Fachhochschule Nordwestschweiz
Hochschule für Angewandte Psychologie

Exploring the role of learning activity when learning with video and virtual reality: a mixed methods study with airport security officers

MASTER THESIS

2021

Kaspar Kaufmann

Co-Supervisor:
Prof. Dr. Adrian Schwaninger

Co-Supervisor:
Thomas Wyssenbach

Institute Humans in Complex Systems
School of Applied Psychology
University of Applied Sciences and Arts Northwestern Switzerland FHNW
Olten, Switzerland

Abstract

The present study explored how media (video vs. virtual reality) and learning activity (passive vs. interactive) affect airport security screeners' learning experiences by applying a 2 x 2 factorial between-subjects design. A mixed methods approach was employed to assess the screeners' (n = 26) learning, cognitive load, intrinsic motivation, and technology acceptance. Results showed that videos led to slightly higher learning outcomes than virtual reality. While screeners believed interactivity to enhance learning, no main effect was discovered. This result may have been influenced by increased cognitive load experienced by the screeners through interactivity. Intrinsic motivation was significantly higher for screeners learning with interactive video, passive virtual reality, and interactive virtual reality compared to passive video. Regarding technology acceptance, screeners perceived virtual reality and interactivity to be more useful than video and passivity, respectively. Overall, this study offers insight into the potentials of multimedia for learning in a practical setting.

Keywords: virtual reality, video, multimedia learning, cognitive load, intrinsic motivation, technology acceptance

Zusammenfassung

Diese Studie untersuchte, wie Medien (Video vs. virtuelle Realität) und Lernaktivität (passiv vs. interaktiv) die Lernerfahrungen von Flughafensicherheitsbeauftragten (Screener) beeinflussen, indem ein 2 x 2-faktorielles between-subjects Design angewendet wurde. Mit einem Mixed-Methods-Ansatz wurde das Lernen, die kognitive Belastung, die intrinsische Motivation und die Technologieakzeptanz der Screener (n = 26) untersucht. Die Ergebnisse zeigten, dass Videos zu leicht höheren Lernergebnissen führten als virtuelle Realität. Obwohl die Screener glaubten, dass Interaktivität das Lernen unterstützt, wurde kein statistischer Haupteffekt gefunden. Dieses Ergebnis wurde möglicherweise durch erhöhte kognitive Belastung beeinflusst. Die intrinsische Motivation war signifikant höher bei Screenern, die mit interaktiven Videos, passiver virtueller Realität und interaktiver virtueller Realität lernten als mit passiven Videos. Hinsichtlich der Technologieakzeptanz empfanden die Screener die virtuelle Realität und die Interaktivität als nützlicher als Videos bzw. Passivität. Insgesamt bietet diese Studie einen Einblick in die Potenziale von multimedialem Lernen in einem praxisbezogenen Kontext.

Keywords: virtuelle Realität, Video, multimediales Lernen, kognitive Belastung, intrinsische Motivation, Technologieakzeptanz

Exploring the role of learning activity when learning with video and virtual reality: a mixed methods study with airport security officers

With one technological innovation following on the heels of another, possibilities of instructional media are steadily evolving. For videos, improved and simplified recording, editing, and broadcasting have already led to the mainstream adoption of the medium in formal and informal learning environments (de Koning et al., 2018). For immersive virtual reality (VR), on the other hand, the prospects of the novel technology are still being explored as it has only recently reached consumer-level affordability (Rupp et al., 2019). While both video and VR offer unique affordances, they assume the same multimedia principle of using words and pictures for instruction (Mayer, 2009).

An understudied facet of multimedia learning has been the role of interactivity which promises enhanced learning effectiveness (Evans & Gibbons, 2007). However, allowing more learning activity in a rich virtual environment (VE) of VR may also overwhelm learners and distract from the task at hand, thus leading to higher cognitive load (e.g., Knight & Tlauka, 2017; Makransky, Terkildsen, et al., 2019; Parong & Mayer, 2018). Apart from cognitive aspects, motivational factors and learners' acceptance of the medium play an important role in understanding the utility and impact of the medium when it is applied in an educational context. It has been shown that learners generally experience more intrinsic motivation when an activity is interesting or enjoyable (Ryan & Deci, 2000). The higher intrinsic motivation may consequently improve learning outcomes (Ryan & Deci, 2009) and working memory capacity (Schnotz & Kürschner, 2007). Further, learners' technology acceptance is vital for successfully implementing any learning medium (Donkor, 2011).

Therefore, the present study explores media (video vs. VR) and learning activity (passive vs. interactive) using a 2 x 2 factorial design. By assessing learning outcomes, cognitive load, intrinsic motivation, and technology acceptance in a practical setting, valuable insights into the use of passive and interactive video and VR are offered.

Video and VR for learning

The use of videos in educational settings has increased considerably in the past decades, making videos one of today's most prevalent teaching and learning methods (de Koning et al., 2018). Due to the wide variety of video types and platforms available, learners of all educational levels and settings can access videos (de Koning et al., 2018). In formal learning environments, videos have recently found use in combination with other forms of instruction. Thus, videos serve as a key component for blended learning (e.g., Coyne et al., 2018), e-learning (e.g., Zhang et al., 2006), massive online open courses (e.g., Watson et al., 2017), and flipped classrooms (e.g., DeLozier & Rhodes, 2017). Accordingly, interest in instructional video research has been rekindled, joining the topics of animation, simulation and VR, which have dominated educational research in recent years (Bétrancourt & Benetos, 2018).

VR is defined as "a computer-mediated simulation that is three-dimensional, multisensory, and interactive, so that the user's experience is 'as if' inhabiting and acting within an external environment" (Burbules, 2006, p. 37). Typically, VR is differentiated into immersive and non-immersive VR (Parong & Mayer, 2018). Immersion is an objective measure dependent on the sensory fidelity of a VR system and the extent to which it shuts out the outside world (Cummings & Bailenson, 2016). For the present study, VR refers to immersive VR which differs from non-immersive VR in hardware used to display the virtual environment (VE; Meyer et al., 2019). Immersive VR is commonly accessed by a head-mounted display (HMD) using two screens close to the eyes (Makransky & Lilleholt, 2018). In contrast, non-immersive VR typically refers to the VE being projected on a computer screen (Lee et al., 2010).

By learning with a HMD, students and trainees are entirely surrounded by the VE, which offers a realistic and lifelike experience (Makransky, Borre-Gude, et al., 2019). This sense of being in the virtual world promises a unique advantage for educational and training purposes and is commonly referred to as presence in VR literature (e.g., Witmer & Singer, 1998). Immersion and its effect on presence have been researched extensively (see Cummings & Bailenson, 2016, for an overview) and VR has consistently shown higher presence than desktop-based learning, such as videos (e.g., Makransky, Terkildsen, et al., 2019; Ulrich et al., 2019).

The role of learning activity

The importance of learning activity in instructional media can be traced back to the notion that learners have to become actively engaged for deep learning to occur (e.g., Mayer, 2009; Renkl et al., 2007; Wittrock, 1991). Although interactivity has played a core role in educational literature, it remains an elusive construct (Mcmillan & Hwang, 2002). Different studies have yielded conflicting results regarding the effect of interactivity on learning (Domagk et al., 2010). Domagk et al. (2010) argued that these ambiguities might reflect diverging definitions of interactivity. In response, they proposed a definition for multimedia interactivity aimed at encompassing shared ideas from different disciplines: "Interactivity in the context of computer-based multimedia learning is reciprocal activity between a learner and a multimedia learning system, in which the (re)action of the learner is dependent upon the (re)action of the system and vice versa" (Domagk et al., 2010, p. 1052). Based on this definition, the present study assesses learning activity by comparing interactive with passive conditions. While the interactive conditions allow learners to control the pace and certain aspects of the multimedia learning system through manipulation (e.g., Moreno & Mayer, 2007), passive (or guided; e.g., Roussou & Slater, 2017) conditions are controlled by the system (e.g., Evans & Gibbons, 2007).

Theoretical background

Multimedia learning

Multimedia learning refers to knowledge acquisition from instruction containing words (e.g., narration, on-screen text) and pictures in static (e.g., illustrations, diagrams) or dynamic (e.g., video, animation) form (Mayer, 2012). The rationale for multimedia instruction, such as video and VR, is that people learn better from words and pictures combined than from words alone (Butcher, 2014; Mayer, 2017). This principle is one of several empirically-based design principles of the Cognitive Theory of Multimedia Learning (CTML) aiming to enhance learning (Mayer & Pilegard, 2014).

While many studies have found instructional videos to be more effective for learning compared to traditional educational methods (e.g., Calandra et al., 2006; Kay & Edwards, 2012; Lin & Tseng, 2012; Santagata, 2009), other scientific works did not find any improvements in learning performance (e.g., Donkor, 2010; Lindgren et al., 2007). Nevertheless, Yousef et al. (2014) stated in their meta-analysis of video-based learning that there is an agreement among researchers that videos have the potential to improve learning outcome when combined with appropriate pedagogical methods. Surprisingly, interactive videos have garnered less scientific interest with only few studies exploring their effect on learning (Giannakos, 2013; Giannakos et al., 2014). Zhang et al. (2006) found positive impacts on learning outcomes when participants could jump directly to any part of the instructional video. However, a recent study found that control over pace did not affect students' learning outcomes when comparing interactive and non-interactive videos (Biard et al., 2018).

Regarding VR, there have been several systematic reviews in recent years exploring the relationship between VR and learning. A recent meta-analysis by Hamilton et al. (2020) reported that around half of the 29 reviewed papers demonstrated a positive effect on learning when using immersive VR over less immersive pedagogical methods. The review indicated that highly complex or conceptual problems requiring spatial understanding and visualisation might benefit the most from VR (see also Jensen & Konradsen, 2018; Wu et al., 2020, for further reviews). However, none of the meta-analyses looked specifically at the role of interactivity. Therefore, similar to videos, indicating that research focusing on the interactive component of VR has been relatively scarce. Notably, Zhang et al. (2019) investigated the effect of three different levels of interactivity (low, medium, high) on objective and subjective learning of immunology concepts. Results showed no evidence of increasing levels of interactivity affecting learning outcomes, even though subjective learning results suggested otherwise (Zhang et al., 2019).

When comparing video and VR, Allcoat and von Mühlenen (2018) showed overall better learning of biology knowledge with VR while using the same instructional visuals. They attributed the better VR scores to either the immersion or the VR environment’s interactivity and suggested further studies comparing VR with other active learning methods. However, in another recent study by Meyer et al. (2019) participants learning with video scored significantly higher in the knowledge retention test than those using VR. Yet, after a post-test delayed by one week, no more differences in knowledge were found between the two media(Meyer et al., 2019). Therefore, research has yet to offer conclusive evidence on the superior media for learning.

Cognitive load

Consistent with CTML, the cognitive load theory (CLT) proposes that human cognitive processing is heavily constrained by limited working memory, inhibiting learning when cognitive processing exceeds the learners' capacity (Sweller et al., 2011). According to CLT, the working memory can only process a limited number of information elements at a time while the long-term memory is limitless (Sweller et al., 2011). Traditionally, cognitive load is differentiated into three independent sources: intrinsic, extraneous and germane cognitive load (e.g., Sweller et al., 1998).

Intrinsic cognitive load is associated with the intrinsic nature of instructional material and thereby determined by both the complexity of the information and the knowledge of the person processing that information (Sweller et al., 2019). Therefore, intrinsic cognitive load generally can only be influenced by altering the quantity of information or its complexity.

Extraneous cognitive load is determined by how the learning material is presented and what activities the learner needs to perform during the learning task (Sweller et al., 2019). Therefore, instructional material should aim to minimise extraneous cognitive load by avoiding design elements that distract the learner and hamper the learning process.

Germane cognitive load emerges during the formation and regulation of mental models, thereby facilitating learning and contributing to transfer performance (Paas et al., 2003). However, learners can only devote resources to germane cognitive load if extraneous cognitive load does not exceed their working memory capacity.

Based on the given description, no conclusive statements or predictions can be made about the cognitive load induced by media or learning activity, as the instructional material and its presentation may both influence cognitive processes. Nevertheless, Makransky and Lilleholt

(2018) have hypothesised that immersive VR simulations could foster generative processing and therefore germane cognitive load by providing a highly realistic experience. On the other hand, researchers have suggested that the rich VE and high-fidelity graphics of VR could distract learners while increasing cognitive load, thereby possibly diverting the learner from the task at hand (e.g., Makransky, Terkildsen, et al., 2019; Parong & Mayer, 2018). Even with equivalent graphics and animations, VR could similarly increase extraneous cognitive load compared to video by too much interaction, leaving less working memory for learning processes (e.g., Zhang et al., 2019)

Intrinsic Motivation

Intrinsic motivation refers to doing an activity for its inherent satisfaction, as opposed to external products, pressures, or rewards (Ryan & Deci, 2000). Ryan and Deci (2000) argue that for intrinsic motivation to occur, an activity must hold intrinsic interest for the learner, appear novel or challenging, or hold aesthetic value. Additionally, the concept of intrinsic motivation has been used as a measure for enjoyment, liking and curiosity (Lepper et al., 2005). Intrinsic motivation entails both personal and situational interest (Linnenbrink & Pintrich, 2002). For multimedia learning, this means that the learning activity as wells as the learning environment is pertinent. Several studies have shown a positive effect of intrinsic motivation on learning outcomes in educational contexts (Ryan & Deci, 2009). Further, motivation can impact cognitive load. For instance, Schnotz and Kürschner (2007) have shown that high motivation can temporarily increase working memory capacity.

Surprisingly, a meta-study by Mutlu-Bayraktar et al. (2019) on the cognitive load in multimedia learning showed only a few studies that have investigated motivation in the past. There is empirical evidence for videos suggesting that interactivity positively affects emotional and motivational factors (e.g., Nikopoulou-Smyrni & Nikopoulos, 2010). Similar results have been found for VR, with recent studies reporting positive motivational outcomes when compared to less immersive instruction (e.g., Makransky, Borre-Gude, et al., 2019; Makransky & Lilleholt, 2018). However, Makransky and Lilleholt (2018) stated that there is still limited empirical evidence of how much value immersive VR holds.

Technology acceptance

The successful implementation of any new learning medium depends on learners' acceptance and willingness to adopt it (Donkor, 2011; Zhang et al., 2006). In order to assess the attitudes and acceptance of learners towards multimedia, Davis (1989) offers a theory-based approach with the technology acceptance model (TAM). TAM is grounded on the theory of reasoned action (Ajzen & Fishbein, 1977) and argues that the decision to accept or reject a system is influenced by two major determinants (Davis, 1989). The first is perceived usefulness which is the degree to which a person believes that a system will help them perform better at their tasks (Davis, 1989). The second is perceived ease of use which is one's beliefs about the effort needed to use the system. Together, these determinants predict an individual's attitude towards using a system called behavioural intention (Davis, 1989).

Recent studies have predominantly focused on the acceptance of instructional videos within learning platforms such as e-learning and have generally identified positive attitudes towards videos within those systems (e.g., Liu et al., 2009; Song & Kong, 2017). In the case of VR, several studies investigated which variables influence an individual's attitude towards using this technology. Notably, Chen et al. (2012) and Huang and Liaw (2018) showed that perceived usefulness directly and positively impacts the intention of students to use VR. This finding led Chen et al. (2012) to the conclusion that VR improves the educational quality and facilitates effective learning. Contradictory findings concerning the effect of perceived usefulness on behavioural intention by Lee et al. (2019), however, indicate that context may have a great influence on VR perceptions. Further, perceived ease of use, learning motivation and enjoyment have all shown to positively affect a learners intention to use VR (Chen et al., 2012; Huang & Liaw, 2018).

Context of the study

The present study is embedded in the research project Systematic Threat Assessment, New Standards, Learning Technology Research - Transfer into Practice (STA^{2}RT) by the Center for Adaptive Security Research and Applications in cooperation with the University of Applied Sciences Northwestern Switzerland. STA^{2}RT focuses in part on X-ray screening of cabin or carry-on baggage at airports. Passenger baggage screening is conducted to prevent terrorist attacks and other unlawful interference against civil aviation. So far, predominantly 2D X-ray imaging systems were used for cabin baggage screening. With novel 3D imaging technology based on computer tomography (CT) set to replace the current X-ray systems, airports face the challenge of preparing airport security officers (screeners) through knowledge building and training. For this purpose, a multimedia lesson was developed in cooperation with training personnel of an airport, focusing on supplementing traditional training protocols with context-specific 3D CT learning material experienced through multimedia. This setting offers great potential to empirically study and compare media (video vs. VR) and learning activities (passive vs. interactive) using identical learning material in a setting with high practical significance.

A quantitative study was scheduled for spring 2020 with a large screener sample at an international airport. However, the study had to be postponed indefinitely as a consequence of the ongoing COVID-19 pandemic. In lieu thereof, cooperation with two smaller Swiss airports was established. Additionally, the study was enlarged by a qualitative part focusing on the screeners' perception of the administered media and learning activity. Thus, the flexibility and small sample of screeners at the cooperating airports was put to an advantage.

Purpose of the study and research questions

Regarding the context and theoretical background, the purpose of the present study is to explore how media (video vs. VR) and learning activity (passive vs. interactive) affect screeners' learning, cognitive load, intrinsic motivation, and technology acceptance. Thereby, it provides preliminary findings on the use of multimedia learning in airport settings and offers important contributions to further studies conducted in the STA^{2}RT research project. The following research questions are addressed in this study:

1. How do media (video vs. VR) and learning activity (passive vs. interactive) affect learning outcomes, cognitive load, intrinsic motivation, and technology acceptance of airport security officers?
2. How do airport security officers perceive learning with passive video, interactive video, passive VR, and interactive VR?

In order to answer these research questions, this study employed a convergent mixed methods design (Creswell & Clark, 2017) Convergent designs intend to obtain different but complementary data on the same topic, thus gaining a better understanding of the research problem (Morse, 1991). This approach entails separate quantitative and qualitative data collection and analysis, thereby ensuring that the methods do not influence each other (Creswell & Clark, 2017). The results are then matched and compared in the discussion. This side-by-side approach allows identifying similarities and contradictions in the data and adds to a more complete understanding (Creswell & Clark, 2017).

Method

Participants

The study was carried out with 26 screeners from two airports (Airport A: $n = 12$; Airport B: $n = 14$) in the German-speaking part of Switzerland. All except for one participant screened mainly cabin baggage. However, due to the airports' small size, all participants also performed hold baggage and staff screening tasks. The participants' mean age was 47.31 years ($SD = 12.81$; Airport A: $M = 42.50$, $SD = 14.07$; Airport B: $M = 51.43$, $SD = 10.41$) and mean work experience was 5.31 years ($SD = 3.98$; Airport A: $M = 4.00$, $SD = 4.13$; Airport B: $M = 6.43$, $SD = 3.61$).[1] Of the 26 participants, approximately half were female (Airport A: 50% female; Airport B: 43% female). Further, slightly less than half of the participants had experienced VR before partaking in this study (Airport A: 42% VR experience; Airport B: 43% VR experience). The participants were informed about the study procedures and goals prior to the study. All participants gave written informed consent and received monetary compensation based on their hourly salary. The study was approved by the institutional ethics review board of the School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland.

Design

The factors media (video vs. VR) and learning activity (passive vs. interactive) were varied in a 2 x 2 between-subjects design, leading to four experimental groups experiencing the multimedia lesson with uniform learning content: Passive video, interactive video, passive VR, interactive VR. Participants at each airport were randomly assigned to one of the four

---

[1] No significant differences for participants' age or work experience were detected between airports using Student's *t* tests.

experimental conditions, equalling eight groups of three or four participants. After the learning intervention, all participants completed a questionnaire measuring learning outcomes, cognitive load, intrinsic motivation, and technology acceptance. Finally, a focus group discussion was conducted with each of the eight groups.

Materials

The multimedia lesson was developed in consideration of the CTML's instructional design principles to manage cognitive load (see Mayer, 2009, for an overview of CTML design principles). The multimedia lesson focused on providing novel and experienced 2D X-ray screeners with introductory information about 3D CT imaging systems. Based on a revision of Bloom's taxonomy by Anderson et al. (2001), the learning material contained mainly factual knowledge (e.g., the 3D CT machine you see in front of you meets the C3 standard of the European Civil Aviation Conference) and some conceptual knowledge (e.g., this means that liquids and laptops can be left in baggage for the scanning process).

While keeping learning content uniform, a tailored version of the multimedia lesson was developed for each experimental group depending on the administered media and learning activity. In line with Moreno and Mayer's (2007) proposed types of interactivity, the interactive versions of the multimedia lesson encompassed control over pace and manipulation. Participants were regularly prompted to initialise the next part of the multimedia lesson by pressing a virtual button (see Figure 1), thus controlling the pace in which the learning material progressed. Further, participants were given the possibility of manipulating a laptop in the last part of the multimedia lesson. By rotating a laptop in any direction, participants could view the laptop's 3D CT image from any chosen angle. For passive groups, on the other hand, pace and laptop rotation was predetermined (see Figure 2).

Figure 1

Screenshot of the virtual button being pressed in the interactive VR version

Note. Participants experiencing the interactive VR version used an Oculus touch controller for interactions.

Figure 2

Screenshot of the rotating laptop in a passive video version

Through head movements, participants of the multimedia lessons' VR versions had the possibility of looking around freely in the VE. As videos do not allow for such control, visuals in the video versions adopted a static camera focusing on the relevant learning content. Therefore, videos were considered displays of an optimal viewing of the VR versions.

The multimedia lesson consisted of five consecutive parts: First participants received an introduction and tutorial, explaining the multimedia lesson's aim and the respective media and learning activity. This allowed participants to acclimate themselves to the video or VE. In the interactive versions, the tutorial further explained how to use the input modalities and let participants practice controlling the pace and manipulation of the learning material. The second part consisted of a historical presentation of security screening at airports, showing the technical progression from preceding 2D X-ray to the novel 3D CT imaging systems. The third part informed participants about 3D CT hardware and working principles. In the fourth part, two baggage trays were scanned, and participants were presented with the new user interface and software features of 3D CT technology. In the last part, participants viewed a rotatable laptop and corresponding 3D CT image.

Procedure

The study took place at the airports' facilities. Workstations were set up in a quiet and normally lit room, familiar to the participants. The VR versions of the learning material were administered with first-generation Oculus Quest HMDs (resolution per display: 1440 × 1600 pixels). For interactive VR, participants used a standard Oculus touch controller (second-generation Oculus Touch) corresponding to their handedness. Videos were displayed on Laptops with 17.3-inch monitors (resolution: 1920 x 1080 pixels). A standard wired computer mouse served the participants as input device for the interactive video condition. The audio was delivered through on-ear headphones for all experimental conditions. In order to reduce visual distractions for video conditions, workstations were fitted with a cardboard visual cover.

Experimental groups first received oral information and instructions concerning the study procedure. Then, before starting the multimedia lesson, VR groups were shown how to use the VR equipment correctly. This entailed ensuring a comfortable fit of the HMD and adjusting the pupillary distance between the HMDs screens. Additionally, participants of the interactive VR groups were shown how to grip the VR controller properly. Depending on the experimental condition and pace of the participant, the multimedia lesson had a mean total duration of 15.66 min (*SD* = 2.73; see Appendix A for an overview of each experimental condition). The tutorial specific to each experimental condition had a mean duration of 3.58 min (*SD* = 2.38). After the learning intervention, the participants were immediately given a questionnaire. Following a subsequent break of 10 to 15 min, the focus groups were conducted. Each focus group lasted around one hour.

Measures

Questionnaire

The quantitative data was measured using a paper-pencil questionnaire (see Appendix B). Questionnaire items were administered in German and adapted to the present study if necessary. For some items, this involved slightly altering the wording (e.g., changing “task” to “learning task”) and translating them when no German version was available (e.g., Beaton et al., 2000). For this process, a native bilingual speaker translated the original items from English to German. Another native speaker of English and German then translated the items back into English. An item revision by the translators followed this process.

Learning outcomes were assessed as a dependent measure by administering a learning performance test in the questionnaire. The performance test consisted of 12 multiple-choice items (e.g., 3D CT machines meet the C3 standard of the European Civil Aviation Conference. What items are passengers allowed to leave in their luggage?). Each correct answer scored one point while incorrect answers scored zero points, resulting in a maximum score of 12 points.

Cognitive load was measured by using two separate subjective rating scales. Subjective rating scales have shown to be similar in validity and reliability as physiological measurement techniques (Szulewski et al., 2017). The main advantage of subjective rating scales is their sensitivity and simplicity (Sweller et al., 2019). The first subjective rating scale used in the questionnaire was developed by Paas (1992) and was measured on a nine-point Likert scale, which is commonly used for this item (Park & Brünken, 2015). Even though the item was originally called mental effort (Paas, 1992), it is considered the most frequently used measure for cognitive load (Mutlu-Bayraktar et al., 2019). The second instrument applied for assessing cognitive load was developed by Klepsch et al. (2017) and contained nine items measured on a seven-point Likert scale. This instrument has the advantage of measuring different types of cognitive load (intrinsic, extraneous, germane) separately. Differentiated measuring scales have found academic interest in recent years, promising better interpretation of the results and linking to the theoretical base (Mutlu-Bayraktar et al., 2019). Cronbach's α values reported by Klepsch et al. (2017) were .81, .86, and .67 for the intrinsic cognitive load, extraneous cognitive load and germane cognitive load scales, respectively.

Intrinsic motivation was assessed with the interest and enjoyment subscale of the intrinsic motivation inventory (IMI; Ryan, 1982). This subscale contained seven items measured on a seven-point Likert scale and is considered the self-reported measure of intrinsic motivation (Cortright et al., 2013).

Technology acceptance was evaluated with the perceived usefulness, perceived ease of use, and behavioural intention subscales of the TAM3 by Venkatesh and Bala (2008), an extension of the original TAM. Each subscale was measured on a seven-point Likert scale. Perceived use and perceived ease of use consisted of four items, and behavioural was composed of three items.

Focus groups

Focus groups were conducted to gain further insight into the participants' experiences when learning with interactive and passive video and VR. Focus groups allow collecting extensive data through active participation and discussion among interviewees (Krueger & Casey, 2014). Furthermore, the group setting makes it possible to uncover a broad range of perspective and gain a deeper understanding of the issues from the viewpoint of the participants (Hennink & Leavy, 2014). In keeping with focus group literature, a discussion guide was crafted (e.g., Barbour & Morgan, 2017; Hennink & Leavy, 2014; Masadeh, 2012; see Appendix C). The focus groups were conducted in a semi-structured format and led by the author of this study. Audio recordings were made of all eight focus groups. Additionally, an observer was present for each focus group noting prominent gestures, themes and quotes (e.g., Hennink & Leavy, 2014).

The opening section of the focus group consisted of a broad question and a brief activity. In the activity, participants had to choose their most and least favoured part of the learning module. This section meant to build rapport among the participants and make them feel at ease before moving to more specific and critical topics (Hennink & Leavy, 2014). Sections two to four served as the main focus and intended to ascertain participants' experiences. Emphasis was laid on using context-specific questions to promote discussion and gain personal insights (e.g., given the current Covid-19 pandemic, what would the reaction be if you had to continue using this learning media at home?). In the last section, the most important topics were revisited if necessary, and participants had the possibility of making final statements.

Data Analysis

Quantitative

All quantitative data was analysed using Jamovi (version 1.2.27). For the rating scales of cognitive load, intrinsic motivation and technology acceptance, reliability was explored using Cronbach's α (Cronbach, 1951). Values are interpreted as unacceptable below .60, undesirable between .60 and .65, minimally acceptable between .65 and .70, respectable between .70 and .80, and very good between .80 and .90 (DeVellis, 2016, p. 136). A low value is often an indicator for a low number of questions, poor inter-relatedness between items or heterogeneous constructs (Tavakol & Dennick, 2011). In this case, Tavakol and Dennick (2011) advise reviewing or even discarding items if the low α value is due to poor correlation between items.

For the measure learning outcome, items of the performance test were analysed using the item difficulty index. The item difficulty index ranges from 0% to 100% and refers to the percentage of participants who correctly answered the item (Quaigrain et al., 2017) According to Boopathiraj and Chellamani (2013), items with a value between 20% and 90% are considered acceptable.

All quantitative measures were investigated using two-way analysis of variance (ANOVA). Prior, normal distribution was assessed using Kolmogorov-Smirnov tests and quantile-quantile plots (Q-Q plots). Additionally, homoscedasticity was tested with the Levene's test. For the ANOVAs, the experimental groups of both airports were joined. Media (video vs.

VR) and learning activity (passive vs. interactive) acted as independent variables, and learning outcomes, cognitive load (intrinsic cognitive load, extraneous cognitive load, germane cognitive load, mental effort), intrinsic motivation, and technology acceptance (perceived usefulness, perceived ease of use, behavioural intention) served as the dependent variables. In order to enhance the interpretation of the results, the effect size ω^2 was calculated (Cumming, 2013). Effect sizes are standardised and objective measures that indicate the magnitude of an observed effect (Field, 2018). For the present analysis, ω^2 was most suitable as it tends to be less biased for small sample ANOVA calculations compared to η^2 (Olejnik & Algina, 2003). Effect sizes of ω^2 are interpreted as small (.01), medium (.06), and large (.14) (Cohen, 1988, p. 368). Interactions of ANOVAs with at least a medium interaction effect were further investigated with post hoc tests.

Focus groups

Qualitative data was transcribed and analysed using MAXQDA (version 20.0.2). Analysis followed the principles of systematic qualitative text analysis (Kuckartz, 2014). By applying deductive-inductive category building, this approach allows for both a theory-driven and open, explorative analysis (Kuckartz, 2014). Deductive categories were derived from thematic groups from the theoretical background and focus group discussion guide. Inductive categories, consisting of additional relevant information and categorical sub-groups, were identified within the collected data.

In order to enhance the interpretive quality of the data, a systematic analysis approach was followed, thereby enhancing the objectivity of the results (Döring & Bortz, 2016). Further, the analytic process was documented to allow for easier intersubjective comprehensibility (Kuckartz, 2014). For reliability, focus group statements were compared within and across focus group sessions (Knodel, 1993). Further, the author discussed each focus group with the observer. Main themes and important statements were identified collectively, thereby increasing the validity of results (Creswell & Clark, 2017).

Results

Questionnaire

Scale reliability

Table 1 shows the reliability of the assessed rating scales. Very good reliability was found for the intrinsic motivation, perceived usefulness, and behavioural intention scales with Cronbach's α ranging from .86 to .89. The extraneous cognitive load (α = .76) and perceived ease of use (α = .73) scales showed respectable reliability. Scales assessing intrinsic cognitive load (α = .34) and germane cognitive load (α = - .44) showed unacceptable internal consistency. Therefore, one item of the germane cognitive load scale, which correlated negatively, was discarded. The two remaining items (α = .84, M = 6.33, SD = 0.75) indicated good internal consistency and were therefore used for the analysis of germane cognitive load.

Table 1

Means, Standard Deviations, and Cronbach's α

Scale	*M*	*SD*	Cronbach's α
Intrinsic cognitive load (2 items)	2.94	1.49	.34
Extraneous cognitive load (3 items)	1.44	0.56	.76
Germane cognitive load (3 items)	5.55	0.74	- .44
Intrinsic motivation (7 items)	6.35	0.74	.86
Perceived usefulness (4 items)	5.86	0.98	.89
Perceived ease of use (4 items)	6.29	0.71	.73
Behavioural intention (3 items)	6.10	1.04	.89

Note. n = 26.

Item difficulty index

One item with a value of .92 on the item difficulty index was identified outside the acceptable range of .20 and .90 (Boopathiraj & Chellamani, 2013). Therefore, the item was excluded from further analysis. The remaining 11 items assessing learning outcomes ranged from .35 to .89 on the item difficulty index.

Analysis of variance (ANOVA)

Interpretation of Q-Q plots (see Appendix D) exhibited potential deviances from the distributions of interest deviations from normality for extraneous cognitive load, germane cognitive load, perceived ease of use, and behavioural intention. However, Kolmogorov-Smirnov tests showed no deviations from normality. Therefore, all scales were used for further analysis. Levene's tests revealed a violation of the assumption of homogeneity by intrinsic cognitive load indicating significantly different variances between the groups.[2]

Two-way ANOVAs were conducted to assess the impact of media (video vs. VR) and learning activity (passive vs. interactive) on learning outcomes, cognitive load, intrinsic motivation, and technology acceptance. Table 2 shows the number of participants per condition, means, standard deviations, and results of the two-way ANOVAs (according to the American Psychological Association, 2020, p. 217).

---

[2] According to Tabachnick & Fidell (Tabachnick & Fidell, 2019), ANOVAs are robust to violations of homogeneity if sample sizes are relatively equal (within a ratio of 4 to 1) and variances between groups do not exceed the ratio of 10 to 1. By satisfying both criteria, intrinsic cognitive load was deemed appropriate for further analysis.

Table 2

Number of participants, means, standard deviations, and two-way ANOVAs

Variable	Video			VR			ANOVA			
	n	*M*	*SD*	*n*	*M*	*SD*	Effect	*F* ratio	*df*	ω^2
LO										
Interactive	7	7.57	2.99	7	6.57	1.27	MD	1.73	1, 22	.03
Passive	6	7.67	2.73	6	6.33	1.51	LA	0.01	1, 22	.00
							MD x LA	0.04	1, 22	.00
ICL										
Interactive	7	2.64	1.95	7	2.79	1.11	MD	0.55	1, 22	.00
Passive	6	2.83	1.94	6	3.58	0.74	LA	0.67	1, 22	.00
							MD x LA	0.25	1, 22	.00
ECL										
Interactive	7	1.14	0.26	7	1.57	0.50	MD	0.01	1, 22	.00
Passive	6	1.72	0.88	6	1.33	0.37	LA	0.64	1, 22	.00
							MD x LA	3.69	1, 22	.10
GCL										
Interactive	7	6.57	0.45	7	6.07	0.79	MD	0.00	1, 22	.00
Passive	6	6.08	1.07	6	6.58	0.58	LA	0.00	1, 22	.00
							MD x LA	2.88	1, 22	.07
ME										
Interactive	7	3.86	2.12	7	3.86	2.97	MD	0.01	1, 22	.00
Passive	6	2.50	1.64	6	2.67	1.37	LA	2.25	1, 22	.05
							MD x LA	0.01	1, 22	.00
IM										
Interactive	7	6.61	0.48	7	6.59	0.50	MD	2.97	1, 22	.06
Passive	6	5.60	1.07	6	6.50	0.39	**LA**	**4.67***	**1, 22**	**.11**
							MD x LA	3.25	1, 22	.07
PU										
Interactive	7	6.07	0.89	7	6.14	0.70	MD	1.30	1, 22	.01
Passive	6	5.17	1.09	6	5.96	1.16	LA	2.07	1, 22	.04
							MD x LA	0.91	1, 22	.00
PEOU										
Interactive	7	6.54	0.55	7	6.43	0.55	MD	0.01	1, 22	.00
Passive	6	6.04	1.22	6	6.08	0.30	LA	2.17	1, 22	.05
							MD x LA	0.07	1, 22	.00
BI										
Interactive	7	6.29	1.18	7	6.14	1.14	MD	0.30	1, 22	.00
Passive	6	5.67	1.15	6	6.28	0.77	LA	0.32	1, 22	.00
							MD x LA	0.79	1, 22	.00

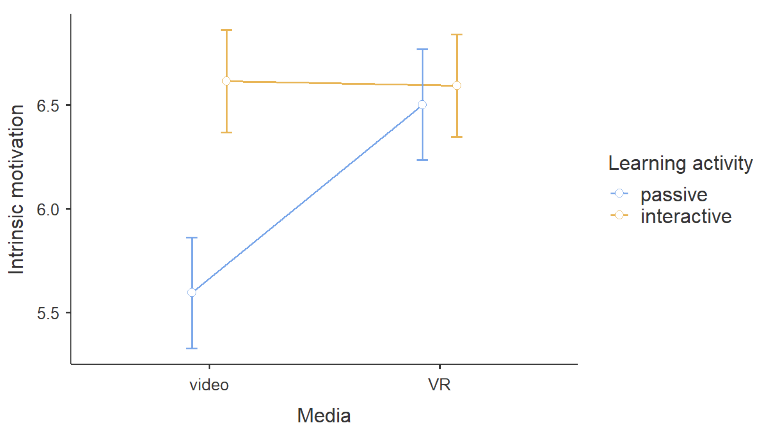
Note. Media = MD; LA = learning activity; LO = learning outcomes; ICL = intrinsic cognitive load; ECL = extraneous cognitive load; GCL = germane cognitive load; ME = mental effort; IM = intrinsic motivation; PU = perceived usefulness; PEOU = perceived ease of use; BI = behavioural intention.

*$p < .05$.

ANOVAs showed a significant main effect for learning activity on intrinsic motivation (passive: *M* = 6.05, *SD* = 0.90; interactive: *M* = 6.60, *SD* = 0.47), as well as an interaction between media and learning activity with medium effect size. Post hoc comparisons investigating interactions of ANOVAs revealed significantly lower intrinsic motivation for passive video (*p* = .010) compared to interactive video (*p* = .025), passive VR and interactive VR (*p* = .012). Figure 3 presents the interaction plot for the intrinsic motivation scale, followed by Figures 4 to 11, showing the interaction plots of the non-significant scales.[3]

Figure 3

Interaction plot for intrinsic motivation



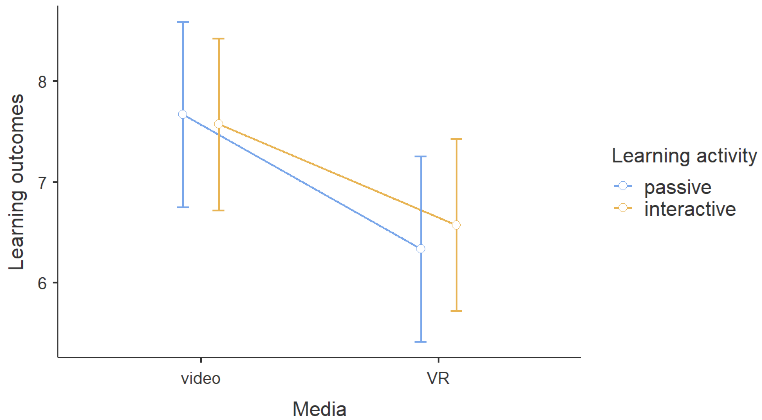
Note. Intrinsic motivation was assessed on a seven-point Likert scale.

Error bars represent the standard error of the mean.

[3] Default Jamovi plots are presented in this study. Therefore, interaction plots are truncated on the y-axis and do not necessarily show the zero-baseline. While this visual exaggeration aims to aid the presentation of results, caution is advised for effect size interpretation (Correll et al., 2020).

Figure 4

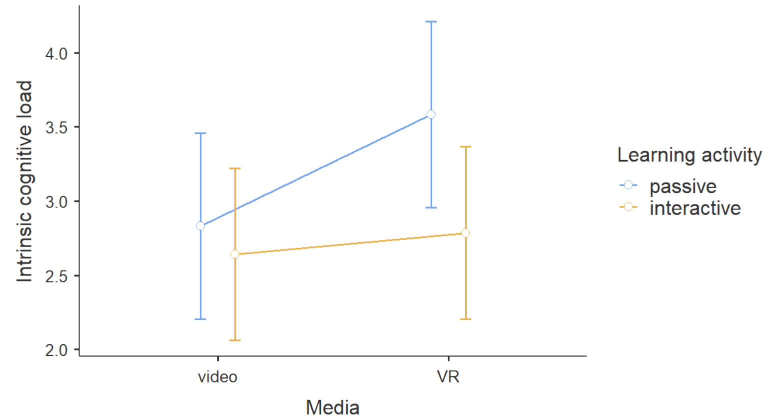
Interaction plot for learning outcomes



Note. Learning outcomes were assessed on a nine-point Likert scale. Error bars represent the standard error of the mean.

Figure 5

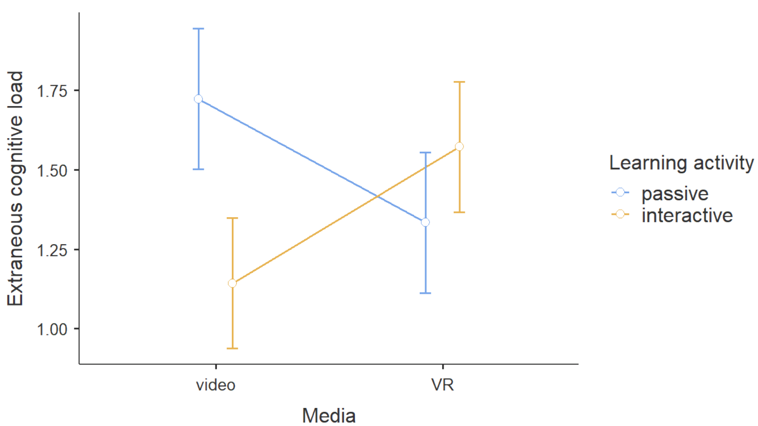
Interaction plot for intrinsic cognitive load



Note. Intrinsic cognitive load was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 6

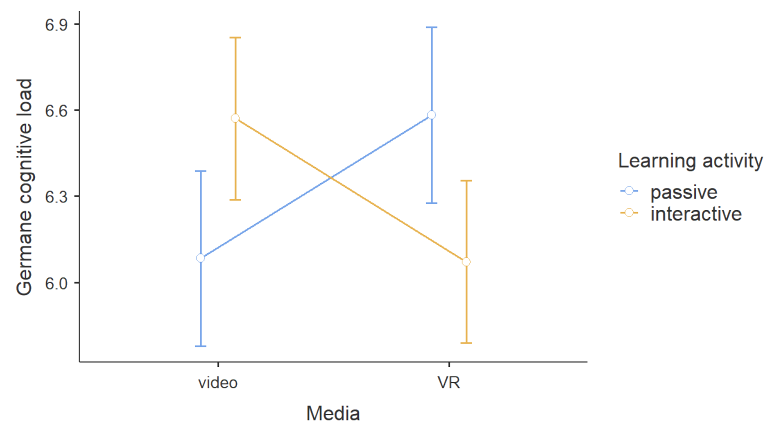
Interaction plot for extraneous cognitive load



Note. Extraneous cognitive load was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 7

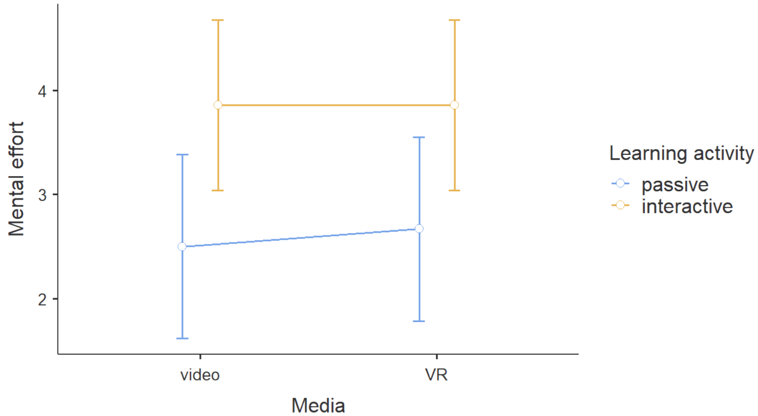
Interaction plot for germane cognitive load



Note. Germane cognitive load was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 8

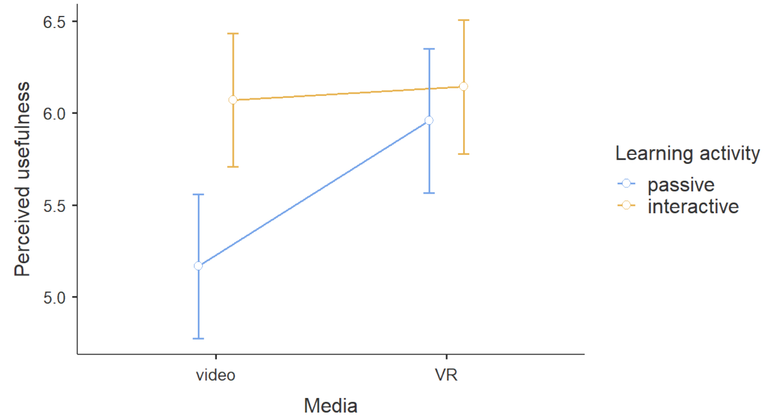
Interaction plot for mental effort



Note. Mental effort was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 9

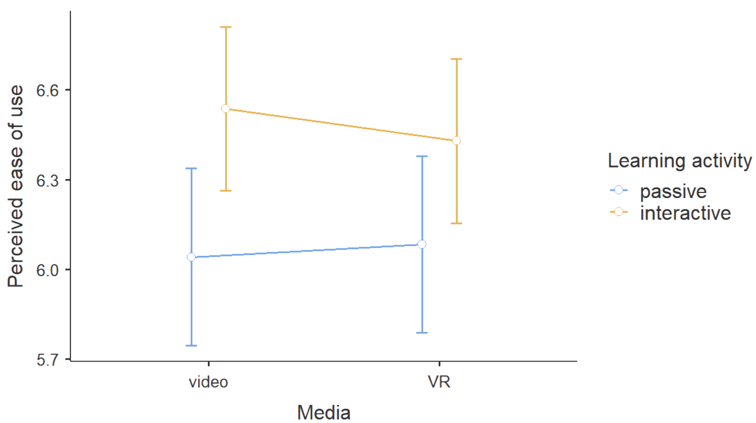
Interaction plot for perceived usefulness



Note. Perceived usefulness was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 10

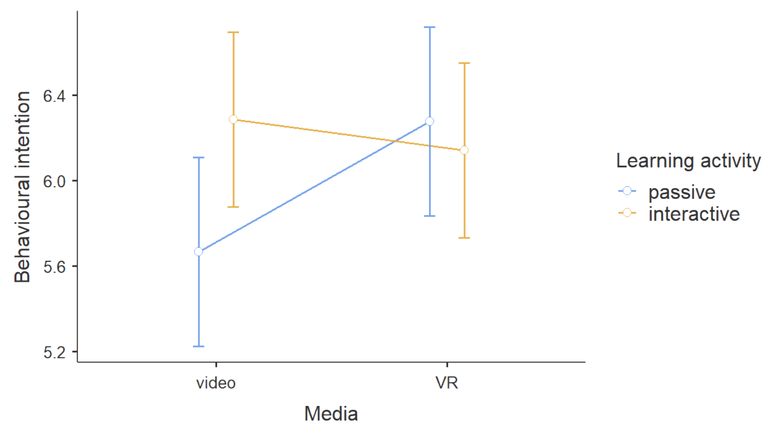
Interaction plot for perceived ease of use



Note. Perceived ease of use was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Figure 11

Interaction plot for behavioural intention



Note. Behavioural intention was assessed on a seven-point Likert scale. Error bars represent the standard error of the mean.

Focus groups

This section describes the perceptions of the participants towards the administered experimental condition (passive video, interactive video, passive VR, interactive VR) captured in the eight focus groups.

Learning

Participants of all groups expressed that they gained knowledge of 3D CT imaging systems through the learning material. When asked if the participants could explain to a colleague what 3D CT is and how it would affect their screening tasks, passive VR participants were diffident. They noted that hands-on experience was required to make conclusive statements about the changes 3D CT entailed.

Participants in interactive video groups believed that being able to interact with the learning material positively affected their learning, with one participant stating: “Well, if I can do something, it just stays with me”. This sentiment was mirrored in interactive VR groups. Further, participants in both passive and interactive VR groups felt that VR facilitated learning processes using more bodily senses compared to traditional learning methods.

Cognitive load

The learning materials’ difficulty level was perceived as appropriate across all groups, and no incomprehensible content was identified. Most participants felt that basic knowledge of cabin baggage screening was necessary in order to understand the learning material fully. Further, participants noted no or minimal prior knowledge about 3D CT imaging systems. The existing knowledge was traced back to personal experiences (e.g., medical procedures) or work-related information. For example, management mentioning future changes from current X-ray to 3D CT imaging systems.

A major theme addressed exclusively in VR groups was feeling distracted by the VE. One passive VR participant with no prior VR experience said: “For me it was the first time. And to look around, everything else is more interesting than what is actually important”. Several other VR participants mentioned that the VE might have hindered their ability to focus on the learning content, especially early in the multimedia lesson. However, participants in interactive VR and interactive video groups felt that regularly pressing the virtual button to initiate the next sequence helped them focus. Further, interactive video and VR participants felt that control over pacing let them adjust the learning pace to their liking. In this context, an interactive VR participant said about the learning material: “I liked it a lot. It adapted to my pace. When I understood it fast, it was also fast”.

Intrinsic Motivation

Most participants stated that they found the learning material interesting and enjoyed the overall experience. While passive and interactive video groups mostly referred their statements to the learning material, passive and interactive VR groups often alluded to the media they were administered. Participants in all VR groups positively described VR as a novelty, or as one passive VR participant put it “just something different”. On the other hand, participants in passive video groups felt that the administered media and learning activity was not particularly exciting, with one participant emphasising that “nowadays, we are used to more”.

Participants in interactive VR and video groups particularly enjoyed manipulating a laptop in the last part of the multimedia lesson. For example, one interactive VR participant said: “That [the manipulation of the laptop], of course, was fascinating. How you can grab that and turn it around yourself. It’s also the playfulness of it”. Conversely, the same part was conceived as rather bland by passive video and VR participants.

Technology acceptance

A majority of participants across all groups found the administered combination of media and learning activity useful for learning information related to airport security. Adverse sentiments were mainly expressed in passive video groups. Participants felt that the absence of human-to-human interactions and not being able to ask questions negatively affected the passive videos' usefulness. Further, participants in passive VR groups voiced displeasure concerning the restrictiveness of a HMD, concluding that it would interfere with them taking notes. On the other hand, interactive video participants mainly mentioned advantages. Being able to simulate real-life working conditions (e.g., loud airport environments) and the possibility of using interactive video for training purposes (e.g., cabin baggage screening training) were seen as having great potential in an airport learning setting. Similarly, passive and interactive VR participants predicted simulating real-life working conditions as VR's most prominent potential for airport security training. Overall, participants across all groups noted a preference for the administered media and learning activity compared to traditional learning methods currently used at airports (e.g., PowerPoint presentations). For example, a passive video participant said: "I'd rather have a learning video than PowerPoint, I have to say. Because when I think back to the theory sessions, it really was just slide after slide - click, click, click".

Participants across all groups generally stated that using the administered media and learning activity was uncomplicated. In this regard, the tutorial at the beginning of the learning material was deemed important and adequate. However, concerning the administered media, participants in passive and interactive VR groups revealed issues regarding the ergonomics and display resolution of the HMD.

Ergonomics and display resolution in VR

Ergonomics and display resolution were two additional themes identified from the focus groups. Both issues were addressed exclusively in passive and interactive VR groups as they corresponded to the HMD used in this study. Participants expressed discontent about the weight and fit of the HMD. Particularly female participants had issues finding a comfortable fit using the HMD's head straps. Further, a few participants found themselves sweating excessively when wearing the HMD. Concerning display resolution, several participants felt that the HMD did not yet display a clear enough picture to warrant future use.

Discussion

The present study employed a mixed methods approach in order to explore learning with media (video vs. VR) and learning activity (passive vs. interactive) in an airport security setting. For this purpose, learning outcomes, cognitive load, intrinsic motivation, and technology acceptance were measured using a questionnaire. Further focus groups were conducted to assess the perceptions of screeners towards the administered media and learning activity.

Learning

With regard to learning, questionnaire data showed a small effect size indicating slightly better learning outcomes with video than VR. Focus group data neither support nor refute the quantitative results, as participants in all groups believed to have gained knowledge of 3D CT imaging systems. Therefore, these findings substantiate the complex relationship between media and learning as literature has previously shown. The superior learning outcomes of screeners with video compared to VR could be ascribed to the conveyed and assessed knowledge in this study. The multimedia lesson aimed at giving screeners introductory information on 3D CT imaging systems, focusing mainly on factual knowledge. However, it has been suggested, that the main instructional potential of VR lies in more complex learning tasks (e.g., Hamilton et al., 2020). Zahn et al. (2004) offer another explanation in suggesting that being unfamiliar with a medium might hinder learning. Therefore, the unfamiliarity of many screeners with VR may have negatively affected learning.

Learning activity did not influence learning outcomes. Interestingly, these results stand in contrast to focus group data. In interactive video and VR groups, screeners believed that being active and using more bodily senses improved their learning. This discrepancy is in line with previous literature by Zhang et al. (2019) showing no objective knowledge gain through increased interactivity, even though participants believed otherwise. The perceived learning advantages of interactivity may be explained due to a higher sense of autonomy through better control over the environment in the interactive groups (e.g., Makransky & Lilleholt, 2018).

Cognitive load

For cognitive load, two instruments were used for assessment. The differentiated measurement instrument showed no effect of media or learning activity on intrinsic cognitive load. This result is consistent with focus group data. Across all groups, participants perceived the level of difficulty of the learning material as appropriate. Combined, these findings reinforce that the learning material did not vary between groups, and learning was not hindered through content complexity.

The extraneous cognitive load scale showed an interaction between media and interactivity with a medium effect size. This result is in accordance with the germane cognitive load scale which showed an inverse interaction with a medium effect size. These findings imply that extraneous cognitive load may have influenced germane cognitive load and therefore learning processes of screeners. In the focus groups, participants in VR groups identified the VE as a possible distraction. On the other hand, interactivity was perceived to enhance focus in interactive video and VR groups. While these qualitative findings do not explain the quantitative results of extraneous and germane cognitive load, they further support the notion that a rich VE can distract learners when using virtual reality (e.g., Makransky, Terkildsen, et al., 2019; Parong & Mayer, 2018).

The mental effort instrument revealed a small effect size indicating slightly higher effort invested for interactive compared to passive groups. This result neither confirms nor contradicts the qualitative and differentiated measurement data but instead offers an additional perspective on the effect of interactivity on cognitive load. While higher cognitive load through more activity is to be expected (e.g., Zhang et al., 2019), it remains unclear how it affected learning in this study.

Intrinsic motivation

Quantitative data revealed that screeners in passive video groups were significantly less intrinsically motivated than in the other groups. These results are reinforced by the qualitative data, as screeners from interactive groups said they particularly enjoyed being able to actively manipulate an object in the multimedia lesson. This finding is consistent with studies assessing the affective value and motivation of interactive video (e.g., Nikopoulou-Smyrni & Nikopoulos, 2010). Further, participants of VR groups highlighted their enjoyment and interest when using the novel media. VR literature confirms this finding with people favouring VR for motivational outcomes compared to less immersive instruction (e.g., Makransky & Lilleholt, 2018). Interestingly, the quantitative data also showed that interactive video is similarly beneficial for intrinsic motivation as passive and immersive VR. Therefore, added immersion may not be the panacea for lacking intrinsic motivation of learners.

Technology acceptance

Regarding technology acceptance, perceived usefulness showed small effect sizes indicating that screeners believed interactivity and VR are more useful for their tasks than passivity and videos, respectively. Drawing on the qualitative data, these findings can be traced back to interactivity and VR offering realistic simulation possibilities of the screeners' real-world tasks. Conversely, drawbacks mentioned almost exclusively in passive video groups were missing human-to-human interactions as well as the possibility of asking questions.

The perceived ease of use subscale showed no difference in the screeners' belief of effort needed to use the assessed media. Regarding learning activity, a small effect size was found showing slightly better scores for interactive compared to passive groups. While focus group data does not provide an explanation for this finding, previous research has shown that perceived control over a system positively affects perceived ease of use (Lee et al., 2007). Considering interactive groups had control of pace and certain aspects of the learning material through manipulation, perceived control may have led to higher perceived usefulness.

Regarding behavioural intention, quantitative results showed no differences for media or interactivity. In view of behavioural intention being determined by perceived usefulness and perceived ease of use, these findings may be explained by a lack of statistical differences found in the other technology acceptance scales.

Ergonomics and display resolution

Through inductive analysis of focus group data, ergonomics and display resolution were identified as major themes when learning with VR. Several participants stated in the focus groups that the HMDs used in this study were too heavy for their linking and did not fit comfortably. Further, some participants felt that the picture displayed in the HMDs was not clear enough. These findings are relevant, as they might have affected the results of this study. Additionally, the findings highlight that even though major technological advancements have made in recent years, state of the art HMDs still face some of the same issues that were identified over 20 years ago (e.g., Nichols, 1999).

Practical implications

This study could not produce strong scientific evidence to support the use of VR for instructional purposes in practical settings such as airports. More specifically, when the goal is to convey facts and basic concepts to employees, the established and more cost-effective videos might offer greater utility. Nevertheless, VR is a novel technology which sparks interest and enjoyment in learners. This can be favourable, especially when instruction aims to motivate and excite learners for new subject matters. If VR is used, findings of this study suggest that novel learners should be given enough time to acclimate to the VE. Additionally, instructional designers and VR developers are advised to minimise potential distractors. Further, the findings of this study show that learning activity plays a key role for intrinsic motivation, perceived usefulness, and perceived learning and should therefore be considered when designing instruction for video and VR.

Limitations and further research

The sample assessed in this study was small, which limits the generalizability of the findings. Moreover, the reduced sample size led to several limitations regarding quantitative analysis and results: First, the small sample did not allow for a control group to assess the knowledge gain of participants. While qualitative data suggests that participants did gain knowledge, future studies may alternatively employ a pretest-posttest design for improved assessment of learning processes. Second, groups of both airports with matching media and learning activity were combined for analysis. While *t* tests showed no significant differences in participants' age or work experience between airports, variance stemming from the respective contexts could have been introduced. Thus, in the case of a larger sample, employing analyses of covariance may allow for a more sensitive assessment of the group means. Third, while Kolmogorov-Smirnov tests suggested normal distributions of the scales, Q-Q plots exhibited potential deviations in some instances. Additionally, the intrinsic cognitive load scale did not meet the assumption of homoscedasticity. However, in view of Tabachnick and Fidell's (2019) recommendations, the scale was considered suitable for analysis. Fourth, analyses predominately yielded non-significant results, which necessitated post hoc tests with no correction. Given a larger sample, a correction applied to the α-level, such as the Holm-Bonferroni correction (Holm, 1979), is advised.

Learning outcomes were assessed using 12 multiple-choice items, of which one item was excluded from analysis for being correctly answered by more than 90% of the participants. As the multimedia lessons' content inherently limits the number of items, future studies using the same learning material should revise the discarded multiple-choice item to enhance overall reliability of the assessment.

Regarding instrument selection, the cognitive load instrument by Klepsch et al. (2017) showed poor reliability for the intrinsic cognitive load and germane cognitive load scales. Considering the advantages of differentiated cognitive load scales for interpreting results and the methodological challenges one-item scales pose (e.g., Leppink et al., 2013; van Gog & Paas, 2008), future studies might consider another differentiated measurement instrument. For example, Leppink et al. (2013) offer a validated alternative with their instrument measuring the three different types of cognitive load (Cook et al., 2017).

Conclusion

Overall, the present study offers insights into the advantages and disadvantages of passive and interactive videos and VR when used in an educational setting. The employed mixed methods design allowed to identify differences in perceived and objective learning. Further, interactive video, passive VR, and interactive VR led to higher intrinsic motivation than passive video. Qualitative data suggests that the novelty of VR and interactions play a crucial role in intrinsically motivating learners. However, rich and exciting VE can distract learners and should therefore be carefully considered when developing learning material for VR. In the near future, the findings of this study will help to improve questionnaires and multimedia material for further studies in the STA^{2}RT research project.

References

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918. https://doi.org/10.1037/0033-2909.84.5.888

Allcoat, D., & von Mühlenen, A. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, *26*, 1–13. https://doi.org/10.25304/rlt.v26.2140

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.).

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Pearson.

Barbour, R. S., & Morgan, D. L. (2017). *A New Era in Focus Group Research*. Palgrave Macmillan UK.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bétrancourt, M., & Benetos, K. (2018). Why and when does instructional video facilitate learning? A commentary to the special issue "developments and trends in learning with instructional video". *Computers in Human Behavior*, *89*, 471–475. https://doi.org/10.1016/j.chb.2018.08.035

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*, *89*, 411–417.

https://doi.org/10.1016/j.chb.2017.12.002

Boopathiraj, C., & Chellamani, D. K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. *International Journal of Social Science & Interdisciplinary Research*, *2*(2), 189–193.

Burbules, N. C. (2006). Rethinking the Virtual. In J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas (Eds.), *The International Handbook of Virtual Learning Environments* (pp. 37–58). Springer.

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 174–205). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.010

Calandra, B., Brantley-Dias, L., & Dias, M. (2006). Using Digital Video for Professional Development in Urban Schools: A Preservice Teacher's Experience With Reflection. *Journal of Computing in Teacher Education*, *22*(4), 137–145.

Chen, C., Shih, B., & Yu, S. (2012). Disaster prevention and reduction for exploring teachers' technology acceptance using a virtual reality system and partial least squares techniques. *Natural Hazards*, *62*(3), 1217–1231. https://doi.org/10.1007/s11069-012-0146-0

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Cook, D. A., Castillo, R. M., Gas, B., & Artino, A. R. (2017). Measuring achievement goal motivation, mindsets and cognitive load: validation of three instruments' scores. *Medical Education*, *51*(10), 1061–1074. https://doi.org/10.1111/medu.13405

Correll, M., Bertini, E., & Franconeri, S. (2020). Truncating the Y-Axis: Threat or Menace? *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376222

Cortright, R. N., Lujan, H. L., Blumberg, A. J., Cox, J. H., & Dicarlo, S. E. (2013). Higher levels of intrinsic motivation are related to higher levels of class performance for male but not female students. *Advances in Physiology Education*, *37*(3), 227–232. https://doi.org/10.1152/advan.00018.2013

Coyne, E., Rands, H., Frommolt, V., Kain, V., Plugge, M., & Mitchell, M. (2018). Investigation of blended learning video resources to teach health students clinical skills: An integrative review. *Nurse Education Today*, *63*, 101–107. https://doi.org/10.1016/j.nedt.2018.01.021

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cumming, G. (2013). *Understanding The New Statistics*. Routledge. https://doi.org/10.4324/9780203807002

Cummings, J. J., & Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, *19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319–339. https://doi.org/10.2307/249008

de Koning, B. B., Hoogerheide, V., & Boucheix, J. M. (2018). Developments and Trends in Learning with Instructional Video. *Computers in Human Behavior*, *89*, 395–398. https://doi.org/10.1016/j.chb.2018.08.055

DeLozier, S. J., & Rhodes, M. G. (2017). Flipped Classrooms: a Review of Key Ideas and

Recommendations for Practice. *Educational Psychology Review*, *29*(1), 141–151. https://doi.org/10.1007/s10648-015-9356-9

DeVellis, R. F. (2016). *Scale Development Theory and Applications* (4th ed.). SAGE.

Domagk, S., Schwartz, R. N., & Plass, J. L. (2010). Interactivity in multimedia learning: An integrated model. *Computers in Human Behavior*, *26*(5), 1024–1033. https://doi.org/10.1016/j.chb.2010.03.003

Donkor, F. (2010). The comparative instructional effectiveness of print-based and video-based instructional materials for teaching practical skills at a distance. *The International Review of Research in Open and Distributed Learning*, *11*(1), 96–116. https://doi.org/10.19173/irrodl.v11i1.792

Donkor, F. (2011). Assessment of Learner Acceptance and Satisfaction with Video-Based Instructional Materials for Teaching Practical Skills at a Distance. *International Review of Research in Open and Distributed Learning*, *12*(5), 74–92. https://doi.org/10.19173/irrodl.v12i5.953

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5th ed.). Springer. https://doi.org/10.1007/978-3-642-41089-5

Evans, C., & Gibbons, N. J. (2007). The interactivity effect in multimedia learning. *Computers and Education*, *49*(4), 1147–1160. https://doi.org/10.1016/j.compedu.2006.01.008

Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE.

Giannakos, M. N. (2013). Exploring the video-based learning research: A review of the literature. *British Journal of Educational Technology*, *44*(6), 191–195. https://doi.org/10.1111/bjet.12070

Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2014). Looking at MOOCs rapid growth through thelens of video-based learning research. *International Journal of Emerging Technologies in Learning*, *9*(1), 35–38. https://doi.org/10.3991/ijet.v9i1.3349

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2020). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*. https://doi.org/10.1007/s40692-020-00169-2

Hennink, M. M., & Leavy, P. (2014). *Understanding Focus Group Discussions*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199856169.001.0001

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Huang, H. M., & Liaw, S. S. (2018). An analysis of learners' intentions toward virtual reality learning based on constructivist and technology acceptance approaches. *International Review of Research in Open and Distance Learning*, *19*(1), 91–115. https://doi.org/10.19173/irrodl.v19i1.2503

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Kay, R. H., & Edwards, J. (2012). Examining the Use of Worked Example Video Podcasts in Middle School Mathematics Classrooms: A Formative Analysis. *Canadian Journal of Learning and Technology*, *38*(3), 1–20. https://doi.org/10.21432/t2pk5z

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*, 1–18. https://doi.org/10.3389/fpsyg.2017.01997

Knight, M. J., & Tlauka, M. (2017). Interactivity in map learning: The effect of cognitive load. *Spatial Cognition & Computation*, *17*(3), 185–198. https://doi.org/10.1080/13875868.2016.1211661

Knodel, J. (1993). The Design and Analysis of Focus Group Studies: A Practical Approach. In D. L. Morgan (Ed.), *Successful Focus Groups: Advancing the State of the Art* (pp. 35–50). SAGE. https://doi.org/10.4135/9781483349008.n3

Krueger, R. A., & Casey, M. A. (2014). *A Practical Guide for Applied Research* (5th ed.). SAGE.

Kuckartz, U. (2014). *Qualitative Text Analysis: A Guide to Methods, Practice & Using Software*. SAGE. https://doi.org/10.4135/9781446288719

Lee, D., Moon, J., & Kim, Y. J. (2007). The effect of simplicity and perceived control on perceived ease of use. *Association for Information Systems - 13th Americas Conference on Information Systems, AMCIS 2007: Reaching New Heights*, *3*, 1764–1776.

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, *55*(4), 1424–1442. https://doi.org/10.1016/j.compedu.2010.06.006

Lee, J., Kim, J., & Choi, J. Y. (2019). The adoption of virtual reality devices: The technology acceptance model integrating enjoyment, social interaction, and strength of the social ties. *Telematics and Informatics*, *39*, 37–48. https://doi.org/10.1016/j.tele.2018.12.006

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and Extrinsic Motivational Orientations in the Classroom: Age Differences and Academic Correlates. *Journal of Educational Psychology*, *97*(2), 184–196. https://doi.org/10.1037/0022-0663.97.2.184

Leppink, J., Paas, F. G. W. C., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J.

G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Lin, C., & Tseng, Y. (2012). Videos and animations for vocabulary learning: A study on difficult words. *Turkish Online Journal of Educational Technology*, *11*(4), 346–355.

Lindgren, R., Pea, R., Lewis, S., & Rosen, J. (2007). Learning from digital video: An exploration of how interactions affect outcomes. *International Society of the Learning Sciences*, *Proceedings of the 8th international conference on Computer supported collaborative learning*, 447–449.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, *31*(3), 313–327. https://doi.org/10.1080/02796015.2002.12086158

Liu, S. H., Liao, H. L., & Pratt, J. A. (2009). Impact of media richness and flow on e-learning technology acceptance. *Computers and Education*, *52*(3), 599–607. https://doi.org/10.1016/j.compedu.2008.11.002

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, *35*(6), 691–707. https://doi.org/10.1111/jcal.12375

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Masadeh, M. A. (2012). Focus Group: Reviews and Practices. *International Journal of Applied Science and Technology*, *2*(10), 63–68.

Mayer, R. E. (2012). *Multimedia Learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, *33*(5), 403–423. https://doi.org/10.1111/jcal.12197

Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning, Second Edition* (pp. 316–344). https://doi.org/10.1017/CBO9781139547369.016

Mcmillan, S. J., & Hwang, J. S. (2002). Measures of Perceived Interactivity: An Exploration of the Role of Direction of Communication, User Control, and Time in Shaping Perceptions of Interactivity. *Journal of Advertising*, *31*(3), 29–42. https://doi.org/10.1080/00913367.2002.10673674

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers and Education*, *140*. https://doi.org/10.1016/j.compedu.2019.103603

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*(3), 309–326. https://doi.org/10.1007/s10648-007-9047-2

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, *40*(2), 120–123. https://doi.org/10.1097/00006199-199103000-00014

Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers and Education*, *141*. https://doi.org/10.1016/j.compedu.2019.103618

Nichols, S. (1999). Physical ergonomics of virtual environment use. *Applied Ergonomics*, *30*(1), 79–90. https://doi.org/10.1016/S0003-6870(98)00045-3

Nikopoulou-Smyrni, P., & Nikopoulos, C. (2010). Evaluating the impact of video-based versus traditional lectures on student learning. *Proceedings of the 7th European Conference on E-Learning, ECEL 2008*, *2*, 214–221.

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Paas, F. G. W. C. (1992). Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology*, *84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Paas, F. G. W. C., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Park, B., & Brünken, R. (2015). The Rhythm Method: A New Method for Measuring Cognitive Load - An Experimental Dual-Task Study. *Applied Cognitive Psychology*, *29*(2), 232–243. https://doi.org/10.1002/acp.3100

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1109/CDC.2017.8264526

Quaigrain, K., Arhin, A. K., & King Fai Hui, S. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, *4*(1). https://doi.org/10.1080/2331186X.2017.1301013

Renkl, A., Atkinson, R. K., Renkl, A., & Atkinson, R. K. (2007). Interactive Learning Environments: Contemporary Issues and Trends. An Introduction to the Special Issue. *Educ Psychol Rev*, *19*, 235–238. https://doi.org/10.1007/s10648-007-9052-5

Roussou, M., & Slater, M. (2017). Comparison of the Effect of Interactive versus Passive Virtual Reality Learning Activities in Evoking and Sustaining Conceptual Change. *IEEE Transactions on Emerging Topics in Computing*, *8*(1), 233–244. https://doi.org/10.1109/TETC.2017.2737983

Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers and Education*, *128*, 256–268. https://doi.org/10.1016/j.compedu.2018.09.015

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, *43*(3), 450–461. https://doi.org/10.1037/0022-3514.43.3.450

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wenzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (pp. 171–195). Routledge.

Santagata, R. (2009). Designing video-based professional development for mathematics teachers in low-performing schools. *Journal of Teacher Education*, *60*(1), 38–51.

https://doi.org/10.1177/0022487108328485

Schnotz, W., & Kürschner, C. (2007). A eeconsideration of cognitive load theory. *Educational Psychology Review*, *19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Song, Y., & Kong, S. C. (2017). Investigating Students' Acceptance of a Statistics Learning Platform Using Technology Acceptance Model. *Journal of Educational Computing Research*, *55*(6), 865–897. https://doi.org/10.1177/0735633116688320

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Szulewski, A., Gegenfurtner, A., Howes, D. W., Sivilotti, M. L. A., & van Merriënboer, J. J. G. (2017). Measuring physician cognitive load: validity evidence for a physiologic and a psychometric tool. *Advances in Health Sciences Education*, *22*(4), 951–968. https://doi.org/10.1007/s10459-016-9725-2

Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Ulrich, F., Helms, N. H., Frandsen, U. P., & Rafn, A. V. (2019). Learning effectiveness of 360° video: experiences from a controlled experiment in healthcare education. *Interactive*

Learning Environments, 1–14. https://doi.org/10.1080/10494820.2019.1579234

van Gog, T., & Paas, F. G. W. C. (2008). Instructional Efficiency: Revisiting the Original Construct in Educational Research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Watson, S. L., Watson, W. R., Yu, J. H., Alamri, H., & Mueller, C. (2017). Learner profiles of attitudinal learning in a MOOC: An explanatory sequential mixed methods study. *Computers and Education*, *114*, 274–285. https://doi.org/10.1016/j.compedu.2017.07.005

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. https://doi.org/10.1162/105474698565686

Wittrock, M. C. (1991). Generative Teaching of Comprehension. *The Elementary School Journal*, *92*(2), 169–184. https://doi.org/10.1086/461686

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, *51*(6), 1991–2005. https://doi.org/10.1111/bjet.13023

Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). Video-based learning: A critical analysis of the research published in 2003-2013 and future visions. *ELmL - International Conference on Mobile, Hybrid, and On-Line Learning*, *June 2015*, 112–119.

Zahn, C., Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos - Design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction*, *14*(3), 275–291. https://doi.org/10.1016/j.learninstruc.2004.06.004

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, *43*(1), 15–27. https://doi.org/10.1016/j.im.2005.01.004

Zhang, L., Bowman, D. A., & Jones, C. N. (2019). Exploring Effects of Interactivity on Learning with Interactive Storytelling in Immersive Virtual Reality. *11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1–8. https://doi.org/10.1109/VS-Games.2019.8864531

Appendix A

Duration of the multimedia lesson for all experimental conditions

Table A1

Total duration of the multimedia lesson

	Video		VR	
	M	*SD*	*M*	*SD*
Interactive	15.24	0.50	19.52	1.52
Passive	12.52	0.00	14.78	0.48

Note. Data is depicted in min.

Table A2

Duration of the tutorial

	Video		VR	
	M	*SD*	*M*	*SD*
Interactive	2.84	0.36	6.91	1.62
Passive	0.95	0.00	3.19	0.49

Note. Data is depicted in min.

Appendix B

Paper-pencil questionnaire

Note. The items “Ich habe mich angestrengt, mir nicht nur einzelne Dinge zu merken, sondern auch den Gesamtzusammenhang zu verstehen.” (germane cognitive load) and “Aus welchem Grund wurden Röntgengeräte an Flughäfen erstmals eingeführt?” (learning outcomes) were not included for ANOVAs due to poor Cronbach’s α and high item difficulty index, respectively.

Fragebogen «3D CT Einführung»

Angaben zur Person

Name ....

Vorname ....

Alter ....

Geschlecht m ☐ w ☐ anders ☐

In welchem Bereich arbeiten Sie derzeit am meisten (Hauptbereich)?

CBS (Passagierkontrolle) ☐

HBS ☐

STAFF-Screening ☐

Arbeitspensum im Hauptbereich ....

Im Hauptbereich tätig seit (Monat, Jahr) ....

Wissen und Erfahrung vor der Lerneinheit «3D CT Einführung»

Wussten Sie vor der soeben erlebten Lerneinheit, was dreidimensionale Computertomographie (3D CT) ist?

Ja ☐ Nein ☐

Haben Sie vor der soeben erlebten Lerneinheit bereits Erfahrungen mit virtueller Realität gemacht?

Ja ☐ Nein ☐

Haben Sie vor der soeben erlebten Lerneinheit bereits Erfahrungen mit virtueller Realität zu Lernzwecken (Trainings, Schulungen etc.) gemacht?

Ja ☐ Nein ☐

Haben Sie vor der soeben erlebten Lerneinheit bereits Erfahrungen mit Lernvideos, die Sie an einem Computerbildschirm geschaut haben, gemacht?

Ja ☐ Nein ☐

Seite 1 von 7

Wie haben Sie die Lerneinheit «3D CT Einführung» erlebt?

Die folgende Frage bezieht sich auf den Inhalt und auf das Lernmedium bzw. -system, mit welchem Sie die Lerneinheit erlebt haben. Bitte kreuzen Sie nur die zutreffendste Antwort an.

Bitte bewerten Sie den Grad Ihrer mentalen Anstrengung, der Sie bei der Lerneinheit «3D CT Einführung» ausgesetzt waren.

Sehr geringe mentale Anstrengung	1	2	3	4	5	6	7	8	9	Sehr hohe mentale Anstrengung
	☐	☐	☐	☐	☐	☐	☐	☐	☐	

Die folgenden Fragen beziehen sich auf den Lerninhalt. Bitte kreuzen Sie jeweils die zutreffendste Antwort an.

Bitte kreuzen Sie nur eine Antwort pro Frage an	**Stimmt gar nicht** **1**	**2**	**3**	**Stimmt teils-teils** **4**	**5**	**6**	**Stimmt völlig** **7**
Die absolvierte Lerneinheit hat mir sehr gefallen.	☐	☐	☐	☐	☐	☐	☐
Die Lerneinheit hat Spass gemacht.	☐	☐	☐	☐	☐	☐	☐
Ich fand die Lerneinheit langweilig	☐	☐	☐	☐	☐	☐	☐
Ich habe bei der Bearbeitung der Lerneinheit gar keine Aufmerksamkeit investiert.	☐	☐	☐	☐	☐	☐	☐
Die Lerneinheit war sehr interessant für mich.	☐	☐	☐	☐	☐	☐	☐
Die Lerneinheit war sehr unterhaltsam für mich.	☐	☐	☐	☐	☐	☐	☐
Ich hatte sehr viel Vergnügen bei der Lerneinheit.	☐	☐	☐	☐	☐	☐	☐

	Stimmt gar nicht 1	2	3	Stimmt teils-teils 4	5	6	Stimmt völlig 7
Bei der Lerneinheit musste man viele Dinge gleichzeitig im Kopf bearbeiten.	□	□	□	□	□	□	□
Diese Lerneinheit war sehr komplex.	□	□	□	□	□	□	□
Bei dieser Lerneinheit ist es mühsam, die wichtigsten Informationen zu erkennen.	□	□	□	□	□	□	□
Die Darstellung bei dieser Lerneinheit ist ungünstig, um wirklich etwas zu lernen.	□	□	□	□	□	□	□
Bei dieser Lerneinheit ist es schwer, die zentralen Inhalte miteinander in Verbindung zu bringen	□	□	□	□	□	□	□
Ich habe mich angestrengt, mir nicht nur einzelne Dinge zu merken, sondern auch den Gesamtzusammenhang zu verstehen.	□	□	□	□	□	□	□
Es ging mir beim Bearbeiten der Lerneinheit darum, alles richtig zu verstehen.	□	□	□	□	□	□	□
Die Lerneinheit enthielt Elemente, die mich unterstützen den Lernstoff besser zu verstehen.	□	□	□	□	□	□	□

Die folgenden Fragen beziehen sich auf das Lernmedium bzw. -system, mit welchem Sie die Lerneinheit erlebt haben. Bitte kreuzen Sie jeweils die zutreffendste Antwort an.

Bitte kreuzen Sie nur eine Antwort pro Frage an	**Stimmt gar nicht** **1**	**2**	**3**	**Stimmt teils-teils** **4**	**5**	**6**	**Stimmt völlig** **7**
Angenommen ich habe Zugang zum Lern-system, dann beabsichtige ich es zu nutzen.	□	□	□	□	□	□	□
Ich plane das Lernsystem in Zukunft wieder zu nutzen.	□	□	□	□	□	□	□
Wenn ich Zugang zum Lernsystem hätte, sage ich voraus, dass ich es nutzen werde.	□	□	□	□	□	□	□
Die Nutzung des Lernsystems würde meine Lernleistung verbessern.	□	□	□	□	□	□	□
Die Nutzung des Lernsystems würde die Produktivität beim Lernen erhöhen.	□	□	□	□	□	□	□
Die Nutzung des Lernsystems würde die Ef-fektivität beim Lernen erhöhen.	□	□	□	□	□	□	□
Ich finde das Lernsystem wäre nützlich zum Lernen.	□	□	□	□	□	□	□
Der Umgang mit dem Lernsystem ist für mich klar und verständlich.	□	□	□	□	□	□	□
Die Benützung des Lernsystems erfordert von mir keine grosse geistige Anstrengung.	□	□	□	□	□	□	□
Ich finde das Lernsystem leicht zu bedie-nen.	□	□	□	□	□	□	□
Ich finde das Lernsystem macht ohne Prob-leme das, was ich möchte.	□	□	□	□	□	□	□

<u>Was haben Sie in der Lerneinheit «3D CT Einführung» gelernt?</u>

Sie haben den letzten Teil des Fragebogens erreicht, in welchem wir herausfinden möchten, was Sie noch zum Inhalt der Lerneinheit «3D CT Einführung» wissen.

Bitte kreuzen Sie jeweils die <u>zutreffendste</u> Antwort an. Für jede Frage gibt es nur <u>eine</u> korrekte Antwort.

1. Aus welchem Grund wurden Röntgengeräte an Flughäfen erstmals eingeführt?

□ Globalisierung
□ Terrorismusbekämpfung
□ Zunehmende Passagierzahlen
□ Neue Waffengesetze

2. Wann wurden die ersten Röntgengeräte an Flughäfen eingesetzt?

□ 1940er Jahre
□ 1960er Jahre
□ 1970er Jahre
□ 1980er Jahre

3. Wozu sind sogenannte Dual-Energie-Röntgenmaschinen in der Lage?

□ Die Materialzusammensetzung von Objekten ermitteln
□ Dreidimensionale Röntgenbilder generieren
□ Röntgenbilder aus zwei Perspektiven darstellen
□ Objekte von zwei Förderbändern gleichzeitig scannen

4. In welchem Feld ausserhalb der Flugsicherheit wird 3D CT bereits seit einiger Zeit erfolgreich eingesetzt?

□ Baugewerbe
□ Robotik
□ Sportevents
□ Medizin

5. In welchem Jahr müssen gewisse Länder 2D-Röntgengeräte durch 3D CT Maschinen ersetzen?

□ 2020
□ 2022
□ 2023
□ 2025

6. 3D CT Geräte erfüllen den C3 Standard der Europäischen Zivilluftfahrt-Konferenz ECAC. Welche Gegenstände dürfen Passagiere demnach in ihren Gepäckstücken belassen?

□ Laptops, aber keine Flüssigkeiten
□ Laptops und Flüssigkeiten
□ Flüssigkeiten, aber keine Laptops
□ Weder Flüssigkeiten noch Laptops

7. Worin unterscheidet sich der Bildaufnahmeprozess von herkömmlichen Röntgenmaschinen und solchen von neuen 3D CT Maschinen?

□ 3D CT Maschinen nehmen farbige Bilder auf
□ 3D CT Maschinen nehmen mehr Bilder auf
□ 3D CT Maschinen sind lauter
□ 3D CT Maschinen sind schneller

8. Wie werden die volumetrischen Pixel genannt, aus welchen die 3D CT Aufnahmen bestehen?

□ Vertex
□ Polygon
□ OLED
□ Voxel

9. Welchen Nachteil haben 3D CT Aufnahmen zurzeit noch gegenüber von 2D-Röntgenbildern?

□ Bildqualität
□ Kosten
□ Grösse
□ Farben

10. Worin liegt ein Unterschied vom Slabbing zum Slicen?

- □ Laptops werden beim Slabbing ebenfalls aus der 3D CT Aufnahme entfernt
- □ Beim Slabbing kann die Schichtdicke benutzerdefiniert eingestellt werden
- □ Der Betrachtungswinkel ist anders
- □ Es gibt keinen Unterschied

11. Welchen Gegenstand erkennt die Leidos-Maschine automatisch?

- □ Smartphone
- □ Mineralwasserflasche
- □ Taschenmesser
- □ Laptop

12. Welche weitere Funktionalität bieten 3D CT Aufnahmen gegenüber klassischen Röntgenbildern?

- □ 3D CT Aufnahmen sind dreidimensional und können somit rotiert werden
- □ 3D CT Aufnahmen sind dreidimensional und können somit in virtueller Realität betrachtet werden
- □ 3D CT Aufnahmen können in mehr Farben dargestellt werden
- □ 3D CT Aufnahmen sehen sind durch ein standardisiertes Bildformat herstellerübergreifend gleich aus

Bitte kontrollieren Sie nochmals, ob Sie alle Fragen beantwortet haben. Vielen Dank für Ihre Teilnahme!

Appendix C

Focus group discussion guide

Note. The discussion guide also included questions focusing mainly on the formative evaluation of the multimedia lesson, which was not assessed in the scope of this study.

Fokusgruppendiskussion «3D CT Einführung» - Leitfaden

Diskussionsleiter:	Beobachter:
Flughafen:	Datum:
Uhrzeit Start:	Uhrzeit Ende:
Teilnehmende:	Experimentalgruppe:

Einleitung	Willkommen zu diesem zweiten Teil der heutigen Studie. Sie haben vor der Pause die Lerneinheit «3D CT Einführung» in Form von …….. erlebt und einen Fragebogen dazu ausgefüllt. Nun möchten wir im Rahmen einer Fokusgruppendiskussion weiter auf Aspekte des **Inhalts** und des **Lernmediums bzw. -systems** eingehen, mit welcher Sie die vorherige Lerneinheit erlebt habt.
Grundregeln	Ich werde die Fokusgruppe moderieren, stelle Ihnen Fragen und gebe hie und da einen Gedankeninput. Während die Fragen stets an euch alle gestellt sind und ich euch bitte diese auch untereinander zu diskutieren, **frage ich allenfalls auch manchmal direkt bei einem von Ihnen nach,** wenn ich länger nichts mehr von Ihnen gehört habe. Es geht bei dieser Diskussion um **Ihre** Gedanken, Einstellungen, Meinungen und Gefühle. **Um Ihre persönliche Sichtweise.** Unabhängig, was im vorherigen Fragebogen bereits abgefragt wurde, möchte ich hiermit den Rahmen öffnen alles zum Inhalt und der Technologie zu besprechen. Es gibt **kein Richtig oder Falsch**. Die Meinungen von jeder Person sind wichtig und egal, ob Sie die gleiche Meinung wie Ihre Gesprächspartner und- Partnerinnen haben oder eine andere – teilen uns Sie diese bitte mit. Die Diskussion wird aufgezeichnet und anschliessend transkribiert. Dabei werden keine Namen genannt. Ihre Aussagen bleiben also **anonym**. Weiter werden die Daten nur für Forschungszwecke genutzt. Dies ist uns wichtig, damit Sie sich in diesem Rahmen wohlfühlen. Für eine gute Gesprächsatmosphäre und Aufnahmequalität ist es von Vorteil, wenn jeweils **nur eine Person spricht.** Diese Fokusgruppendiskussion dauert voraussichtlich ca. **60 Minuten.** Bitte melden Sie sich jeder Zeit, falls Sie irgendwas stört (Lärm, Durchzug etc.) oder etwas unklar sein sollte (z.B. unklare Frage. Haben Sie eine **Frage** bevor wir beginnen?

AUFNAHMEGERÄTE STARTEN!

Seite **1** von **6**

Teil 1: Einstieg und Warm-up

Dauer:	~ 10 min
Start:	Ende:
	Als Einstieg würde ich Sie gerne kurz durch die Runde gehen und Sie bitten zu sagen...
F1 **Allgemeine Einstellungen / erste Gedanken**	Wie haben Sie die Lerneinheit erlebt? *Falls niemand sich meldet, jemanden direkt ansprechen.*
	Vielen Dank. Das war bereits ein interessanter Einstieg. Nun würde ich die Frage gerne etwas ausweiten.
F2 **Wichtigkeit der Teile und Bedürfnisse**	Die Bilder, die ich Ihnen nun präsentiere, stellen in chronologischer Reihenfolge die Inhalte der 3D CT Einführung dar. Sie sehen... 1. … die Einleitung oben, 2. … die ersten historischen Informationen nach dem Wechsel nach unten, 3. ... die single-energy Maschine, 4. … die dual-energy Maschine, 5. … die Einführung der 3D CT Maschine, 6. ... die Erklärung der Funktionsweise der 3D CT Maschine, 7. ... den Scanprozess eines Rucksacks mit Video, 8. … den Scanprozess zweier Handtaschen mit Video, 9. ... und das abschliessende Drehen des Laptops. Sie erhalten nun je 2 Kleber. Einen grünen und einen roten. Nehmen Sie sich einen Moment Zeit und kleben Sie bitte den grünen Kleber zu dem Teil der «3D CT Einführung», der Ihnen am besten gefallen hat, und den roten Kleber zu dem Teil, der ihnen am wenigsten Gefallen hat. Bitte erläutern Sie, wieso Sie die jeweiligen Kleber an die entsprechenden Orte geklebt haben. Was meinen die anderen dazu?

Teil 2: Inhalt	
Dauer:	~ 20 min
Start:	Ende:
	Gerne würde ich nun weiter auf den Inhalt der Lerneinheit eingehen. Doch zuvor...
F3 **Vorerfahrung / Wissenszuwachs**	Was wussten Sie vor dieser Lerneinheit über 3D CT? Hat Ihnen die Lerneinheit geholfen Ihr Wissen über 3D CT zu verbessern? Wie würden Sie nun einer Kollegin oder einem Kollegen erklären, was 3D CT ist und welche Veränderungen die Technologie für Ihre Tätigkeiten bedeutet?
F4 **3D CT Schulung**	Die «3D CT Einführung» hat zum Ziel, Screener auf den Wechsel von Röntgengeräten zu 3D CT Maschinen vorzubereiten. Gehen wir nun davon aus, dass Ihr Flughafen in den kommenden Monaten eine 3D CT Maschine anschafft und Sie bald an einer Schulung dazu teilnehmen werden. Würden Sie auf Grund der heutigen «3D CT Einführung» bereit und informiert genug dazu fühlen diese Schulung zu starten? Was meinen die anderen dazu?
F5 **Inhaltliche Erweiterungen**	F5: Zusammenfassend kann ich festhalten, dass Euch folgende Aspekte *notierte Erweiterungen aufzählen* gefehlt haben. Ist das korrekt? Wenn Ihr nun einen dieser Erweiterungen auswählen müsstet, um in einer nächsten, überarbeiteten Version der Lerneinheit zu ergänzen, welcher wäre es?

Teil 3: Gestaltung	
Dauer:	~ 10 min
Start:	Ende:
	Gerne würde ich nun auf die Gestaltung der Lerneinheit eingehen.
F6 **Lernumge-bung**	Sie haben sich während dieser Lerneinheit in einem Terminal befunden, wie war dies für Sie? • Geräuschkulisse? • 360°? • Realitätsnähe?
F7 **Erzähl-stimme**	Wie würden Sie die Stimme, welche Ihnen die Informationen zur 3D CT Technologie erzählt hat, beschreiben?
F8 **Maschinen-typ**	Im Entwicklungsprozess dieser Lerneinheit haben wir sowohl eine herstellerspezifische Version ……, und eine herstellerunspezifische Version entwickelt. Sie haben die herstellerspezifische Version mit der …… Maschine erlebt, was bietet Ihrer Meinung nach einem grösseren Mehrwert?

Teil 4: Lernmedium	
Dauer:	~ 20 min
Start:	Ende:
	Vielen Dank. Gerne würde ich den Fokus nun auf das Lernmedium, mit welchem ihr die Lerneinheit erlebt habt, eingehen.
F9 **Vorerfah-rung Lern-medium**	Welche Erfahrungen hatten sie vor der heutigen Studie mit ……? • Zum Lernen?
F10 **COVID-19**	Gehen wir davon aus, dass aufgrund der derzeitigen COVID-Situation Schulungen nötig sind, die Sie zuhause machen müssten. Dazu erhalten Sie ein System um regelmässig mit …… zu Lernen. Was ist Ihre Reaktion dazu? • Positives? Wieso? Interesse? Motivierend? Spass? • Negatives? Wieso? Ablenkend? Anstrengung? • Privater Gebrauch?
F11 **Ausbildung**	Wenn Sie sich an Ihre Screener-Ausbildung zurückbesinnen, mit welchen Medien oder didaktischen Mitteln (Powerpoint etc.) haben Sie gelernt und könnten Sie sich vorstellen, dass …… die ergänzen oder gar ersetzen könnte? • In welchem Bereich/für welche Schulungen?

Teil 5: Abschluss		
Dauer:	~ 5 min	
Start:		Ende:
F12	*(Gerne würde ich an dieser Stelle nochmals auf ….. zurückkommen)*	
Relevante Themen / Fragen		
	Haben Sie noch eine Frage/Input/Anliegen?	

AUFNAHMEGERÄTE STOPPEN!

Appendix D

Q-Q plots of all scales

Figure D1

Q-Q plot for learning outcomes

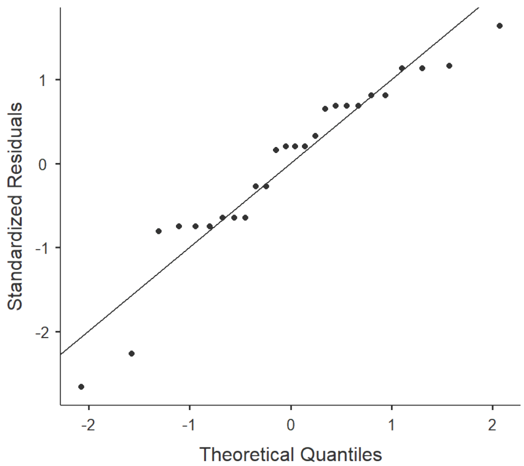
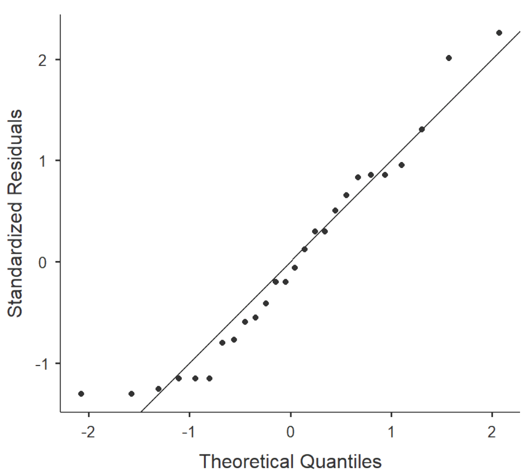


Figure D2

Q-Q plot for intrinsic cognitive load



6

Figure D3

Q-Q plot for extraneous cognitive load

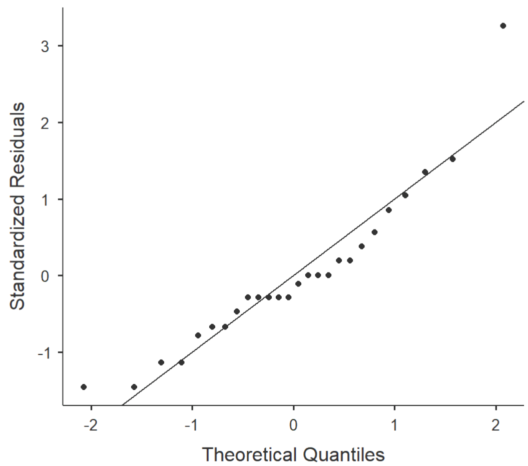


Figure D4

Q-Q plot for germane cognitive load

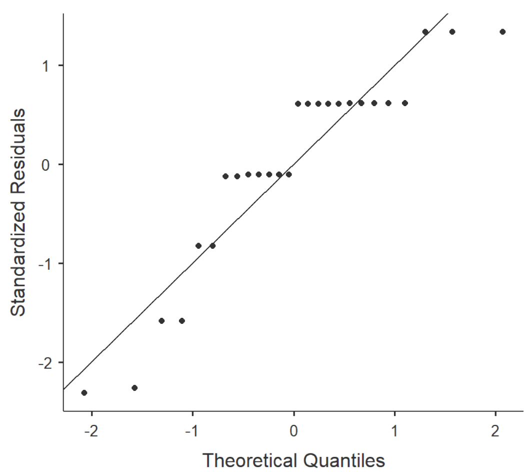


Figure D5

Q-Q plot for mental effort

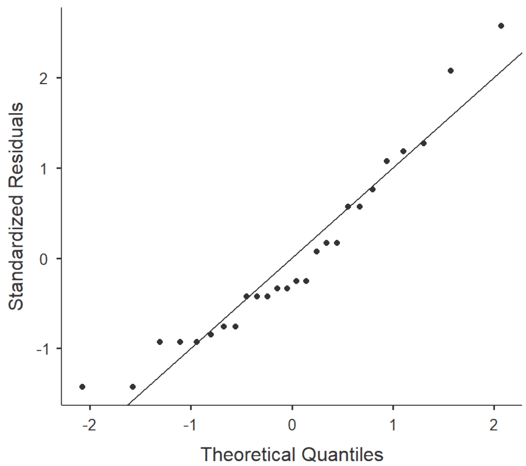


Figure D6

Q-Q plot for intrinsic motivation

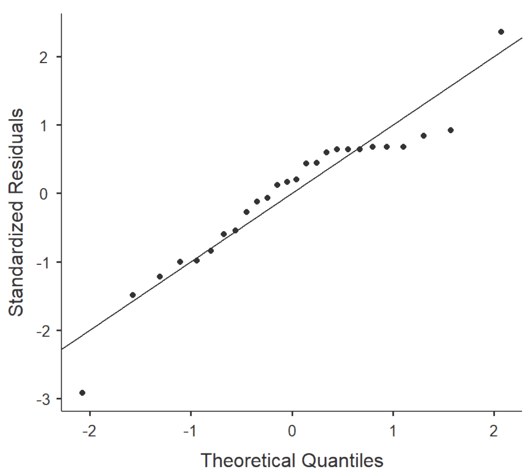


Figure D7

Q-Q plot for perceived usefulness

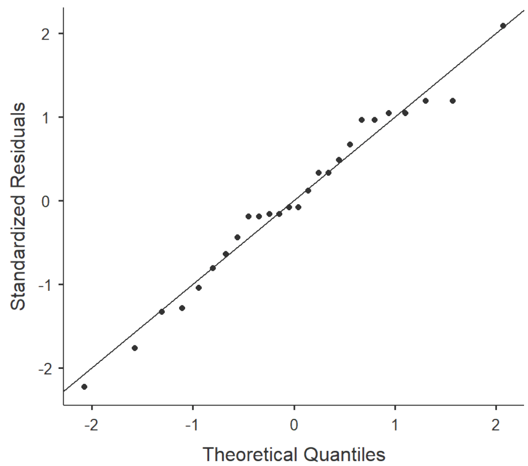


Figure D8

Q-Q plot for perceived ease of use

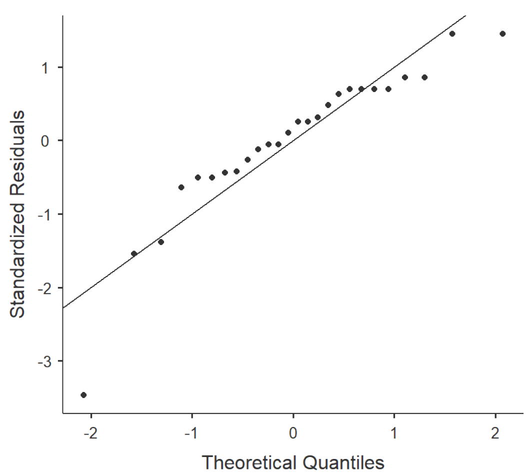


Figure D9

Q-Q plot for behavioural intention