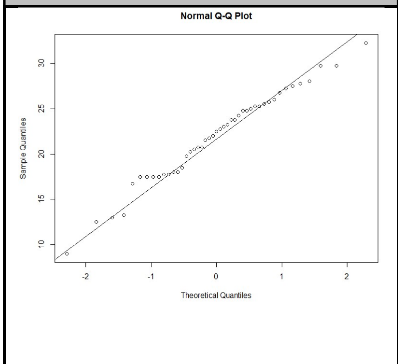
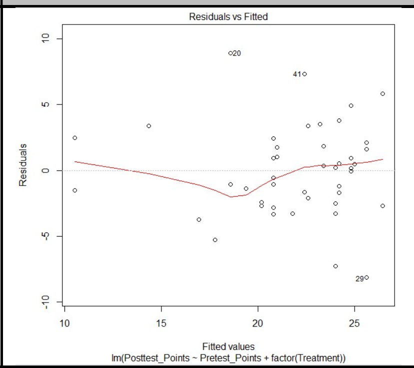
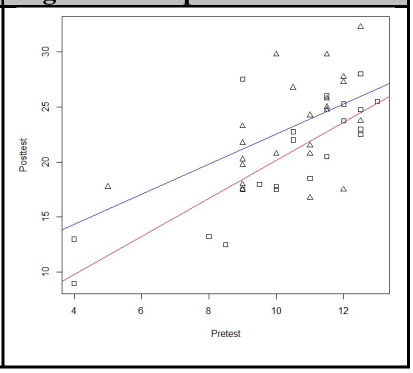
3. Problem reconstruction	What is a possible solution for the problem? Which article could apply here? What does the given law article say?	3
4. Solution presentation and communication	How can the law article be applied in the context? Are other articles also applicable? Who can raise which claims?	3
5. Reflection and Negotiation	How can you justify the solution? How can the related fundamental rights be restricted?	2
Total		**12**

Appendix G: Post-survey items

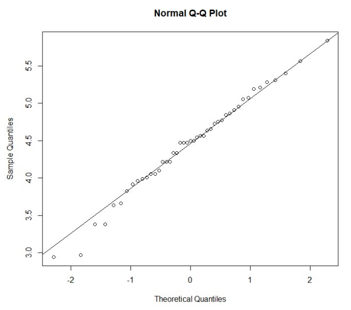
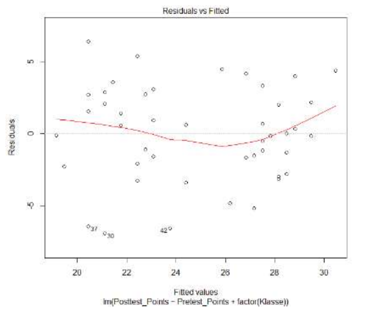
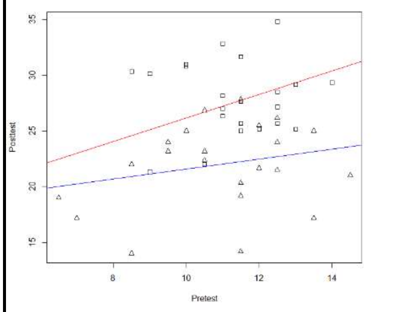
Post-survey		
Variable	Item	Scale
Question 1	To what extent do you feel that the learning materials has helped you? Why?	Open
Question 2 (treatment-specific)	What are your experiences when using the Smart Personal Assistant as a tutor?	open

Appendix H: Graphical check for assumptions for ANCOVA

Experiment 1 in high school

Test for normality	**Test for homogeneity of variance**	**Test for homogeneity of regression slopes**
		

Experiment 2 in vocational business school

<table>
<tr><th>Test for normality</th><th>Test for homogeneity of variance</th><th>Test for homogeneity of regression slopes</th></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
</table>

Appendix I: R Statistics for Pretest ANOVAS and Posttest ANCOVAS and T-tests

Experiment 1 in high school

Gender:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1  0.054  0.0538   0.209   0.65
Residuals  43 11.057  0.2571
```

School grades in business and law:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1  0.076 0.07564   0.348  0.558
Residuals  43  9.335 0.21710
```

Pre-experience with SPAs:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1   0.02  0.0159   0.006  0.941
Residuals  43 121.90  2.8348
```

Personal innovativeness:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1  0.754  0.7537   1.769  0.191
Residuals  43 18.324  0.4261
```

Pre-test results:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1   0.38   0.380   0.084  0.774
Residuals  43 194.82   4.531
```

Post-test ANCOVA:

```
                Df Sum Sq Mean Sq F value   Pr(>F)
Pretest_Points   1  520.1   520.1  42.005 8.07e-08
Treatment        1   55.9    55.9   4.514   0.0396
Residuals       42  520.0    12.4

Pretest_Points ***
Treatment      *
Residuals
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Gain scores T-test:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    8.009      1.761   4.547  4.4e-05 ***
Treatment      2.343      1.107   2.117     0.04 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Experiment 2 in vocational business school

Gender:

```
          Df Sum Sq Mean Sq F value Pr(>F)
Treatment  1  0.288  0.2881   1.179  0.284
Residuals 43 10.512  0.2445
```

School Grades in Business and Law:

```
          Df Sum Sq Mean Sq F value Pr(>F)
Treatment  1  0.892  0.8918   2.818    0.1
Residuals 43 13.608  0.3165
```

Pre-experience with SPAs:

```
          Df Sum Sq Mean Sq F value Pr(>F)
Treatment  1   1.64   1.636    0.36  0.552
Residuals 43 195.34   4.543
```

Personal Innovativeness:

```
          Df Sum Sq Mean Sq F value Pr(>F)
Treatment  1   1.66   1.657   1.005  0.322
Residuals 43  70.92   1.649
```

Pre-test results:

```
           Df Sum Sq Mean Sq F value Pr(>F)
Treatment   1   2.49   2.488   0.801  0.376
Residuals  43 133.49   3.104
```

Post-test ANCOVA:

```
               Df Sum Sq Mean Sq F value   Pr(>F)
Pretest_Points  1   42.4    42.4   3.356   0.0741
Treatment       1  386.4   386.4  30.573 1.88e-06
Residuals      42  530.8    12.6

Pretest_Points .
Treatment      ***
Residuals
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Gain scores T-Test:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.436      1.736   3.131  0.00313 **
Treatment      5.600      1.106   5.065 8.17e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```