When Learning Turns To Surveillance – Using Pedagogical Agents in Organizations

Andreas Bucher	Birgit Schenk	Gerhard Schwabe
University of Zurich	University of Public Administration Ludwigsburg	University of Zurich
bucher@ifi.uzh.ch	birgit.schenk@hs-ludwigsburg.de	schwabe@ifi.uzh.ch

Abstract

Workplace learning is often used to train employees systematically. New in this context is workplace learning with the help of a pedagogical agent (PA). Following Actions Design Research (ADR), this paper describes organizational training for telephone service using such PA. To develop the training, existing employee telephone service problems were analyzed, and the content of the learning program was determined based on this analysis. Subsequently, a PA was developed, implemented, and used in three municipalities. The evaluation of the learning outcome shows promising results but also yields some challenges: even though the employees improved in various aspects of the learning, they also developed a perception of surveillance. This research concludes with the formulation of design principles and suggestions for the organizational embedding of a PA in a workplace setting.

Keywords: Pedagogical Agent, Conversational Agent, Organizational Training, Workplace Learning

1. Introduction

Workplace learning includes not only the implicit learning of job-related skills, but also concerns the systematic training of employees' competencies, knowledge, and behavior. Herein, workplace learning can take place online, offline, and in single or group settings. As human capital is an essential factor for organizations and necessary to their competitive advantage, the development of human capital is a key element of modern management practices in organizations. However, workplace learning requires careful planning and execution, as it might be perceived as ineffective by employees who are not motivated to participate in organizational training programs (Illeris, 2003). Furthermore, workplace learning is often seen as expensive and associated with

unpredictable results, thus being cut early in economic downturns (Abd Rahman et al., 2013). Even as online training programs are more cost-effective than face-to-face training and easier to distribute among the workforce, such training is often based on click-through solutions that do not provide an immersive training experience (Baceviciute et al., 2022; Farrell, 2018). Besides failing to cover the needs of today's workforce, online training programs are also frequently interrupted by the employees' daily business or other activities negatively impacting the learning outcome (Ashton, 2004; Illeris, 2003).

Recent developments in Artificial Intelligence enabled by breakthroughs in Machine Learning (ML), Natural Language Processing (NLP), and Information Technology (IT) have opened new possibilities to integrate conversational user interfaces into everyday life. Nowadays, digital agents allow users in various private and professional contexts to naturally interact and communicate with computer systems in spoken or written language, similar to human-to-human interaction. In pedagogical contexts, like higher education or academia, previous studies have investigated the effect of applying an adaptation of digital agents, so-called Pedagogical Agents (PAs), in different learning situations. For instance, Winkler et al. (2020) demonstrated that PAs can guide learners in programming tasks using voice and text scaffolds. Similarly, Wambsganss et al. (2020) found that PAs can support learners in developing better argumentation and reasoning skills. Therein, PAs have been shown to effectively train children and students by providing an immersive and natural learning environment.

Despite the differences between workplace learning and educational or academic learning (e.g., its high focus on self-directiveness or its embeddedness in the organizational context), PAs are also seen fit for training the workforce on job-related skills or behavior (Meyer von Wolff et al., 2019). In this context, PAs might improve independence from trainers and sparring partners in terms of time and place or introduce the possibility for immediate feedback

HÍCSS

during the learning process (Curtis & Thomas, 2008). Therefore, developing PAs for organizational training to overcome the limitations of current learning programs for employees seems to be a promising direction. The lack of research on using PAs for workplace learning leaves practitioners and designers of organizational training programs confronted with a multitude of design options and possibilities for developing a PA. To provide more clarity about the feasibility and design of PAs in workplace learning, this study attempts to answer the following research questions by exploring the application of a PA to improve telephone services:

RQ1: *What are the effects of training employees with a Pedagogical Agent in telephone service?*
RQ2: *What can we learn from these effects for the design of a Pedagogical Agent for workplace learning?*

To answer these research questions, we investigate the possibilities and effects of using a PA for training employees of three public administrations in Germany. Establishing a high-quality telephone service is an important aspect of a service culture and of providing good citizen service. Previous approaches in telephone service training primarily focused on direct instructions, which required high effort and were difficult to repeat (Schenk & Gaeng, 2022). We believe that a PA could be a suitable solution to reduce the administrations' efforts, provide the possibility for continuous learning, and create a novel learning experience instead of relying on direct instructions. By following Actions Design Research (ADR) and combining design and research activities, we developed a PA to improve the telephone service and to establish a shared understanding of the administrations' policies. Our PA can independently call employees, analyze their telephone service, and provide spoken feedback to them (e.g., the correctness of their salutation). In a large-scale study, after performing more than 2,400 training calls over 2.5 months, we collected and analyzed the experiences of more than 280 employees who participated in the training. To further analyze the effects of the training, we evaluated the employee's behavior and performance with mystery calls before and after the training. Despite the positive effects of encouraging the employees to rethink and adapt their telephone service according to the organizations' policies, our study also shows that a large share of the employees perceived the training as means of surveillance. It seems that organizations need to carefully consider the training design when supporting organizational training with a PA as it is a balancing act between the perception of an immersive training or a tool for surveillance. Our study contributes to research on PAs in Information Systems (IS), Computer-Supported Cooperative Work (CSCW), and Human-Computer Interaction (HCI) by exploring the effects of using PAs for workplace training and by deriving five design principles for the development of such systems. Also, organizations and practitioners are informed about the possibilities and potential pitfalls that should be considered when applying PAs in this context.

2. Related Work

2.1. Organizational Training and Learning

Workplace learning has become a valuable tool in organizations' modern management practices and knowledge management processes. Research on workplace learning has been picked up by (human resource) management and by adult learning scholars (Illeris, 2003). In management science, workplace learning is often divided into formal (or explicit) and informal (or implicit) learning to characterize the learning processes. Formal learning can be understood as *learning at work* and considers formal education and training courses to systematically develop employees' knowledge, skills, or understanding for performing a certain task (Jacobs & Park, 2009; Kyndt et al., 2009; Sambrook, 2005). Therein, formal learning comprises institutionally sponsored training programs and takes place within a context intended for learning (Jacobs & Park, 2009).

Instead of dealing with high-level learning processes (i.e., whether learning occurs *at* or *in* work), adult learning scholars study the fundamental mechanisms of adult learning (Illeris, 2003). In particular, adult learning, also called *andragogy*, deals with the *"art and science of helping adults learn"* (Knowles, 1984). Contrary to *pedagogy*, andragogy describes learning as a self-directed inquiry. According to Knowles (1980), adult learners are 1) autonomous and self-directed, 2) goal-oriented, 3) relevancy-oriented, 4) practical, 5) have accumulated a foundation of experiences and knowledge, and 6) need to be shown respect (Russell, 2006). Accordingly, adult learners need to understand the meaning and importance of what is being learned and control the nature, timing, and direction of the learning process (Illeris, 2003; Rogers, 1969; Russell, 2006). Adult learners should also engage with practical or social problems to facilitate the learning process. In general, adult learning requires careful consideration of the learning process to build a shared understanding of why something is important to learn and to provide an environment that supports self-directiveness.

Considering these characteristics of adult learning, organizations can provide necessary conditions for formal and informal learning. As informal learning primarily occurs incidentally or implicitly, it is more difficult for organizations to steer the learning process. However, organizations can still foster informal learning by promoting a culture of learning and knowledge sharing, including developing knowledge management processes (Kyndt et al., 2009; Marsick, 2009). In contrast, there exist numerous approaches in formal learning to create knowledge transfer to and between employees and to create a shared understanding of the learning content, including face-to-face instructor-led training, e-learning, or blended learning (Martins et al., 2019; Noe et al., 2014; Shaw et al., 2009). Additionally, self-assessment, which is the process of deliberately assessing your own performance informed by external and trusted feedback, is another important method in formal workplace learning (Embo, 2015). Given the needs and demands of today's technology-savvy, cross-cultural, and geographically dispersed personnel, organizations have enforced the use of digital tools, like mobile apps or complex e-learning management systems and intelligent tutoring systems (ITS). Furthermore, research has started exploring virtual reality (VR)-based training to create an immersive learning experience and support formal learning (Baceviciute et al., 2022; Farrell, 2018). We believe such developments are a step in the right direction for motivating employees and providing an effective knowledge transfer. Nonetheless, organizations still rely frequently on rudimentary e-learning systems, and research on using immersive technologies in organizational training has primarily taken place in laboratory settings.

As there has been much research on using digital, voice-based PAs for educational and academic purposes, we believe that such interactive technology might become an effective addition to modern workplace learning programs. However, the literature does not yet address whether a PA is well received and accepted by employees in a workplace setting. There are also no indications of how the aspects of adult learning can and must be implemented in the design of a PA for workplace training.

2.2. Pedagogical Agents

Pedagogical agents are virtual characters that engage with learners in a virtual environment and support them in their learning process (Hobert & Wolff, 2019). Compared to traditional e-learning systems, PAs can offer more personalized feedback and tailored guidance throughout the learning process while establishing a perception of human-like interaction. Integrating PAs in learning settings has been proven effective in various studies within the fields of IS, CSCW, and HCI. For instance, Wambsganss et al. (2020) found that learners could develop better argumentation and reasoning skills when interacting with a PA compared to learning with proven argumentation writing support. In other studies, Winkler et al. (2020) showed positive effects of applying a scaffolding-based PA for teaching programming skills or that PAs can support the development of problem-solving skills (Winkler et al., 2019). Despite the success of integrating PAs in academic and educational contexts, only a few studies have investigated the use of PAs outside of these environments (Bendel, 2003; Curtis & Thomas, 2008; Meyer von Wolff et al., 2019). Hence, research on using PAs for workplace learning is still very nascent.

PAs allow for natural communication using speech recognition and NLP capabilities to simulate human-like interactions (Hobert & Wolff, 2019). PAs might also exhibit a virtual representation or embodiment (K. Kim et al., 2018) and be integrated into e-learning environments or ITS (Cook et al., 2017; Keller & Brucker-Kley, 2020; Kulik & Fletcher, 2016). Considering different design elements in the development of PAs is crucial. For instance, Heidig and Clarebout (2011) found that explanatory feedback is better for learning than purely corrective feedback, and it is easier to absorb when the feedback is spoken rather than presented as text. To better navigate the complex design space of PAs, Wellnhammer et al. (2020) provide a framework that characterizes PAs along various design dimensions including their physical appearance (e.g., gender or immediacy), communication forms (e.g., support of written or spoken interaction, type of voice) or didactical elements (e.g., learning setting, form or content).

Besides technical and didactical elements, PAs can also be distinguished according to their specific role in the learning process. Research by Kim and Baylor (2016) found significant changes in learning when changing the PA's role. Accordingly, PAs in the role of an *expert* result in better acquisition of knowledge, whereas PAs in the role of a *motivator* lead to increased motivation, and PAs in the role of a *mentor* increase learning and motivation (Y. Kim & Baylor, 2016). Most research on the design of PAs is associated with the impact on learners' motivation, including learners' self-efficacy, engagement, or anxiety (Martha & Santoso, 2019). However, only limited research is available on incorporating adult learning theories in the design of PAs.

3. Method

3.1 Research Approach

This study is situated in a larger research project to enhance the telephone service of three public administrations in Germany by new ways of digitally training the employees. Therein, this research intends to identify and address real problems. For this, Actions Design Research provides a suitable framework (Sein et al., 2011). ADR is a methodology that combines design and research activities. It includes the steps of problem formulation, intervention, evaluation of the intervention, reflection and learning, and formalization of learning. Therefore, our research included all steps:

- Problem formulation based on mystery calls and discussions with experts in the field
- Conceptualization of the problem "telephone reachability of public service employees" by analyzing and clustering the problems,
- Identification of an intervention by discussing the findings with experts.
- Reflection and evaluation of the intervention by conducting final mystery calls and discussions with experts in the field.
- Finally, we reflected on the gained insights and concluded the requirements for the further development of the PA.

To address our research questions, we followed a mixed-method research approach (see Table 1). In the first phase, "the problem identification," we conducted 1,383 mystery calls to identify citizens' experiences while trying to reach public administrations by phone. The mystery calls were conducted by five researchers over a period of 3 weeks in each municipality, starting from April to June 2021. The researchers documented their experiences after each call in a matrix. The matrix was developed according to the expected outcome (e.g., if the employee could be reached on 1st call, if the answering machine (AM) was switched on, or if the salutation was correct according to the administrations' policies) (Wiele et al., 2005). Analyzing the experiences, we identified and categorized problems that arose. This helped us understand the type of problems (e.g., organizational, personnel, and technological problems) and determine which issues could be addressed with digital training. Further discussion with five experts from practice and academia led to the decision to build a PA for the training. Followingly, we derived the requirements, such as the possibility to provide the

Steps	Methods	Products
Phase I – Zero Measurements		
Analysis of accessibility of telephone service	Mystery calls (5 field researchers)	Record of experiences and problems identified during the process in 3 public administrations (M; NU; S) (M: 180 phone-numbers, 360 calls; NU: 204 phone-numbers, 408 calls; S: 301 phone-numbers, 615 calls)
Identification of interventions	Expert discussions	Discussion of problems to cluster them, to identify next possible steps (measurements), and to decide on the most promising solution
Phase II – Measurements during the training		
Development and use of prototype	Recording of trainings	Developed prototype, which conducted 2.412 calls (M: 490 calls; NU: 749 calls; S: 1.173 calls) Record of experiences and problems identified during the training
Evaluation of the digital trainer	Questionnaires	Structured feedback about the prototype – pros and cons
Phase III – Final Measurements		
Analysis of accessibility of telephone service	Mystery calls (5 field researchers)	Record of experiences and problems identified during the process in 3 public administrations (M; NU; S) (M: 180 phone-numbers, 360 calls; NU: 202 phone-numbers, 404 calls; S: 301 phone-numbers, 615 calls)
Formalization of the findings	Expert discussions	Record of lessons learned

Table 1: Overview of steps, used methods, and products.

training at any time without additional personnel resources or the need for easy-to-understand exercises without time burden for the participants, to improve telephone reachability and the quality of public service employees.

Through reciprocal and user-centered design, we developed a prototype of a PA for training service skills. Our PA then automatically conducted 2,412 training calls over 11,5 weeks from June 23rd, 2021, to September 10th, 2021. During the training, the PA called every employee or phone number of the public administrations at least three times during business hours with a minimum of eight days between each call. Depending on the reachability, correct salutation, or call-back, an employee or phone number could be contacted by the PA up to six times. Overall, the training with the PA was intended to be 20min per employee. After the training, we used questionnaires to collect feedback from the trained people to evaluate the PA. The online questionnaires were distributed to the employees and covered aspects like the usefulness of the training, perceptions of the PA, or reflections on their own behavior. The final evaluation consisted of 1,379 mystery calls to answer whether the training improved the reachability of the employees by phone in the municipalities. Again, the mystery calls were conducted by five researchers over a period of 3 weeks in each municipality, starting July 19th till September $29^{th,}$ 2021. After the second round of mystery calls, we

reflected on the insights to derive design principles for the development of PA for workplace learning similar to the reflective approach by Möller et al. (2020).
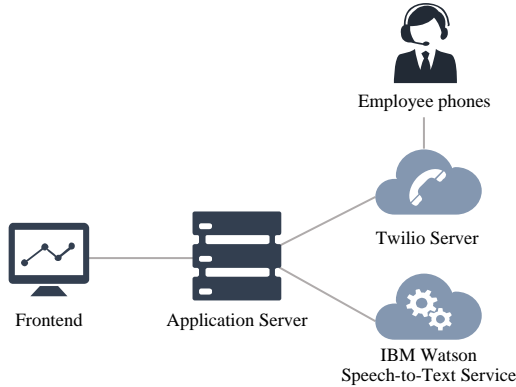
3.2 Design of the prototype

Figure 1: High-level architecture of COTA, our prototype.

The general idea behind our prototype was to develop a PA, called COTA (Conversational Organizational Training Agent), that can automatically call employees during their business hours, analyze their salutation (including the organizations' and employees' names, the bureau and department, and the overall greeting), to provide the possibility for self-reflection and to offer spoken feedback about the correctness of their salutation. Additionally, COTA gave recommendations to improve the salutation. Furthermore, COTA was designed to recognize AMs and analyze their announcements, leave a message for callbacks, and track the time until an employee returns the call. Overall, COTA took over the role of an expert (Y. Kim & Baylor, 2016) that managed the training, including the number of times and timing when it called an employee and allowed for natural interaction and explanatory feedback using spoken language (Heidig & Clarebout, 2011).

The high-level architecture of COTA is displayed in Figure 1. To enable the functionalities, we developed a Java-based backend (application server) that handles the call management, implements basic NLP functionalities to analyze the calls, and provides statistical evaluations. To establish a call between COTA and the employees of the public administrations, the prototype implements the "Twilio Voice" API. The audio snippets of these calls are then transcribed using the IBM Watson Speech-to-Text (STT) service and are analyzed by the backend based on pre-defined patterns and rules. After the training, the employees could review and compare their performance through a web-based user interface

(frontend). Also, representatives and managers of the public administrations were provided with aggregated statistics about their departments. From a technical point of view, the frontend allowed for configurations of the COTA (e.g., setting rules). Generally, the prototype was developed to support simultaneous training of different administrations. To account for the employees' privacy, the audio recordings of the calls were deleted after each call.

4. Results

4.1 Observations from the Mystery Calls

Measurements	City					
	M_0	M_1	Nu_0	Nu_1	S_0	S_1
Mystery Calls (N)	360	360	408	404	615	615
Reached at 1st call (%)	58%	53%	58%	60%	64%	52%
Correct salutation (%)	13%	45%	55%	61%	25%	68%
AM turned on (%)	37%	52%	12%	19%	52%	53%
AM salutation correct (%)	19%	66%	45%	30%	36%	69%
Call back (%)	5%	16%	5%	3%	18%	14%
Understandability (%)	89%	92%	75%	98%	98%	86%

Table 2: Results of the Mystery Calls (0 = baseline measurement, 1 = measurement after the training).

The mystery calls brought the weaknesses of the telephone reachability and the salutation of citizens by the administrative staff to the surface. As shown in Table 2, it was observable that two municipalities (*M, Nu*) performed worse than the third municipality (*S*) in the baseline measurement. This can be explained by the fact that municipality *S* has been sensitizing its employees to this issue for some time. In about 42% of cases, the responsible public employee in the two municipalities, *M* and *Nu,* could not be reached immediately. Only in municipality *S*, about 32% of the employees were not reachable at the first attempt. When no one could be reached, it was checked if the AMs were activated. Furthermore, we left a message for returning the call if an employee could not be reached and if the AM was activated. As a result, about 37% of the AMs were activated in municipality *M*, 12% in municipality *Nu*, and 52% in municipality *S*. After leaving a message on the AMs, we received a call-back from only 5% of the unreachable employees of municipality *M*. The same was true for municipality *Nu*. Only the employees of municipality *S* returned the call 18% of the time.

In all public administrations, there was no consistent salutation - either in the direct response from the employees or in the salutation on their AM. It ranged from a simple *"hello"* to *"<city>, <name>,*

what can I do for you?". When reaching the AM, the announcements differed from *"The call cannot be answered, please call back later."* or *"You are calling out of hours. These are...."*, regardless if the call was made within opening hours or not. Sometimes, we also encountered announcements like *"The telephone is not manned due to illness. Please dial the number...."*. Additionally, across the municipalities, we frequently faced the issue that employees spoke unclearly, too silent, or too fast making it impossible to understand the salutation.

After the training with COTA, we again conducted mystery calls to check if there had been any improvements because of the training. To provide comparable results, we again called the same phone numbers twice. Only in municipality *Nu*, two telephone numbers were removed after the baseline measurement, which reduced the number of calls to 404. As shown in Table 2, the final measurements showed an improvement for 12 of 18 items. These improvements primarily concern qualitative aspects of the telephone service, like the correct salutation, the correct announcement on the AMs, or the understandability. For instance, the percentage of correct salutations improved greatly in all municipalities. The largest improvement was in municipality S, where the percentage of correct salutations increased from 25% of the baseline-measurement to 68% after the training. Similar results were also observable in municipality *M*. Also, all municipalities increased the percentage of activated AMs, and most municipalities could enhance the announcements on the AMs.

However, we could not find improvements in all aspects of the training. For example, the reachability of the employees at the first call dropped from 64% to 52% in municipality *S* and from 58% to 53% in municipality *M*. There could be an explanation for this: one-third of the last measurement was conducted during the country's school holidays. Nonetheless, this shows that the training, which only refers to the greetings and does not include the rules, may not be sufficient on its own.

4.2 Experiences of the Employees

In total, 369 trained employees filled out the distributed online questionnaires. After removing all questionnaires that were not answered completely, the sample size was reduced to 281 questionnaires (*M*: 101; *Nu*: 80; *S*: 100). Table 3 presents the aggregated results for five single-choice questions. Overall, 32.38 % of the employees described having changed their greeting or behavior regarding reachability after the training. In an open question, some employees stated

Answers	N	Percentage
Question 1: Based on the trainer's feedback, have you changed your greeting (formula or pronunciation) or your behavior regarding accessibility?		
Yes	91	32.38 %
No, because everything was correct	150	53.38 %
No, other reason	40	14.24 %
Question 2: How do you feel about the trainer giving you personal feedback?		
Good	203	72.24 %
Bad	78	27.76 %
Question 3: When asked for a self-assessment, did you take the opportunity to think about it?		
Yes	166	59.07 %
No	115	40.93 %
Question 4: How did you feel about the trainer personally? As ...		
Advantageous	112	39.86 %
Annoying	169	60.14 %
Question 5: What did you perceive the trainer to be?		
Possibility for training	59	21.00 %
Tool for surveillance and monitoring	186	66.19 %
Other	36	12.81 %

Table 3: Questionnaire results after training with COTA.

to speak more slowly and clearly after the training, and indicated to answer the phone now in more detail. One employee even stated that his department called a meeting to discuss possible improvements and that since then everyone in the department has set their AMs appropriately. Nonetheless, 67.62 % stated that they did not change their greeting or behavior after the training. Here, the majority mentioned they did not change their salutation as it was correct in first place.

When asked about their perception of receiving personal feedback from COTA, most of the employees welcomed it (72.24%). One employee even mentioned in an open question that the personalized feedback *"was a bit unusual and also somewhat amusing to be praised by a machine. However, I remember it very positively as result. And to hear a criticism from a "real" person a is more annoying and has more of a reprimanding aspect"* [Employee 5]. Similarly, others mentioned that personalized feedback by COTA is perceived as neutral and therefore welcomed. However, COTA apparently did often not understand the employees correctly, which reduced the effectiveness of personalized feedback according to the comments of many employees. COTA unexpectedly called employees during their business day. Therefore, some employees mentioned that it was difficult to pay attention to the feedback when they had engaged in some work activity and were busy. For example, one employee stated that he/ she *"just had one of the most stressful days of the year, so every call was inopportune"* [Employee 73].

COTA also offered the possibility for self-assessment, in which the employees were able to comment on their performance while listening to the

recording of the call. According to our survey data, most employees took the chance to reflect on their performance (59.07%). Such self-assessment was perceived as a further tool for raising awareness of the learning content. Nonetheless, employees again criticized that COTA did not correctly understand them. For instance, one employee stated that he/ she *"[doesn't] see any reason to think about feedback that apparently didn't work."* [Employee 231]. Regarding the self-assessment possibility, another employee denied using this functionality by saying, *"I know that I'm doing it correctly and reliably. In my position, one is aware of one's external image."* [Employee 248].

In another question, we wanted to know how the employees felt about COTA personally. On the one side, 39.86% of the employees believed that the PA was advantageous. Despite the mentioned technical issues in understanding the employees, some employees stated that the training had a playful character and led to amusement among the employees. Even among those employees that perceived the training as advantageous, many raised the wish for a more human-like voice and noted that *"it feels very strange to be 'controlled' or evaluated by an artificial intelligence"* [Employee 114]. On the other side, a large share of the employees (60.14%) felt that COTA was annoying. Some stated that the PA called too often within in a short period of time and that it was frustrating to hear negative feedback in case the employees had answered correctly.

This negative attitude towards COTA was also apparent in the question about the employees' perceptions towards the PA. Here, the results clearly show that 66.19% of the employees perceived the training with COTA as an instrument for surveillance and control. Only 21.00% of the employees described the PA as a possibility for training. One employee stated that COTA was *"a nuisance, a control and an additional stress"* [Employee 258], whereas another employee described that he/ she *"felt 'checked' or 'controlled'. not trained."*. Clearly, the employees were skeptical towards COTA and had the feeling of being controlled and monitored.

5. Discussion

The results suggest that the training with COTA positively influenced the employees' learning process and outcome concerning the telephone service, especially their reachability. For example, the qualitative aspects of the baseline and final measurements improved greatly. Also, many employees noted in the questionnaires that they had changed their salutation in response to the training.

Overall, these positive results indicate the potential of using a PA in workplace learning and provide a starting point for expanding existing knowledge on PA from educational and academic contexts to organizational settings. By discussing the employees' perceptions and deriving five design principles (**DP**) for developing PAs in organizational settings, particularly in telephone services, we provide answers to our research questions. The composition of the DPs is inspired by Gregor et al. (2020) and followed an approach similar to Möller et al. (2020).

DP1: *A PA for workplace learning should offer employees in telephone service personalized but neutral feedback to create an engaging and respectful learning environment:* According to the results, the employees value personalized and neutral feedback from COTA. By combining individual and spoken feedback, a PA creates a feeling of personalization, not standardization. This is important to directly pinpoint the need for improvement and relevance for the training and to show the learners that they are valued and are not simply provided with standardized feedback (Illeris, 2003; Knowles, 1980; Russell, 2006). Additionally, the employees valued the neutral feedback of COTA. Some employees even mentioned that the neutral criticism of COTA was less intriguing than being criticized by a human. Apparently, the knowledge that a PA does not resemble a human being makes it easier to accept criticism. Designers of PAs can leverage this effect by actively creating a learning environment that is neutral rather than judgmental to the learner. For instance, designers can manipulate the PA's social cues, like its choice of words or tonality, to create an impression of neutrality. Hence, we propose that designers should provide the PA with the ability to offer personalized yet neutral feedback to create a learning environment that is neutral, respectful, and relevant to the learner's abilities.

DP2: *A PA for workplace learning should provide employees in telephone service with self-assessment possibilities to create acceptance of the learning*: providing the ability for self-assessment was a core element of COTA and was well perceived by the employees. An important component for the acceptance of the learning content was the recording and playback of the greetings. These were played as part of the feedback to better compare the actual greeting of the employee with the defined greeting of the administration. Therein, the employees could hear their salutation and recognize whether they spoke too slowly or quickly, too indistinctly or clearly. At the

same time, it became transparent to them to what extent they deviated from the given rules. By recognizing one's mistakes, the basis for accepting change is created (Embo, 2015; Illeris, 2003). Hence, self-assessment is a core factor for accepting and learning with a PA and should thus be considered in the design of the PA.

DP3: *A PA for workplace learning should consist of a playful character to foster motivation of employees in telephone service*: the playful character of COTA, which allowed for natural interaction using spoken language and provided spoken feedback and compliments, positively influenced the learning process. This is reflected in the comments of the employees, who mentioned that COTA created amusement and motivation for the participants. Especially in adult learning, it is necessary that learners are self-motivated to accept and engage in the learning process (Illeris, 2003). Hence, we believe that providing a playful learning environment with the help of a PA can increase motivation and improve the learning process. This is in line with earlier research on PAs, which primarily focused on social-cognitive frameworks to increase learners' motivation (Martha & Santoso, 2019). Similarly, we propose that designers should leverage social cues like natural speech or non-verbal behavior to increase the playfulness and interactivity of the learning environment and foster motivation.

Despite the positive effects of the training, we also encountered some negative feedback. Especially, the high number of employees that perceived COTA as a tool for surveillance and monitoring provides evidence that the integration of a PA in an organizational context must be considered carefully. Even as we deleted the recordings and only granted access to the employee's own results and aggregated statistics, the majority still perceived the PA as surveillance and monitoring. Therefore, we assume that this issue is rooted deeper in the underlying characteristics of the PA, but also in the overall context of the training.

DP4: *Instead of acting autonomously, a PA for workplace learning should offer employees in telephone service the possibility to control the learning process*: According to theories on adult learning, learners should be provided with the possibility to control the nature, timing, and direction of their learning (Illeris, 2003; Knowles, 1980; Russell, 2006). Hence, the learners' self-directiveness and autonomy should not be harmed by the PA. However, COTA automatically triggered the calls

during business hours without the intervention of the employees. Besides creating a feeling of being externally controlled, the autonomous PA frequently interfered with the employees' work schedule. Unsurprisingly, many employees reported being disturbed in their everyday work. Especially for those employees that were heavily burdened by their work, COTA's autonomous and unexpected calls presented another source of stress and led to frustration. To counteract the feeling of being monitored and to better account for the peculiarities of adult learning, we propose that PAs should inhibit mechanisms to control the learning process instead of being autonomous.

Eventually, the overall context of the training embedded in the baseline and final measurements might have also reinforced the perception of surveillance. Before the mystery calls, the municipality management informed the employees about the upcoming evaluations to comply with privacy regulations. It is conceivable that some employees could not distinguish the training with COTA from the evaluation phase, thus confusing the PA with a tool for surveillance and monitoring. To avoid this, we propose to refrain from embedding the training with a PA into such evaluations.

Besides perceiving the training with the PA as surveillance, we also encountered negative effects related to technical issues, leading to the last DP for developing PA in workplace learning.

DP5: *A PA for workplace learning should provide employees in telephone service transparency and detailed explanations about what was understood and why certain feedback was given*: The results indicate that a PA became a nuisance when the underlying technology is not yet sophisticated to produce results that are very close to reality. For example, the speech recognition was often a point of criticism from the employees, as COTA frequently failed to understand correct salutations. However, this criticism might also be rooted in the lacking transparency of the feedback because the PA is non-deterministic and does not follow simple logic or rules (e.g., correct/ incorrect). Even though COTA explained what aspects of the salutations were not understandable, it did not make it clear how it arrived at its conclusions. For instance, instead of only telling that the employee's name and department were not understandable, COTA could have elaborated more on the reasons why it could not understand it (e.g., the employee spoke too fast or mumbled) or provide hints (e.g., the transcript of the recording). However, since the results were essential

to the employees, they were annoyed when the PA did not share the same judgment. Therefore, designers should carefully consider ways to create transparency in the PA's feedback mechanism to allow for a better understanding of why certain feedback was given.

Lastly, the results show that the overall reachability of employees (i.e., the percentage of employees reached at the 1st attempt) had hardly changed from the baseline to the final measurement and thus because of the training. This may be because COTA did not give feedback on all aspects of telephone service, but rather trained on qualitative aspects related to the salutation. Hence, it can be assumed that there are hardly any spillover effects beyond the training's core. To develop the employees in all aspects of the underlying subject (e.g., telephone service), the PA should include all aspects.

6. Conclusion, Limitations, & Future Research

The results of developing, implementing, and evaluating a PA in an organizational setting show that shared knowledge can be built, and behavior change is possible. To achieve such learning outcomes, a PA in a workplace setting must meet the requirements of adult learning and the learner's expectations. These requirements are summarized under five design principles that should guide designers and practitioners in developing PAs for workplace learning. Followingly, a PA should: 1) provide **personalized and neutral feedback** on learning outcomes (i.e., respond directly to the employee's performance). 2) allow for **self-regulation and autonomy** in the learning process (i.e., employees determine the time and duration of the training to keep their minds free from learning and not to increase the burden in times of high workload). 3) motivate through the **playful character** of the learning environment and increase the excitement of the learning process. 4) **enable self-assessment** of the employees' performance to foster the understanding of their performance compared to the overall training objective. 5) **provide transparency and detailed explanations** of how and why the PA arrived at specific feedback and what was expected. Additionally, organizational aspects must be considered. For example, a training with PA should not be integrated into performance evaluations as this could lead to monitoring fears among participants.

This study is not free of limitations, which also lead to future research. First, the observed and tested PA for workplace learning was limited to a small

aspect of the telephone service: the greeting. Other knowledge elements such as rules for good telephone reachability, how to use the telephone and answering machine, etc., and behavioral aspects such as friendliness were not recorded. Here it would be interesting to see what design recommendations are necessary to train these aspects successfully. Additionally, the PA had only the role of providing feedback. In learning settings, however, different roles (e.g., tutor, sparring partner, etc.) are necessary to create a holistic learning experience. Adding these roles might require other design elements. Further, due to the high practicality of this study, the questionnaires for collecting the employees' feedback after their training with our PA were constructed without scientifically validated constructs. Future work could include such constructs for greater rigor. Finally, it was not verified if the employees recognized the difficulties that arise for citizens from the sometimes understandable or incomplete forms of greeting. Whether a PA could help identify such causal chains would also be an interesting aspect to be investigated in further work.

Besides the limitations and possibilities for future work, we believe this research expands the existing knowledge on PAs by developing a PA for workplace learning and evaluating its effect on the learning outcome. This study adds to current research in IS, CSCW, and HCI on PAs and to research on workplace learning. Further, the design principles support practitioners and designers in developing PAs for an organizational context.

Acknowledgments

References

Abd Rahman, A., Imm Ng, S., Sambasivan, M., & Wong, F. (2013). Training and organizational effectiveness: Moderating role of knowledge management process. Eur. J. Train. Dev., 37(5), 472–488.

Ashton, D. N. (2004). The impact of organisational structure and practices on learning in the workplace. Int. J. Train. Dev., 8(1), 43–53.

Baceviciute, S., Cordoba, A. L., Wismer, P., Jensen, T. V., Klausen, M., & Makransky, G. (2022). Investigating

the value of immersive virtual reality tools for organizational training: An applied International study in the biotech industry. J. Comput. Assist. Learn., 38(2), 470–487.

Bendel, O. (2003). Pädagogische Agenten im Corporate E-Learning [Doctoral Thesis]. University of St. Gallen.

Cook, S. W., Friedman, H. S., Duggan, K. A., Cui, J., & Popescu, V. (2017). Hand Gesture and Mathematics Learning: Lessons From an Avatar. Cogn. Sci., 41(2), 518–535.

Curtis, C. K., & Thomas, K. M. (2008). Animated Pedagogical Agents in Interactive Learning Environment: The Future of Air Force Training? Airforce Research Laboratory.

Embo, M. (2015). Integrating workplace learning, assessment and supervision in health care education. Maastricht University.

Farrell, W. A. (2018). Learning Becomes Doing: Applying Augmented and Virtual Reality to Improve Performance. Performance Improvement, 57(4),19–28.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. J. Assoc. Inf. Syst., 21, 1622–1652.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? Educational Research Review, 6(1), 27–54.

Hobert, S., & Wolff, R. M. von. (2019). Say Hello to Your New Automated Tutor—A Structured Literature Review on Pedagogical Conversational Agents. WI.

Illeris, K. (2003). Workplace learning and learning theory. Journal of Workplace Learning, 15(4), 167–178.

Jacobs, R. L., & Park, Y. (2009). A Proposed Conceptual Framework of Workplace Learning: Implications for Theory Development and Research in Human Resource Development. Hum. Resour. Dev. Rev., 8(2), 133–150.

Keller, T., & Brucker-Kley, E. (2020). Design hints for smart agents as teachers in virtual learning spaces. In P. Kommers, B. Bontchev, & P. Isaias (Eds.), Proc. of the 18th Int. Conf. e-Society 2020 (pp. 111–118).

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., & Welch, G. F. (2018). Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. 2018 IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR), 105–114.

Kim, Y., & Baylor, A. L. (2016). Research-Based Design of Pedagogical Agent Roles: A Review, Progress, and Recommendations. Int. J. Artif. Intell. Educ., 26(1), 160–169.

Knowles, M. S. (1980). The modern practice of adult education: From pedagogy to andragogy (Rev. and updated). Cambridge Adult Education.

Knowles, M. S. (1984). The Art and Science of Helping Adults Learn. In Andragogy in Action: Applying Modern Principles of Adult Learning (pp. 1–21).

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. Review of Educational Research, 86(1),42–78.

Kyndt, E., Dochy, F., & Nijs, H. (2009). Learning conditions for non-formal and informal workplace learning. Journal of Workplace Learning, 21(5), 369–383.

Marsick, V. J. (2009). Toward a unifying framework to support informal learning theory, research and practice. Journal of Workplace Learning, 21(4), 265–275.

Martha, A. S. D., & Santoso, H. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. Journal of Educators Online, 16(1).

Martins, L. B., Zerbini, T., & Medina, F. J. (2019). Impact of Online Training on Behavioral Transfer and Job Performance in a Large Organization. J. Work Organ. Psychol., 35(1), 27–37.

Meyer von Wolff, R., Masuch, K., Hobert, S., & Schumann, M. (2019). What Do You Need Today? - An Empirical Systematization of Application Areas for Chatbots at Digital Workplaces. 25th Am. Conf. on Inf. Sys.

Möller, F., Guggenberger, T. M., & Otto, B. (2020). Towards a Method for Design Principle Development in Information Systems. In S. Hofmann, O. Müller, & M. Rossi (Eds.), Designing for Digital Transformation. Co-Creating Services with Citizens and Industry (Vol. 12388, pp. 208–220).

Noe, R. A., Clarke, A. D. M., & Klein, H. J. (2014). Learning in the Twenty-First-Century Workplace. Annu. Rev. Organ. Psychol. Organ., 1(1), 245–275.

Rogers, C. R. (1969). Freedom to learn: A view of what education might become. C. E. Merrill Pub. Co.

Russell, S. (2006). An overview of adult-learning processes. Urologic Nursing, 26, 349–352, 370.

Sambrook, S. (2005). Factors Influencing the Context and Process of Work-Related Learning: Synthesizing Findings from Two Research Projects. Human Resource Development Int., 8(1), 101–119.

Schenk, B., & Gaeng, M. (2022, March 7). Ein „Must have" für einen guten Bürgerservice. Future4Public. https://f4p.online/2022/03/07/ein-must-have-fuer-einen-guten-buergerservice/

Shaw, R. S., Chen, C. C., Harris, A. L., & Huang, H.-J. (2009). The impact of information richness on information security awareness training effectiveness. Computers & Education, 52(1), 92–100.

Wambsganss, T., Niklaus, C., Cetto, M., Söllner, M., Handschuh, S., & Leimeister, J. M. (2020). AL: An Adaptive Learning Support System for Argumentation Skills. Proc. of the 2020 CHI Conf. on Hum. Fac. in Com. Sys., 1–14.

Wellnhammer, N., Dolata, M., Steigler, S., & Schwabe, G. (2020). Studying with the Help of Digital Tutors: Design Aspects of Conversational Agents that Influence the Learning Process. Hawaii Int. Conf. on Sys. Sci.

Wiele, T. V. der, Hesselink, M., & Iwaarden, J. V. (2005). Mystery shopping: A tool to develop insight into customer service provision. Total Quality Management & Business Excellence, 16(4), 529–541.

Winkler, R., Büchi, C., & Söllner, M. (2019). Improving Problem-Solving Skills with Smart Personal Assistants: Insights from a Quasi Field Experiment. ICIS.

Winkler, R., Hobert, S., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020). Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent. Proc. of the 2020 CHI Conf. on Hum. Fac. in Com. Sys., 1–14.