frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 06 June 2025
DOI 10.3389/fcomp.2025.1553441

OPEN ACCESS

EDITED BY
Bin Yang,
Jiangnan University, China

REVIEWED BY
Michele Gattullo,
Politecnico di Bari, Italy
Chris Young,
Texas Department of Transportation,
United States

*CORRESPONDENCE
Sarah Depenbusch
✉ sarah.depenbusch@uni-paderborn.de

RECEIVED 09 January 2025
ACCEPTED 15 May 2025
PUBLISHED 06 June 2025

CITATION
Depenbusch S, Schaper N, Schürmann M and Schumacher J-P (2025) VR-based avatar videos as an effective tool for process training in the context of digitalization? *Front. Comput. Sci.* 7:1553441.
doi: 10.3389/fcomp.2025.1553441

VR-based avatar videos as an effective tool for process training in the context of digitalization?

Sarah Depenbusch[1]*, Niclas Schaper[1], Mirko Schürmann[1] and Jan-Philip Schumacher[2]

[1]Department of Industrial and Organizational Psychology, Institute for Human Sciences, Paderborn University, Paderborn, North Rhine-Westphalia, Germany, [2]Department of Industrial and Organizational Psychology, Institute of Psychology, Osnabrueck University, Osnabrueck, Lower-Saxony, Germany

In the context of digitalization, work processes are subject to constant change. To achieve overall process efficiency, it should be ensured that employees have a deep understanding of the work processes in which they are involved. Preliminary research has shown that the utilization of virtual reality (VR) environments, which visualize employees' workspaces and present VR avatars that demonstrate work process steps, can enhance employees' understanding of (future) work processes. However, implementing such virtual environments entails certain challenges, such as the necessity of training employees in the utilization of VR technology. Thus, the delivery of VR avatar simulations in a video format (VR-based avatar video) may present a flexible alternative solution. Focusing on related work, it can be assumed that VR-based avatar videos (VRA videos) help learners build a coherent mental model of their work processes by providing contextualized visual information that is close to real life. Furthermore, the visual design elements included in a VRA video (e.g., the VR avatar and virtual workspace) may increase employees' motivation to learn. Despite the potential benefits of VRA videos, critics may argue that these videos contain an excessive amount of visual detail, thus increasing learners' cognitive load. Due to these contradicting opinions, the present study investigates the potential advantages of a VRA video in enhancing employees' understanding of work processes compared to a schematically designed voice-over slides video (VOS video). Furthermore, the study compares the motivational impact of both videos. In an online experimental study, participants ($N = 121$) were randomly assigned to either the VRA or the VOS video group. One-way ANOVAs revealed that the VRA video group achieved significantly better transfer scores than the VOS video group. Results of the motivation questionnaires (based on the ARCS model) demonstrated that attention (ARCS-A), relevance (ARCS-R), and satisfaction (ARCS-S) were significantly higher in the VRA video group than in the VOS video group.

KEYWORDS

learning video, virtual reality, avatar, work processes, motivation to learn, contextual learning, multimedia learning

1 Introduction

The digitalization of work processes can range from implementing new information and communication technologies to fully automating entire production systems (Mueller et al., 2022, 2023). Regardless of the digitalization strategy that is adopted, the competencies required of employees change when work processes are digitalized. Therefore, employees need to develop a deep understanding of the digitalized work processes in order to carry them out effectively (Hirsch-Kreinsen et al., 2020; Leyer et al., 2021). Against this background, process training that develops process understanding is becoming increasingly important (Leyer et al., 2019, 2021). Virtual reality (VR) environments offer new methods to facilitate employees' understanding of recently digitalized work processes (Aysolmaz et al., 2016; Leyer et al., 2019, 2021). VR can be described as immersive technologies using a Head-Mounted Display (HMD) to simulate interactive virtual environments in which users can interact with virtual objects in an intuitive way (e.g., Mueller et al., 2022; Wohlgenannt et al., 2020). Using VR, it is possible to visualize abstract process models (diagrams that represent work processes using standard process notations) in virtually replicated work environments (Aysolmaz et al., 2016; Leyer et al., 2019, 2021). For process training, these environments are typically enriched with VR avatars simulating the practical execution of the process steps (e.g., Aysolmaz et al., 2016; Guo et al., 2013; Leyer et al., 2019, 2021). Previous research has already demonstrated that respective contextualized simulations can support employees' understanding of work processes as well as their motivation to learn (e.g., Leyer et al., 2021). Motivation to learn is considered a key factor in facilitating employees' engagement to comprehending work processes (Leyer et al., 2021; Mayer, 2005). Notwithstanding the advantages described, the utilization of VR environments in process training is not without its challenges.

Due to the sophisticated nature of VR hardware (e.g., HMD, VR controller), VR usage is restricted to designated rooms or spaces (Mueller et al., 2022). In addition, time-consuming introductory sessions are required for employees to learn how to use VR environments (e.g. how to operate VR controllers; Mueller et al., 2022, 2023). Consequently, the utilization of contextualized VR avatar sequences presented in a conventional 2D video format emerges as a potentially efficacious alternative to facilitate employees' process understanding and enhance their motivation to learn. As demonstrated in previous studies, learning videos offer greater flexibility and lower costs compared to VR environments, while also having similar learning effects (cf. Grassini et al., 2020).

However, it is also important to acknowledge a critical perspective on the utilization of VRA videos. Existing research suggests that visualizations like those in the VRA video (e.g., an animated VR avatar or a virtually replicated work environment) can be cognitively demanding (Scheiter et al., 2009; Um et al., 2012). From this perspective, schematic and static visualizations are preferred for process training as they can be processed with less cognitive effort (cf. Scheiter et al., 2009). Against this background, the present study investigates whether a VRA video is more effective in promoting process understanding and motivation to learn than a voice-over-slides video (VOS video) that utilizes static and schematic graphics to convey work processes. For this purpose, an online experiment has been conducted. Participants ($N = 121$) were randomly assigned to either the VRA or the VOS video group.

Abbreviations: VRA video, VR-based avatar video; VOS video, Voice-over slides video.

The two videos presented the same newly digitalized warehouse process of a small and medium-sized glass wholesaler. Both the VRA and VOS videos were viewed by the participants on a 2D screen (such as a smartphone or PC). After viewing the respective videos, participants worked on a cloze test (measuring retention), answered two problem-solving questions (measuring transfer) and completed a motivation questionnaire (measuring attention, relevance, confidence, and satisfaction; ARCS model, Keller, 2010). One-way ANOVAs were conducted to identify significant group differences in terms of acquired process understanding (retention and transfer scores), and motivation to learn (attention, relevance, confidence, and satisfaction).

2 Theoretical and conceptual background

2.1 Fostering process understanding using learning videos

Process understanding is defined as the comprehension of individual process elements (e.g., process steps or roles/activities involved in the process) and their relationships (cf. Burton-Jones and Meso, 2008). According to Recker and Dreiling (2011), process understanding can be conceptualized using two variables: retention and transfer. Retention is defined as the ability to comprehend and recall the work process, with particular reference to its inherent process elements and their relations (Recker and Dreiling, 2011). Transfer, on the other hand, refers to the ability to apply the aforementioned process understanding to problem-solving questions (Recker and Dreiling, 2011). A high level of retention but a low level of transfer indicates a superficial process understanding (Recker and Dreiling, 2011). Conversely, a high level of both retention and transfer signifies a deep process understanding (Recker and Dreiling, 2011). The development of process understanding through the utilization of learning videos, such as the VRA and VOS videos, can be conceptualized as a form of multimedia learning, which encompasses the acquisition of knowledge from both words and pictures (Mayer, 2014, 2021). In learning videos, words are presented primarily as audio commentary and/or as printed text on a screen (Mayer, 2021). Pictures can be presented as static graphics, animations, schematic drawings, or real pictures (photos; Koese et al., 2021). According to cognitive theory of multimedia learning (CTML), the processing of words and pictures occurs in two separate channels (Mayer, 2005). The capacity of both channels is limited, meaning that only a restricted amount of information can be processed simultaneously (Mayer, 2005). First, visual and acoustic information is selected (using sensory memory) and transferred into working memory (Mayer, 2005). Second, the selected information is organized (in working memory) into two "channel-specific" mental models, a verbal and a pictorial model (Stiller et al., 2020). Subsequently, the verbal and pictorial models are integrated into a coherent mental model with the help of prior knowledge drawn from long-term memory (Scheiter et al., 2020). The establishment of such a coherent mental model is considered a prerequisite for learners to apply the acquired knowledge in new situations (transfer; Mayer, 2005). CTML further distinguishes between three types of information processing (Mayer, 2005). Essential processing refers to mentally representing the learning content in working memory (Mayer, 2005). Generative processing is defined as the process of comprehending the learning content (constructing meaning; Mayer, 2014). The focus here is on integrating the verbal and pictorial models formed into a coherent mental model (Mayer, 2014). Generative processing is reflected in good transfer performance (Mayer, 2014). In contrast, extraneous processing is not focused on the learning content (Mayer, 2005). It is evoked by poorly designed learning material (Mayer, 2005). According to Mayer (2014), essential processing should be managed, extraneous processing should be reduced, and generative processing should be encouraged (Mayer, 2005).

One method of encouraging generative processing is to utilize pedagogical agents (Mayer, 2014). Pedagogical agents (PAs) are characters integrated into multimedia learning material to support learning (Peng and Wang, 2022; Wang et al., 2018). The VR avatar presented in the VRA video is also considered a pedagogical agent, as it is used to support employees' process understanding. According to social agency theory (Mayer, 2014), pedagogical agents or avatars exhibit social cues (e.g., facial expressions, gestures/body movements), which can induce a feeling of actually being in a social interaction situation ("social presence"). This, in turn, has been shown to motivate learners to invest more effort to understand the presented material (generative processing), thus leading to better transfer test scores (Mayer and DaPra, 2012). Mayer (2014) proposes various principles to guide the implementation of social cues, with the aim of stimulating the aforementioned social processes. The present study primarily focuses on the embodiment principle, according to which individuals learn better when the agents or avatars display human-like gestures, movements, eye contact, or facial expressions (Mayer, 2014). In accordance with the principles of CTML and social agency theory, the VRA video displaying a VR avatar with human-like body movements may also have the capacity to facilitate employees' engagement in understanding the presented work process. It may be argued that the presence of the avatar, as well as its human-like body movements, induce a social presence, resulting in higher generative processing (cf. Mayer, 2014).

The VR avatar in the VRA video does not perform its work activities in front of a white background, but rather in the virtually replicated 3D warehouse environment of the glass wholesaler. Consequently, the actions of the VR avatar are situated within a virtual work context. This approach has the potential to facilitate contextual learning (e.g., Setyowati et al., 2023). The concept of contextualization entails the establishment of relationships between learning content and its application in specific situations (e.g., Parchmann and Kuhn, 2018). In this manner, the processing and comprehension of the learning content is facilitated (cf. Chen et al., 2019; Setyowati et al., 2023). Consequently, it can be hypothesized that the VRA video fosters enhanced process understanding by providing a contextual framework for the content to be learned.

2.2 Increasing motivation to learn using learning videos

In addition to the potential of the VRA video to facilitate process understanding, it is investigated whether it leads to higher motivation to learn than a static and schematic voice-over-slides video. Motivation to learn is defined as a person's intention or willingness to learn certain content or skills (cf. Zander and Heidig, 2019). According to the ARCS model (Keller, 2010), instructional material should be designed to increase learners' attention (ARCS-A), their perceived relevance of the material (ARCS-R), their confidence in their ability to learn (ARCS-C), and their satisfaction with the learning experience (ARCS-S; Keller, 2010). Attention (ARCS-A), for example, can be increased by the presentation of visually appealing and engaging visualizations, such as animations, bright colors, or pedagogical agents/avatars (Chin et al., 2016; Zander and Heidig, 2019). The perceived individual relevance of the subject matter to the learner (ARCS-R) can be emphasized by relating the learning content to the context of application in the real world. This corresponds to the recommendations of contextualized learning described above (cf. Parchmann and Kuhn, 2018; Schmid, 2023). Moreover, delivering instructional content via a VR avatar can serve to underscore the significance of the learning material (Zander and Heidig, 2019). This phenomenon can be explained by the social agency theory (Mayer, 2014), which asserts that the avatar is perceived by the learner as a social interaction partner, thereby rendering the learning content more significant (cf. Stiller et al., 2020). Learners' confidence in successfully completing the learning unit (ARCS-C) can be fostered by the clear structuring of learning material or transparent explanation of learning objectives (cf. Keller, 2010). Furthermore, the integration of avatars is recommended, as the resulting humanization of the learning environment has been shown to enhance learners' confidence in their ability to learn (cf. van der Meij et al., 2015). Finally, to increase learners' satisfaction with the learning experience (ARCS-S), the use of appealing learning material (e.g., warm colors, anthropomorphism) is advocated as it has been shown to induce positive emotions, such as joy (cf. Keller, 2010; Um et al., 2012).

Against this background, it becomes clear that the VRA video provides special potential to enhance motivation to learn. For example, presenting the anthropomorphic VR avatar and the virtually replicated warehouse environment can increase learners' attention (ARCS-A) as well as their satisfaction with the learning experience (ARCS-S). Through creating the illusion of a social interaction, the VR avatar may also increase the perceived relevance of the learning content (ARCS-R) as well as learners' confidence in their ability to learn (ARCS-C). The relevance factor (ARCS-R) can additionally be enhanced by the contextualization of the learning content using the virtually replicated work environment. With respect to the confidence factor (ARCS-C), it can be further mentioned that the representation of the process steps by the VR avatar in the course of action can also strengthen learners' confidence in successfully completing the learning session.

3 Related work and hypotheses

While the aforementioned theoretical explanations posit the potential benefits of VRA videos in supporting process understanding and motivation to learn, the extant research is inconsistent as to whether the visualizations of VRA videos are conducive or detrimental to learning (e.g., Scheiter et al., 2009). In contrast to the expected positive effects, there are views that the visualizations presented in VRA videos are a source of extraneous processing (cf. Hegarty, 2004; Hoeffler and Leutner, 2007; Yarden and Yarden, 2010). Consequently, VOS videos may be the preferred choice, as their schematic and static nature may result in less cognitive load (cf. Scheiter et al., 2009). However, the current state of research provides insufficient empirical evidence to substantiate the assumption that the visual elements of the VRA video actually evoke extraneous processing (cf. Scheiter et al., 2009). In contrast, prior research has revealed that presenting work processes demonstrated by humanlike VR-avatars in a virtually replicated work environment can foster employees' process understanding (e.g., Leyer et al., 2019, 2021). In a comparative study, Leyer et al. (2021) examined the learning efficacy of VR-based process and avatar visualizations with that of a conventional 2D process model (e.g., visualizing work processes using abstract geometric forms, cf. Kathleen et al., 2014). The results show that the VR-based process and avatar visualizations led to significantly better process understanding in terms of faster and more accurate recall of process information (retention; Leyer et al., 2021). In their conceptual study, Guo et al. (2013) also emphasize the advantages of employing contextualized process visualizations with VR avatars demonstrating respective process steps. The researchers argue that the realistic and contextualized presentation of work processes enables employees to connect the process information with their existing knowledge or practical experiences. This, in turn, frees cognitive capacity for meaningful learning (Guo et al., 2013). In view of this, the 3D warehouse environment, which is virtually replicated in the VRA video, may facilitate process understanding.

According to social agency theory, it can be further assumed that the VR avatar induces a social presence, thus increasing learners' active cognitive processing to understand the warehouse process (generative processing; Mayer, 2014). In their meta-analysis, Castro-Alonso et al. (2021) show that the mere presence of pedagogical agents or avatars, regardless of whether they are embodied or static, results in enhanced retention and transfer test scores. In contrast, Davis (2018) reveals that the embodiment of PAs or avatars (e.g., human-like body movements, gestures, or facial expressions) is central to supporting better retention and transfer, thus highlighting the embodiment principle (Davis, 2018). Wang et al. (2018) have obtained analogous results when comparing the learning effectiveness of an online learning unit (for synaptic transmission) containing an embodied female PA (with a female voice, human-like posture, gaze, and pointing gestures) with the same online learning unit without this PA. In accordance with the embodiment principle, the results indicate that the learning unit containing the pedagogical agent led to superior retention and transfer scores (Wang et al., 2018).

Based on these findings, the VRA video may offer potential advantages for improving employees' process understanding in terms of retention and transfer. As the VOS video does not include these visual elements, respective positive effects on process understanding are not expected. This leads us to the following hypotheses:

Hypothesis 1a: The VRA video leads to better retention scores than the VOS video.

Hypothesis 1b: The VRA video leads to better transfer scores than the VOS video.

As already indicated, the VRA video not only provides special potential to facilitate process understanding but also to enhance motivation to learn. For instance, Leyer et al. (2021) found that using virtually replicated work environments comprising human-like VR avatars not only facilitated employees' process understanding but also their motivation to learn. Focusing on the ARCS motivation model (Keller, 2010), Jong (2023) examined how a VR environment simulating 3D virtual classrooms with teaching scenarios influenced the motivation of prospective teachers. The results demonstrate that an authentically modeled learning environment (which corresponds to the later application context) fosters curiosity and interest among the prospective teachers, thereby enhancing their attention (ARCS-A). In addition, the contextuality and realism generated by the virtual classroom helped clarify the importance of the learning content for the teachers' future professional lives (ARCS-R). Furthermore, the positive feelings of learners (e.g., joy) were increased, which can positively contribute to their satisfaction with the learning experience (ARCS-S; Jong, 2023). However, the learners' confidence in their ability to learn (ARCS-C) was not enhanced due to concerns about the comfort and user-friendliness of the VR environment (Jong, 2023).

The motivational potential of the VRA video can be further attributed to its animated VR avatar (e.g., Chin et al., 2016; Dinçer and Doganay, 2017). Chin et al. (2016) investigated the benefits of an animated, cartoon-like pedagogical agent (PA) in a digital learning platform to promote primary school students' motivation to learn (ARCS factors) in science education. The results show that the use of the PA led to an increase in all ARCS factors. The high attention of the learners (ARCS-A) is attributed to the observation that the learning content appears more engaging and interesting through the use of the PA (Chin et al., 2016). The perceived relevance (ARCS-R) of the learning content to the school students can be ascribed to its delivery by a PA, who creates a sense of social interaction. The high level of confidence exhibited by the learners (ARCS-C) is attributed to the utilization of human-like language and gestures by the PA. These elements serve to engender a sense of familiarity during the learning process, thereby fostering learners' confidence in their capacity to learn (Chin et al., 2016). Finally, it is argued that learner satisfaction (ARCS-S) was increased by the interesting and visually appealing design of the learning material. In particular, learners' satisfaction was expressed in a higher level of joy during learning (Chin et al., 2016).

Dinçer and Doganay (2017) analyzed the effects of PAs in a digital learning platform (used to promote Excel skills) on the motivation to learn (ARCS factors) of fifth-grade students. They also investigated whether the possibility of choosing between several PAs (with different designs) leads to different effects on the ARCS factors. The results obtained demonstrate that there is no significant difference between the effects of "fixed" and "selectable PAs" on the ARCS factors. However, it is generally found that the use of PAs (e.g., human-like, cartoon-like) contributes to significantly higher ARCS factors than using the digital learning platform without PAs.

Based on the above study results, it can be postulated that both the virtual replica of the warehouse environment and the animated and anthropomorphic VR avatar presented in the VRA video have great potential for increasing the ARCS factors. As the VOS video does not contain these visual elements, positive effects on the ARCS factors may not be realized. Accordingly, we assume:

Hypothesis 2a: The VRA video leads to higher attention scores (ARCS-A) than the VOS video.

Hypothesis 2b: The VRA video leads to higher relevance scores (ARCS-R) than the VOS video.

Hypothesis 2c: The VRA video leads to higher confidence scores (ARCS-C) than the VOS video.

Hypothesis 2d: The VRA video leads to higher satisfaction scores (ARCS-S) than the VOS video.

4 Materials and methods

4.1 Research design

A single-factor study design was used to examine the differences in process understanding and motivation to learn between the VRA and the VOS video groups. The independent variable was the video design variable (coded as a binary variable with VRA video = 1 and VOS video = 0). The dependent variables were retention and transfer (process understanding), as well as the ARCS factors (motivation to learn). To control for the effects of potentially confounding variables, respondents' prior theoretical knowledge and practical experience in warehouse management, their frequency of using learning videos, their frequency of using VR, as well as their age, gender, and employment status were assessed.

4.2 Design of the VRA and VOS videos

The VRA and VOS videos present the same newly digitalized warehouse process of a small and medium-sized glass wholesaler. Both videos demonstrate the storage and retrieval of glassware using a digital warehouse management system and digital handheld scanners to book the glassware into the system. The formal structure of the VRA and VOS videos is the same. First, the title of the respective work process step is mentioned in the audio commentary. Subsequently, the work equipment required for the process step is described using audio commentary and supplementary bullet points. Afterwards, the practical execution of each process step is demonstrated by static graphics in the VOS video and by the animated VR avatar in the VRA video. In both videos, the same audio commentary, encompassing a female human voice, is used. The VRA video was produced using VR technology (HTC-Vive Pro VR-Headset, "Layout and Performance" software provided by Halocline GmbH). For this purpose, a person entered the virtual warehouse environment by means of a Head-Mounted Display (HMD). Within this environment, the person was represented as the VR avatar. Using the teleportation and gripping functions of the VR controllers, the practical execution of future process steps was simulated and documented in the VR environment using the recording function. The VR recordings were then recapitulated using a playback function within the VR software to convert them into a 2D format. Open Broadcaster Software (OBS) was utilized to create a screen recording of the event. The recording was then enriched with an audio commentary that explained the avatar action sequences. The animated VR avatar was designed to resemble an anthropomorphic character (a human-like robot) with human-like movements. The avatar was used to vividly demonstrate the physical execution of operational process steps, such as scanning the barcodes on the glassware using a handheld scanner. The virtual warehouse environment has been designed using a low-fidelity approach with a color scheme based on reality. The prototypical interfaces of the digital warehouse management system and the handheld scanners were presented as overlays depicting detailed warehouse data (e.g., estimated inbound and outbound glassware and information on available storage locations). Figure 1 shows respective excerpts of the VRA video.

The VOS video is the counterpart to the VRA video. It only contains schematic and static two-dimensional graphics in grayscale, including standard geometric shapes and pictograms provided by Microsoft PowerPoint. The graphics and pictograms are presented on white presentation slides. Throughout the video, black arrows are used between static graphics to visualize the dynamics of workflows (e.g., scanning barcodes using the handheld scanner). The interfaces of the warehouse management system and the handheld scanner are depicted schematically, providing an overview of the interface's structure without the incorporation of concrete warehouse data. Figure 2 presents respective excerpts from the VOS video.

4.3 Measures

Process understanding was assessed based on retention and transfer scores achieved by the study participants (Recker and Dreiling, 2011). To measure retention, the participants completed a cloze test (an exercise in which individuals are asked to fill in gaps with terms removed from the text, Taylor, 1953). The content of the cloze test was related to the sequence and physical execution of the warehouse process steps. Participants who correctly filled in more gaps on the test demonstrated superior retention of the warehouse process. The cloze test comprised seven gaps (e.g., "than you scan the [____] of the boxes" with the missing gap being "barcode"). The maximum attainable score for the cloze test was seven points. Participants were awarded one point for each gap filled with a correct term. Half a point was awarded for gaps filled with correct synonyms, and no points were awarded for unfilled gaps or gaps containing incorrect terms. To assess transfer, participants were required to respond to two problem-solving questions.

The answers for the transfer test were not directly presented in the videos but had to be deduced by the participants from

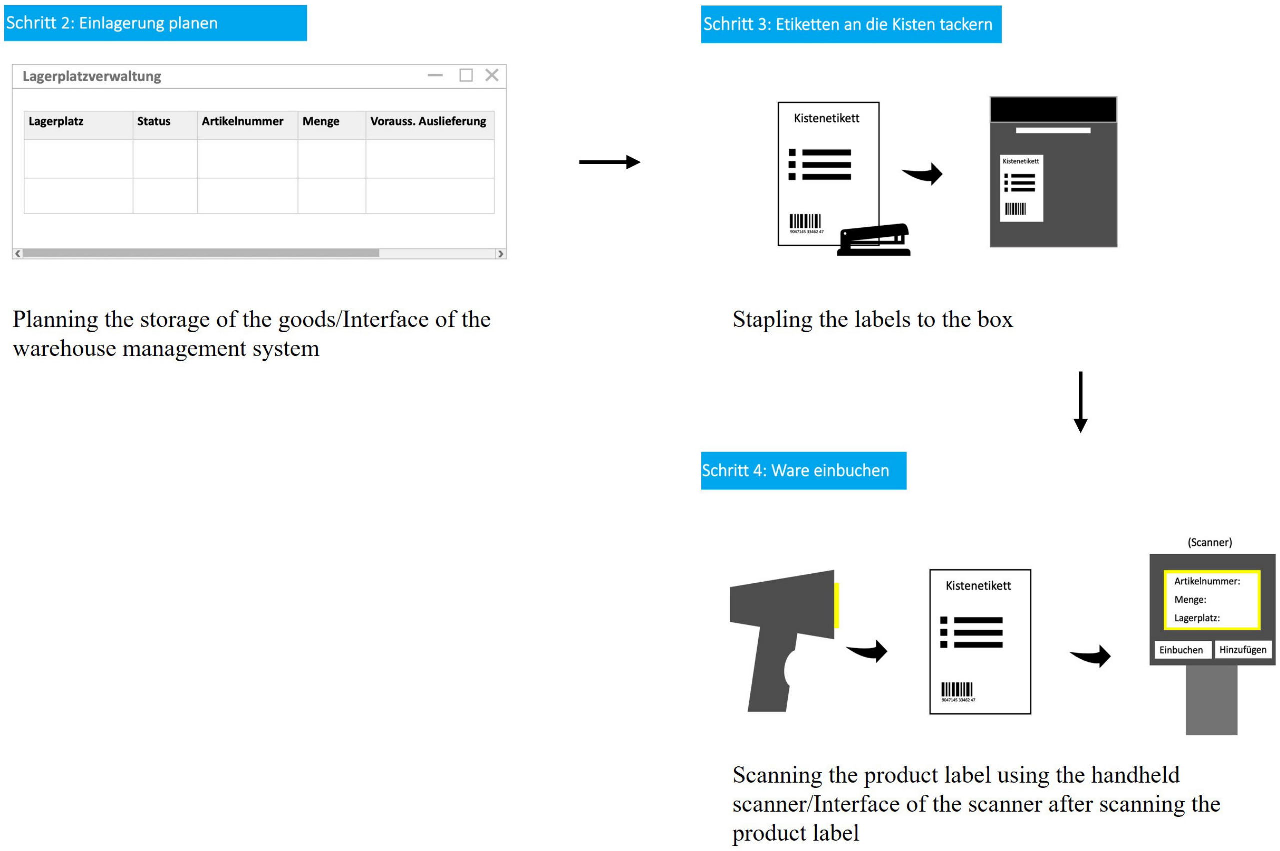

FIGURE 1
Graphics from the VOS video illustrating three example steps in the warehouse process.

the information provided. The first problem-solving question asked how to use the digital warehouse management system to organize the storage of newly arrived glassware when storage space is limited. The second problem-solving question was about how to use the digital warehouse management system to implement a short-term increase in the order quantity of glassware requested by the customer (the problem-solving questions can be found in the Supplementary material). The maximum attainable score for the transfer task was two points. One point was awarded for a plausible answer that included all relevant information from the video. Half a point was given for a partially plausible answer that included some, but not all, of the relevant information from the video. No points were awarded for answers that were implausible. Two independent raters with expertise in industrial and organizational psychology evaluated the retention (cloze test) and transfer tests (problem-solving questions) based on the evaluation criteria described above. Interrater reliability was assessed using Cohen's Kappa (κ). The kappa value for retention (cloze test) was $\kappa = 0.887$, and for transfer (problem-solving questions), the kappa value was $\kappa = 0.704$.

Motivation to learn was assessed using a self-report motivation questionnaire consisting of 16 items. The questionnaire was based on the Instructional Materials Motivation Survey (IMMS) developed by Keller (2010) to measure the ARCS factors. As the items of the IMMS are mainly designed to evaluate the motivational capacity of a face-to-face educational context, they were revised with regard to their application in the context of video-based instruction. To ensure adequate study duration, we only selected the IMMS items relevant to our study (e.g., items related to the impact of the visual design elements). Attention (ARCS-A) was measured with five items (e.g., "the video contained elements that aroused my interest"). The attention scale included two negatively coded items, each recoded for further analysis. Relevance (ARCS-R) was assessed using two items (e.g., "the video provided examples that demonstrated how relevant the content is to real users"). Due to its low discriminatory power, one of the original three items was removed. Confidence (ARCS-C) was assessed using six items (e.g., "numerous video segments contained an overwhelming amount of information, making it challenging to recall the critical points"). Three of the confidence items were recoded (as they were originally stated negatively). Satisfaction (ARCS-S) was measured with three items (e.g., "I enjoyed watching the video"). All ARCS items were assessed using a five-point Likert scale (1 = I strongly disagree to 5 = I strongly agree). Cronbach's alpha values for the ARCS scales were: $\alpha = 0.763$ for attention (ARCS-A), $\alpha = 0.664$ for relevance

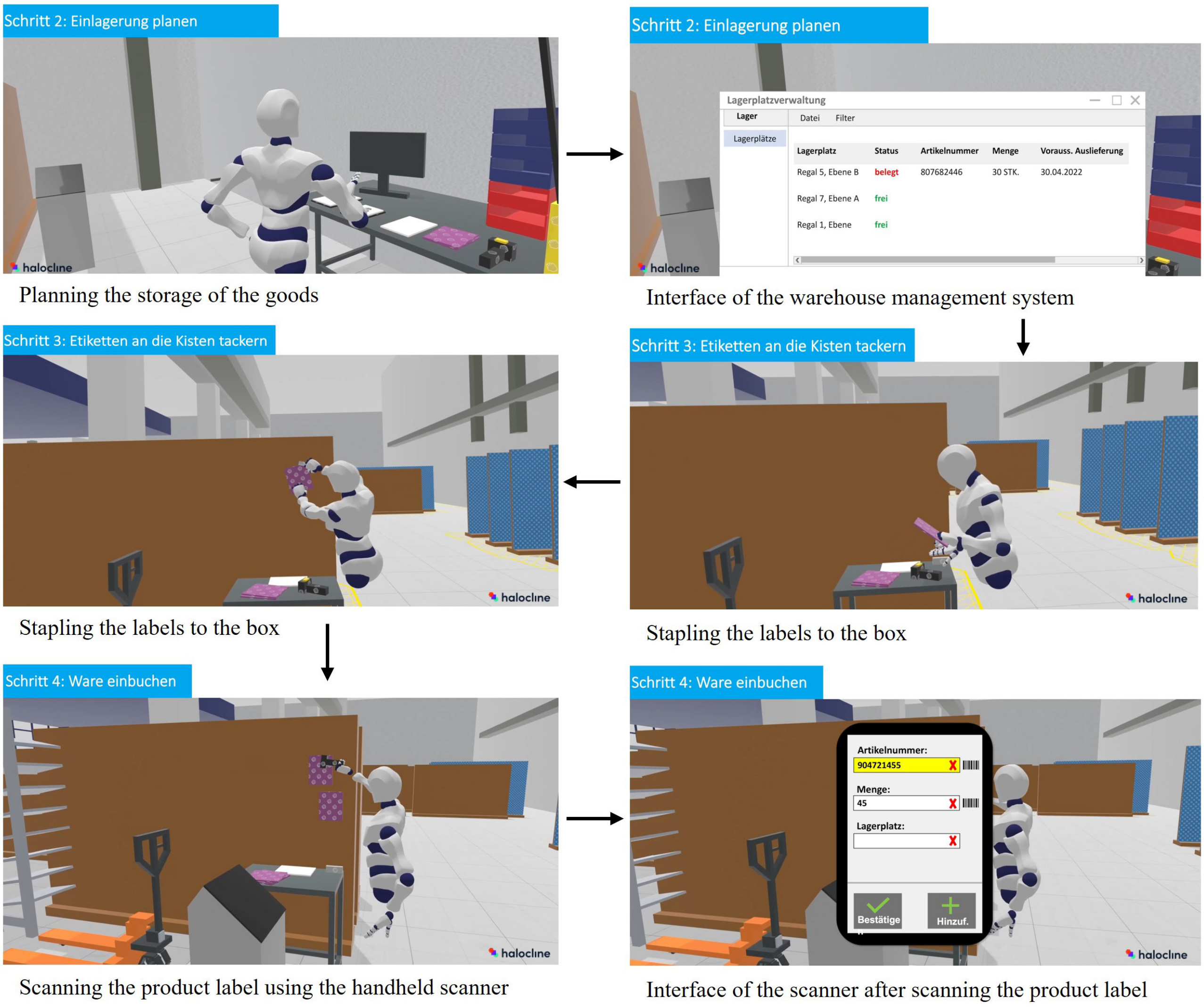

FIGURE 2
Graphics from the VRA video illustrating three example steps in the warehouse process. The virtual environments were created using Planning App "Halocline" from Halocline GmbH & Co. KG; www.halocline.io.

(ARCS-R), $\alpha = 0.823$ for confidence (ARCS-C), and $\alpha = 0.896$ for satisfaction (ARCS-S).

In addition, control variables were measured that potentially affect process understanding and motivation to learn. These variables were prior knowledge and experience in warehouse management, the frequency of use of instructional videos and VR environments, as well as the age, gender, and employment status of the test subjects. Prior theoretical knowledge and practical experience in warehouse management were measured using one question each (e.g., "do you have theoretical knowledge in warehouse management?"). Both questions could be answered using a binary response scale (0 = no; 1 = yes). The frequency of using learning videos and VR was measured on a five-point Likert scale (1 = never to 5 = regular). Age was assessed as a metric variable using a free-text field. Gender was measured using three answer options (1 = male, 2 = female, 3 = diverse). However, as no participant selected the option "diverse," gender was recoded into a binary variable (0 = female, 1 = male). Employment status was assessed using a binary answer scale (0 = student, 1 = employee in a German enterprise).

4.4 Procedures

An online survey was conducted, with participants recruited from two German universities and different German companies. Students were recruited via access to university courses (e.g., seminars). Employees were contacted personally or through social media. The survey link could be accessed on a PC or other mobile devices (e.g., a mobile phone or tablet). Therefore, the VRA and VOS videos were watched on a conventional 2D screen.

Before starting the online questionnaire, participants provided written informed consent to participate in this study[1]. Upon accessing the survey link, participants were randomly assigned to either the VRA or the VOS video group. Participants were instructed to watch the video twice in succession. This was to ensure that they were able to obtain all the relevant information presented. Participants were instructed not to take notes while watching the videos. The duration of both videos was ∼9 min. Subsequent to watching the videos, the participants completed the cloze test, answered the two problem-solving questions, and completed the motivation questionnaire. The total duration of the study was approximately 35 min on average. Upon successful completion of the study, employees were eligible to receive a financial incentive of 10 euros. Students had the opportunity to earn bonus points for their exams or test subject credits. The statistical analyses were carried out using SPSS 29.0. One-way ANOVAs were conducted to evaluate the differences between the VRA and VOS video groups regarding retention and transfer scores (process understanding), as well as ARCS scores (motivation to learn).

5 Results

5.1 Participants

G*Power indicated that a one-way ANOVA with a sample of $N = 128$ participants across two conditions would be sensitive to the effects of $f = 0.25$ with 80 % power (alpha = 0.05). In this study, a total sample size of $N = 121$ participants could be attained (76 % female; $M_{age} = 24.5$ years, $SD = 6.5$). A total of 62 individuals viewed the VRA video, while 59 individuals viewed the VOS video. The sample comprised 73.6 % students (mainly psychology and economics) from two German universities and 26.4 % employees in German enterprises. With respect to the total sample, 11.6 % of the participants had prior practical experience in warehouse management, and 14.9 % had prior theoretical knowledge about warehouse management. Concerning the frequency of using learning videos, 6.6 % of the participants never used learning videos, 23.1 % of the participants rarely used learning videos, 43.0 % of the participants occasionally used learning videos, 19 % of the participants frequently used learning videos, and 8.3 % of the participants even regularly used learning videos. Regarding the frequency of using VR, 77.7 % of the participants reported never using VR, 17.4 % of the participants reported rarely using VR, 2.5 % of the participants reported occasionally using VR, and 2.5 % of the participants reported frequently using VR.

5.2 Descriptive results

The mean scores achieved in the VRA video group for retention and transfer, as well as for attention (ARCS-A), relevance (ARCS-R), confidence (ARCS-C), and satisfaction (ARCS-S), were higher than in the VOS video group (see Table 1). The mean values of the relevance (ARCS-R) and confidence (ARCS-C) factors were the highest in both test groups (relevance: $M_{VRA} = 3.750$ and $M_{VOS} = 3.093$ and confidence: $M_{VRA} = 3.758$ and $M_{VOS} = 3.489$). The mean value of the attention factor (ARCS-A) was moderate in both groups ($M_{VRA} = 3.107$ and $M_{VOS} = 2.661$) and the mean value of the satisfaction factor (ARCS-S) was the lowest in both groups ($M_{VRA} = 2.441$ and $M_{VOS} = 2.034$).

TABLE 1 Means and standard errors for all variables of the VRA and the VOS video group.

Variable	VOS ($N = 59$) M (SD)	VRA ($N = 62$) M (SD)
Retention	4.873 (1.379)	5.081 (1.446)
Transfer	1.585 (0.617)	1.831 (0.384)
Attention	2.661 (0.836)	3.107 (0.891)
Relevance	3.093 (1.12)	3.750 (0.927)
Confidence	3.489 (0.786)	3.758 (0.743)
Satisfaction	2.034 (0.894)	2.441 (1.14)

Maximum retention score = 7 points. Maximum transfer score = 2 points. ARCS factors are assessed with a five-point Likert scale 1 = I strongly disagree to 5 = I strongly agree. VOS, voice-over slides video group; VRA, VR-based avatar video group.

TABLE 2 Mann-Whitney-U Test results.

Variable	U	Z	p
Age	1,734.5	−0.492	0.623
Frequency of using learning videos	1,794	−0.191	0.848
Frequency of using VR	1,661.5	−1.198	0.231

5.3 One-way ANOVA

Before conducting one-way ANOVAs, it was analyzed whether the control variables needed to be considered in further analysis. Both Chi-squared Test and Mann-Whitney-*U* Test showed that there were no statistically significant differences in the distributions of the control variables (see Tables 2, 3). Therefore, these variables were not included in the subsequent analysis.

Afterwards, the hypotheses (H1a-b, and H2a-d) were tested using one-way ANOVAs. First, it was found that the VRA video group achieved significantly higher transfer scores than the VOS video group, thus confirming hypothesis H1b (see Table 4). However, the effect size was relatively small. Second, there were no significant differences in retention scores. Therefore, Hypothesis 1a has to be rejected.

Furthermore, it was revealed that the VRA video group had significantly higher scores in attention (ARCS-A), relevance (ARCS-R), and satisfaction (ARCS-S). There were middle effect sizes for attention (ARCS-A) and relevance (ARCS-R) and a small effect size for satisfaction (ARCS-S). No significant group difference was found about the confidence factor (ARCS-C). Accordingly,

1 Before the participants could take part in the survey, they had to agree to a privacy policy in accordance with the DSGVO.

TABLE 3 Chi-squared Test results.

Variable	χ^2	df	p
Gender	0.134	1	0.714
Prior practical experience in warehouse management	0.010	1	0.921
Prior theoretical knowledge about warehouse management	0.391	1	0.532
Employment status	0.062	1	0.804

TABLE 4 One way ANOVA results.

Variable	$F_{(1,119)}$	p	η^2
Retention	0.653	0.421	0.005
Transfer	7.006	0.009**	0.056
Attention	8.017	0.005**	0.063
Relevance	12.404	<0.001**	0.094
Confidence	3.757	0.055	0.031
Satisfaction	4.742	0.031*	0.038

*$p < 0.05$, **$p \leq 0.01$.

hypotheses 2a, 2b, and 2d were confirmed, and hypothesis 2c was rejected.

6 Discussion

The purpose of the present study was to compare the potential of a VRA video to a VOS video in enhancing employees' understanding of work processes (retention and transfer) and increasing their motivation to learn (ARCS factors).

6.1 The effects of the VRA and VOS videos on process understanding

A central result was that transfer scores were found to be significantly higher in the VRA video group than in the VOS video group, yet retention scores were not. It seems that the VRA video enabled subjects to apply process understanding to problem situations. However, it cannot be concluded that the VRA video supported them in focusing on specific details required to perform the retention test. From the perspective of contextualized learning, the superior transfer scores of the VRA video group may be attributed to learners' ability to relate the learning content to real-world concepts. This suggests that prior implicit knowledge necessary for problem-solving may have been activated, thereby facilitating transfer skills (cf. Guo et al., 2013). Another potential explanation for the enhanced transfer performance of the VRA video group is the presentation of detailed interface overlays in the digital warehouse system and the handheld scanners. The overlays depicted detailed information (e.g., expected deliveries and retrievals of glassware) that may have supported the learners to respond more efficaciously to the problem-solving questions.

6.2 The effects of the VRA and VOS videos on motivation to learn

The significantly higher scores of attention (ARCS-A), relevance (ARCS-R), and satisfaction (ARCS-S) in the VRA video group can be attributed to the special visualizations in the VRA video. Based on the study by Jong (2023), it can be deduced that the significantly higher level of attention (ARCS-A) may be due to the authentic presentation of the replicated virtual 3D warehouse, which aroused the interest and curiosity of the learners. Furthermore, in accordance with the findings of Chin et al. (2016) as well as Dinçer and Doganay (2017), it may be concluded that the anthropomorphic appearance and the human-like body movements of the VR avatar contributed to the increased learners' attention (ARCS-A; cf. Chin et al., 2016; Dinçer and Doganay, 2017). The significantly higher scores of the relevance factor (ARCS-R) in the VRA video group can be attributed to the virtually replicated 3D warehouse, which creates a relation to the real-world application context (cf. Parchmann and Kuhn, 2018; Schmid, 2023). As the practical execution of the process steps was demonstrated by the anthropomorphic VR avatar, it can be further inferred that the avatar was perceived as a social interaction partner (Mayer, 2014). Consequently, the process information conveyed by the avatar was automatically considered more relevant (cf. Stiller et al., 2020). As indicated above, no significant differences were observed between the VRA and VOS video groups in terms of learners' confidence in their ability to learn (ARCS-C). However, the differences may have become significant when the sample size was increased to the required 128 test subjects. Finally, it can be assumed that the visually appealing and interesting design of the virtual 3D warehouse environment and the VR avatar enhanced the level of enjoyment experienced by learners during process training, consequently leading to a substantial increase in satisfaction scores (ARCS-S). Nevertheless, despite these superior scores, the level of satisfaction in the VRA video group was also relatively low ($M = 2.441$). Consequently, it is imperative to implement additional enhancements to the VRA video to improve learner satisfaction. According to the ARCS model, learner engagement in the learning process is a prerequisite for satisfaction and can, for instance, be achieved with exercises that actively involve them in the process of learning (cf. Zander and Heidig, 2019).

7 Implications for science and practice

This study extends current VR-based approaches to process training using VR environments (e.g., Aysolmaz et al., 2016; Leyer et al., 2019, 2021) with an alternative approach using VR-based avatar videos. The study posits that VRA videos can serve as a flexible medium for the effective communication of novel (digitalized) work processes to employees. In addition to the learners' ability to transfer the acquired process understanding, VRA videos offer the potential to increase their attention (ARCS-A), perceived relevance (ARCS-R) and satisfaction with the learning experience (ARCS-S). Despite these advantages, future research should empirically investigate whether these positive or even better learning and motivational effects can be achieved using immersive VR environments.

This study provides practitioners with preliminary insights into the potential of VRA videos to enhance employees' comprehension of work processes and to boost their motivation to learn. The utilization of VRA videos demonstrates several advantages over VR environments, primarily due to their flexibility and accessibility. Moreover, in contrast to the utilization of immersive VR environments, there is no requirement for extensive introductory training in the use of hardware and software. However, it is important to note that the creation of VRA videos also requires a certain level of expertise in the use of VR technologies. This includes, for example, the creation of a virtual work environment and dynamic VR avatar simulations. Consequently, the employment of VRA videos may prove particularly advantageous for organizations that are already utilizing VR for human resources or business process management. In this respect, organizations can leverage their expertise and experience in using VR environments

8 Limitations and future directions

This study is not without its limitations. First, a major limitation is that the effects of the VR avatar and the virtually replicated warehouse environment on retention and transfer, as well as the ARCS factors, were not disentangled. Accordingly, future research should carry out further comparative studies that differentiate between those effects. Second, the present study only includes one point of measurement. Therefore, it was not possible to investigate increases or decreases in retention and transfer or the ARCS factors during the process of learning. The decision to focus on one point of measurement was made since participation in the study was time consuming (∼35 min). This would have made it difficult to recruit the participants again at a second point of measurement. Nevertheless, future research should conduct a study using a pre-post-design. Third, this study is limited by the use of a fictitious experimental setting, which reduces the generalizability of the findings. Instead of including "real" warehouse employees as participants, students or employees from various other work areas were asked to participate. Fourth, the applicability of the study results is limited to work process scenarios involving simple manual activities that can be easily taught using VRA videos (e.g., simple warehouse processes, simple quality controls, processes for series production, simple sorting of products and packaging of goods). However, complex manual tasks, such as operating technically sophisticated tools or performing difficult assembly tasks, may be more effectively trained in immersive VR environments (e.g., Eversberg et al., 2021; Tichon and Scott, 2019). Finally, it should be noted that the VRA video (in contrast to the grayscale VOS video) contained different colors, but the present study did not consider potential color blindness of the test subjects. Consequently, subsequent research should incorporate this potential confounding variable.

9 Conclusions

In the context of digital transformation, where work processes are subject to constant change, it becomes imperative to implement effective training approaches that foster process understanding (e.g., Leyer et al., 2021). The present study demonstrates that a VRA video—in comparison to a VOS video—offers particular potential to facilitate employees' process understanding in terms of transfer. Furthermore, the VRA video was found to be more effective in increasing motivational factors, in particular employees' attention (ARCS-A), their perceived relevance of the learning content (ARCS-R), and their satisfaction with the learning experience (ARCS-S). Consequently, VRA videos emerge as a cost-effective and flexible alternative to immersive VR environments.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements.

Author contributions

SD: Writing – original draft, Writing – review & editing. NS: Writing – review & editing. MS: Writing – review & editing. J-PS: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is part of the research project SoDigital and was funded by the German Federal Ministry of Education and Research, co-funded by the European Social Fund under Grant [02L18B571].

Acknowledgments

We would like to thank all participants of this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2025.1553441/full#supplementary-material

References

Aysolmaz, B., Brown, R., Bruza, P., and Reijers, H. A. (2016). "A 3D visualization approach for process training in office environments," in *On the Move to Meaningful Internet Systems: OTM 2016 Conferences: Confederated International Conferences: CoopIS, CandTC, and ODBASE 2016, Rhodes, Greece, October 24–28, 2016, Proceedings,* eds. C. Debruyne et al. (Springer: Cham), 418–443. doi: 10.1007/978-3-319-48472-3_24

Burton-Jones, A., and Meso, P. (2008). The effects of decomposition quality and multiple forms of information on novices' understanding of a domain from a conceptual model. *J. Assoc. Inf. Syst.* 9, 748–802. doi: 10.17705/1jais.00179

Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., and Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories: a meta-analysis. *Educ. Psychol. Rev.* 33, 989–1015. doi: 10.1007/s10648-020-09587-1

Chen, M. P., Wang, L. C., Zou, D., Lin, S. Y., and Xie, H. (2019). Effects of caption and gender on junior high students' EFL learning from iMap-enhanced contextualized learning. *Comput. Educ.* 140:103602. doi: 10.1016/j.compedu.2019.103602

Chin, K.-Y., Hong, Z.-W., Huang, Y.-M., Shen, W.-W., and Lin, J.-M. (2016). Courseware development with animated agents in the learning system to improve learning motivation. *Interact. Learn. Environ.* 24, 360–381. doi: 10.1080/10494820.2013.851089

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: a meta-analysis. *Educ. Res. Rev.* 24, 193–209. doi: 10.1016/j.edurev.2018.05.002

Dinçer, S., and Doganay, A. (2017). The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. *Comput. Educ.* 111, 74–100. doi: 10.1016/j.compedu.2017.04.005

Eversberg, L., Grosenick, P., Meusel, M., and Lambrecht, J. (2021). "An industrial assistance system with manual assembly step recognition in virtual reality," in *2021 International Conference on Applied Artificial Intelligence (ICAPAI)* (Halden: IEEE), 1–6. doi: 10.1109/ICAPAI49758.2021.9462061

Grassini, S., Laumann, K., and Rasmussen Skogstad, M. (2020). The use of virtual reality alone does not promote training performance (but sense of presence does). *Front. Psychol.* 11:1743. doi: 10.3389/fpsyg.2020.01743

Guo, H., Brown, R., and Rasmussen, R. (2013). "A theoretical basis für using virtual worlds as a personalised process visualisation approach," in *CAiSE 2013 Workshops, LNBIP, 148*, eds. X. Franch and P. Soffer (Berlin, Heidelberg: Springer), 229–240. doi: 10.1007/978-3-642-38490-5_22

Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learn. Instr.* 14, 343–351. doi: 10.1016/j.learninstruc.2004.06.007

Hirsch-Kreinsen, H., Ten Hompel, M., and Kretschmer, V. (2020). "Digitalisierung industrieller arbeit: entwicklungsperspektiven und gestaltungsansätze," in *Handbuch Industrie 4.0: Band 3: Logistik*, eds. M. ten Hompel, T. Bauernhansl, and B. Vogel-Heuser (Springer: Vieweg), 495–512. doi: 10.1007/978-3-662-58530-6_21

Hoeffler, T. N., and Leutner, D. (2007). Instructional animation versus static pictures: a meta-analysis. *Learn. Instr.* 17, 722–738. doi: 10.1016/j.learninstruc.2007.09.013

Jong, M. S. Y. (2023). Flipped classroom: motivational affordances of spherical video-based immersive virtual reality in support of pre-lecture individual learning in pre-service teacher education. *J. Comput. High. Educ.* 35, 144–165. doi: 10.1007/s12528-022-09334-1

Kathleen, N., Ross, B., and Kriglstein, S. (2014). "Storyboard augmentation of process model grammars for stakeholder communication," in *2014 International Conference on Information Visualization Theory and Applications (IVAPP)* (Portugal: IEEE), 114–121 doi: 10.5220/0004668101140121

Keller, J. M. (2010). *Motivational Design for Learning and Performance: The ARCS Model Approach.* New York, Dordrecht, Heidelberg, London: Springer Science + Business Media. doi: 10.1007/978-1-4419-1250-3

Koese, E., Taşlibeyaz, E., and Karaman, S. (2021). Classification of instructional videos. *Technol. Knowl. Learn.* 26, 1079–1109. doi: 10.1007/s10758-021-09530-5

Leyer, M., Aysolmaz, B., Brown, R., Türkay, S., and Reijers, H. A. (2021). Process training for industrial organisations using 3D environments: an empirical analysis. *Comput. Ind.* 124:103346. doi: 10.1016/j.compind.2020.103346

Leyer, M., Brown, R., Aysolmaz, B., Vanderfeesten, I., and Turetken, O. (2019). "3D virtual world BPMN training systems: process gateway experimental results," in *Advanced Information Systems Engineering. CAiSE 2019. Lecture Notes in Computer Science. vol. 11483,* eds. P. Giorgini and B. Weber (Cham: Springer). doi: 10.1007/978-3-030-21290-2_26

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning.* New York: Cambridge University Press. doi: 10.1017/CBO9780511816819

Mayer, R. E. (2014). *The Cambridge Handbook of Multimedia Learning. 2. Edition.* New York: Cambridge University Press. doi: 10.1017/CBO9781139547369

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *J. Appl. Res. Mem. Cogn.* 10, 229–240. doi: 10.1016/j.jarmac.2021.03.007

Mayer, R. E., and DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *J. Exp. Psychol. Appl.* 18, 239–252. doi: 10.1037/a0028616

Mueller, K., Hamborg, K.-C., Straatmann, T., Schumacher, J.-P., Koßmann, C., Teuteberg, F., et al. (2023). "Sozio-digitale Innovation durch partizipative Prozessgestaltung im virtuellen Raum," in *Digitalisierung der Arbeitswelt im Mittelstand 3,* eds. V. Nitsch, C. Brandl, R. Häußling, P. Roth, T. Gries, and B. Schmenk (Berlin, Heidelberg: Springer), 239–290. doi: 10.1007/978-3-662-67024-8_7

Mueller, K., Straatmann, T., Schumacher, J.-P., and Depenbusch, S. (2022). Virtual reality bei der digitalen neugestaltung von geschäftsprozessen. *Pers. Q.* 74, 34–39.

Parchmann, I., and Kuhn, J. (2018). "Lernen im Kontext," in *Theorien in der naturwissenschafts-didaktischen Forschung*, eds. D. Krüger, and H. Schrecker (Berlin: Springer Spektrum), 193–207. doi: 10.1007/978-3-662-56320-5_12

Peng, T. H., and Wang, T. H. (2022). Developing an analysis framework for studies on pedagogical agent in an e-learning environment. *J. Educ. Comput. Res.* 60, 547–578. doi: 10.1177/07356331211041701

Recker, J., and Dreiling, A. (2011). The effects of content presentation format and user characteristics on novice developers' understanding of process models. *Commun. Assoc. Inf. Syst.* 28, 65–84. doi: 10.17705/1CAIS.02806

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., and Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learn. Instr.* 19, 481–494. doi: 10.1016/j.learninstruc.2008.08.001

Scheiter, K., Richter, J., and Renkl, A. (2020). "Multimediales Lernen: Lehren und Lernen mit Texten und Bildern," in *Handbuch Bildungstechnologie*, eds. H. Niegemann and A. Weinberger (Berlin, Heidelberg: Springer), 31–56. doi: 10.1007/978-3-662-54368-9_4

Schmid, A. (2023). "Authentische Kontexte für MINT-Lernumgebungen. Eine zweiteilige Interventionsstudie in den Fachdidaktiken Physik und Technik," in *Studien zum Physik- und Chemielernen*, eds. M. Hopf and M. Rophol (Berlin: Logos Verlag). doi: 10.30819/5605

Setyowati, R. R., Rochmat, S., and Nugroho, A. N. P. (2023). Virtual reality on contextual learning during covid-19 to improve students' learning outcomes and participation. *Int. J. Instr.* 16, 173–190. doi: 10.29333/iji.2023.16110a

Stiller, K. D., Schworm, S., and Gruber, H. (2020). "Learning with and from illustrations: cognitive, motivational, affective, social and metacognitive processes," in *Challenging the iconic turn – Den Iconic Turn neu denken,* eds. D. E. Delarue, and C. Wagner (Paderborn: Wilhelm Fink Verlag).

Taylor, W. L. (1953). Cloze proecedure: a new tool for measuring readability. *Journal. Q.* 30, 415–433. doi: 10.1177/107769905303000401

Tichon, J., and Scott, S. (2019). Virtual reality manual handling induction training: Impact on hazard identification. *Asia Pac. J. Contemp. Educ. Commun. Technol.* 5, 49–58. doi: 10.25275/apjcectv5i1edu5

Um, E. R., Plass, J. L., Hayward, E. O., and Homer, B. D. (2012). Emotional design in multimedia learning. *J. Educ. Psychol.* 104, 485–498. doi: 10.1037/a0026609

van der Meij, H., van der Meij, J., and Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Educ. Technol. Res. Dev.* 63, 381–403. doi: 10.1007/s11423-015-9378-5

Wang, F., Li, W., Mayer, R. E., and Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: effects on eye-fixations during learning and learning outcomes. *J. Educ. Psychol.* 110, 250–268. doi: 10.1037/edu0000221

Wohlgenannt, I., Simons, A., and Stieglitz, S. (2020). Virtual reality. *Bus. Inf. Syst. Eng.* 62, 455–461. doi: 10.1007/s12599-020-00658-9

Yarden, H., and Yarden, A. (2010). Learning using dynamic and static visualizations: students comprehension, prior knowledge and conceptual status of a biotechnological method. *Res. Sci. Educ.* 40, 375–402. doi: 10.1007/s11165-009-9126-0

Zander, S., and Heidig, S. (2019). "Motivationsdesign bei der Konzeption multimedialer Lernumgebungen," in *Lernen mit Bildungstechnologien*, eds. H. Niegemann, and A. Weinberger (Deutschland: Springer-Verlag GmbH), 1–23. doi: 10.1007/978-3-662-54373-3_37-1