



The challenges of entering the metaverse: An experiment on the effect of extended reality on workload

Nannan Xi^{1,2} · Juan Chen^{3,1} · Filipe Gama¹ · Marc Riar⁴ · Juho Hamari¹

Accepted: 13 January 2022
© The Author(s) 2022

Abstract

Information technologies exist to enable us to either do things we have not done before or do familiar things more efficiently. Metaverse (i.e. extended reality: XR) enables novel forms of engrossing telepresence, but it also may make mundane tasks more effortless. Such technologies increasingly facilitate our work, education, healthcare, consumption and entertainment; however, at the same time, metaverse bring a host of challenges. Therefore, we pose the question whether XR technologies, specifically Augmented Reality (AR) and Virtual Reality (VR), either increase or decrease the difficulties of carrying out everyday tasks. In the current study we conducted a 2 (AR: with vs. without) × 2 (VR: with vs. without) between-subject experiment where participants faced a shopping-related task (including navigating, movement, hand-interaction, information processing, information searching, storing, decision making, and simple calculation) to examine a proposed series of hypotheses. The NASA Task Load Index (NASA-TLX) was used to measure subjective workload when using an XR-mediated information system including six sub-dimensions of *frustration, performance, effort, physical, mental, and temporal demand*. The findings indicate that AR was significantly associated with overall workload, especially mental demand and effort, while VR had no significant effect on any workload sub-dimensions. There was a significant interaction effect between AR and VR on physical demand, effort, and overall workload. The results imply that the resources and cost of operating XR-mediated realities are different and higher than physical reality.

Keywords Mixed reality · Virtual reality · Augmented reality · Metaverse · Workload · NASA Task Load Index

✉ Nannan Xi
nannan.xi@tuni.fi

Juan Chen
lynn2012@163.com

Filipe Gama
filipe.gama@tuni.fi

Marc Riar
marc.riar@tu-berlin.de

Juho Hamari
juho.hamari@tuni.fi

- ¹ Faculty of Information Technology and Communication Sciences, Tampere University, Kalevantie 4, 33100 Tampere, Finland
- ² School of Technology and Innovations, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland
- ³ School of Business Administration, Anhui University of Finance and Economics, Benghu 233030, China
- ⁴ Chair of Information and Communication Management, Technical University of Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany

1 Introduction

Extended reality (XR) technologies are some of the most prominent new developments in information systems, processing, and management (Kim & Hall, 2019; Rauschnabel, 2021; Xi & Hamari, 2020), and offers an umbrella term that covers all forms of virtual and augmented reality technologies, which is often used interchangeably with the term mixed reality (Fast-Berglund et al., 2018; Kwok & Koh, 2021). Among XR-related literature, Augmented Reality (AR) and Virtual Reality (VR) have increasingly been investigated in various areas (Zhang et al., 2017; Pfeiffer et al., 2020; Manis & Choi, 2019; Yim et al., 2017; Klinker et al., 2020) showcasing the impact of this information systems development on a variety of fields. For example, XR has been used to increase the learning and working efficiency of students and employees by providing real-time information (Bednar & Welch, 2020; Lal et al., 2021; Lee, 2012) in education and training, as well as being considered as a way of improving the efficiency of physical rehabilitation

(Afanasiev et al., 2018) in health information systems. In addition, in business, applications such as virtual try-on technologies (Kim & Forsythe, 2008), AR product catalog presentation (Poushneh & Vasquez-Parraga, 2017; Rese et al., 2014) and virtual reality shops (Peukert et al., 2019) provide opportunities for high-efficiency consumption and hedonic experiences. Especially under the global impact of the coronavirus pandemic, XR technologies are inheriting increasingly important roles in social and economic development. According to a report from Technavio (2020), the AR and VR market will accelerate at a compound annual growth rate of over 35% through 2020–2024 considering the impact of COVID-19. However, a significant number of people either anticipate or have determined based on first experiences that there are unsurmountable limitations and hurdles in using XR in terms of e.g., usability, comfort, mental effort, operation and physical interaction. Also, the cost and degree of user challenge is unclear when carrying out these types of virtually-assisted activities in XR-based information systems¹.

In the literature, the challenges and difficulties of using information systems based on XR technologies have been interpreted as the workload that users would experience. The current understanding of the impacts of different XR technologies on workload is still in its infancy. There is a research gap in distinguishing the similarities and differences in the difficulties and costs of operating realities created by XR technologies (e.g., completing tasks and conducting activities). In addition, workload has a multidimensional nature and can be influenced by various factors, for example an individual's motivation, past experience and ability, and the specific characteristics of tasks (Hart, 1982; Meshkati, 1988). A number of XR-related studies have examined the users' workload when using specific devices (Caria et al., 2020; Wang et al., 2019) or methods (Barré et al., 2019; Jost et al., 2020), as opposed to assessing the efficiency and usability of XR technologies in completing tasks from a more general perspective. Furthermore, many studies have failed to observe whether these effects were due to the mediating technology, or to content that was intentionally added or modified. More importantly, granular research on how XR technologies affect different aspects of workload has been slow to emerge, and current literature is limited regarding research on specific dimensions of workload (see e.g., mental workload, Zhao et al., 2017; physical workload, Chihara & Seo, 2018; cognitive workload, Tremmel et al., 2019).

Thus, the *research objective* of this study is to develop an in-depth understanding of what kinds and how much workload users can experience when carrying out information processing tasks in XR environments. Accordingly, the current study design takes the form of a 2 (VR: with vs. without) × 2 (AR: with vs. without) between-subject experiment in a shopping context. A brick-and-mortar record store, as well as a 1-to-1 digital 3D “replica” of the store (using laser scanning and Unity 3D for modeling) were constructed as the experimental environments. The NASA Task Load Index (NASA-TLX) was applied as a subjective measure for evaluating six sub-dimensions of workload. The paper is structured as follows. The extant literature related to XR and workload is discussed for proposing hypotheses in the next section. The research method including participants, materials, measures and procedures is presented in section 2.1. Sections 2.2 and 2.3 present and discuss the results in terms of the effects of XR on overall workload and its sub-dimensions. The conclusion, research implications and practical implications are presented in section 2.3.1. Section 2.3.2 presents the limitations of the current study and future research directions.

2 Background and hypotheses

2.1 Workload

In the field of information systems, individual workload or the lack thereof is often operationalized simply as the general ease of use of an information system in a given task (e.g. ‘freedom from difficulty’ according to the technology acceptance model (TAM): Davis, 1989). While this conceptualization and construct has functioned as a workhorse in the bulk of the technology acceptance literature produced in the last decades (Lah et al., 2020), it provides a limited and one-dimensional understanding of workload and the usability of information systems. Specifically, it does not take into consideration the barriers that would prevent system adoption (see Taylor & Todd, 2001), and it does not reveal how such perceptions are formed (see Mathieson, 1991). In this study, we are interested in the usability of a multimodal information system when used in a rich information processing task. Therefore, this paper ventures to investigate more holistic conceptualizations and constructs of workload. One instrument that is widely used in similar situations where attention and cognitive processes are limited by the context and complexity of the task is the NASA-TLX instrument, which measures the perceived workload of an individual in a given information processing task that can include mental, physical, frustration-inducing, temporal, performance-related and effort-related workload dimensions. Therefore, beyond investigating a novel context of using information systems, this study also builds bridges between fields in the

¹ The unifying factor of all activities utilizing XR is the use of virtualization in a certain form. Steffen et al. (2019) adapted a previous definition of activities, and referred to these as “virtually assisted activities”.

Table 1 Explanation of Each Dimension of Workload in NASA-TLX Based on Hart (2006)

Sub-dimensions of workload	Explanation
Mental demand	Perceived mental and perceptual activity required by an individual to accomplish a given task (e.g., thinking, deciding, calculating, remembering, looking, searching).
Physical demand	Perceived physical activity required by an individual to accomplish a given task (e.g., pushing, pulling, turning, controlling, activating).
Temporal demand	Perceived time pressure due to rate or pace of the given task.
Effort	Perceived level of work (mental and physical) to realize performance level.
Performance	Perceived success in accomplishing the goals that are tied to the performed tasks.
Frustration	Perceived insecurity, discouragement, irritation, stress and annoyance versus perceived security, contentment, relaxation and complacency during task performance.

frontier between adoption and ease of use research, in relation to new technologies.

Workload is largely considered as the effort or cost (e.g., physical, mental, emotional) an individual devotes to accomplish a task (Hart, 2006; Hart & Wickens, 1990). It can be influenced by internal aspects such as an individual's motivation or their past experience and ability, as well as external aspects such as the type, novelty, difficulty, and number of tasks that an individual completes (Hart, 1982; Meshkati, 1988). Originally rooted in attempts to measure the effort of flight-related tasks (see e.g., Li et al., 2020), the practice of assessing workload also became more and more relevant in contexts that transcend aviation. For example, due to the rapid technological advancements and the growing number of novel systems that aim at enhancing, among other factors, convenience, productivity and efficiency, it became increasingly relevant to scrutinize the workload demands of all sorts of information systems. Respectively, today workload is also used to evaluate the interface design of conventional computer systems and portable devices, including the technologies that support virtual and augmented vision (Hart, 2006). The goal often involves gaining a better understanding of how to design and improve systems so that the intended benefits are not compromised by excessive workload during their use. Generally speaking, ergonomists and designers of information technology are interested in creating technology in a way that reduces workload or at least keeps it within an acceptable range (e.g., Grier et al., 2008), as workload management is vital for user acceptance (Dang et al. 2020), productivity, performance, and user health (Jung & Jung, 2001; MacDonald, 2003). An important consideration is that users can only deal with a finite capacity of workload. Kantowitz (1987) describes the concept of spare capacity, which understands that as long as task demands are below an individual's maximum workload capacity, then performance should not be impaired. However, with increasing task complexity or difficulty, the perceived workload intensifies and if the acceptable level is exceeded, performance will suffer. In the light of this rationale, it is not surprising that in the past decades, theories of task-technology-fit

(Goodhue & Thompson 1995) and disciplines devoted to the usability of systems (e.g., Hoehle & Venkatesh, 2015; Lewis, 2014) have garnered great attention in the realm of human-computer-interaction.

The practical necessity to assess the workload involved in human-computer interaction brought forth a number of different evaluation approaches, including objective measures based on performance indicators and psychophysiological cues, as well as measures based on subjective experiences (e.g., Cain 2007; Tsang & Vidulich 2006). Whereas objective measures collect real-time performance data or measure physiological reactions (e.g., via electrodes), subjective measures rely on the self-assessment of the experienced workload by the subjects (Tsang & Vidulich, 2006). A general issue pertaining to the assessment of workload considers the circumstance that different tasks tend to be subject to different sources of workload (e.g., mental and physical), as well as the varying degrees to which each specific source is accountable for an individual's perceived overall workload (i.e., weighted workload) (Hart & Staveland 1988). A weighting scheme aims at measuring workload more accurately, and requires users to evaluate the degree to which different dimensions of workload contribute to the overall workload of a specific task (Hart, 2006). One particular measurement that has been widely accepted and deemed functional to cover the multidimensional nature of workload and capable of accounting for the individual differences of humans with regard to their weighted perceptions of workload is the NASA Task Load Index (TLX) (Hart & Staveland, 1988). This index belongs to the class of subjective measurement instruments, and allows individuals to quantify their experienced workload via a weighted scheme and consists of six dimensions. These include (1) physical demand, (2) mental demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration. Table 1 specifies each dimension in more detail, based on the work by Hart (2006). What should be mentioned here is that a single effort scale (combining physical effort and mental effort) cannot capture the information needed to address the specific source of demands (Hart & Staveland, 1998). Thus, instead of

asking subjects to introspect about the amount of mental or physical effort exerted, the NASA TLX instrument requires them to assess the objective physical and mental demands that are placed on them (Hart & Staveland, 1998).

2.2 Extended reality

Even though there seems to be a lack of consistency in the use of reality-related terms (e.g., Virtual Reality-VR, Augmented Reality-AR, Mixed Reality-MR, and Extended Reality-XR) in academic and professional fields (Flavián et al., 2019), AR and VR have been considered as the two core reality-virtuality technologies. With the development of multi-sensory technologies and modalities and the deepened conceptualized understanding of AR and VR, there is a consensus that any sensory experience can be augmented in a digital way (Harley et al., 2018) and also be virtualized (Boyd & Koles, 2019). In terms of AR, multimodal information such as smell, touch, taste and sound can be digitally overlaid on the current world (Azuma, 1997; Carmigniani et al., 2011; Riar et al., 2021) and AR users are not isolated from it (Rauschnabel, 2021). AR has been defined as the term for technologies for augmenting or altering the current reality (Riar et al., 2021), while in VR, all of the sensory information and stimulus of the ‘real reality’ is rather blocked and inhibited (Manis & Choi, 2019; Yim et al., 2017). Therefore, while VR has been considered as the digital technologies of choice for substituting the perceived reality (Xi & Hamari, 2021), AR and VR provide different kinds of experiences to users (Fromm et al., 2021). Regarding AR, the “augmenting” information and content can bring users interactivity, vividness and novelty (McLean & Wilson, 2019; Yim et al., 2017). Alternatively, VR has been believed to create immersiveness (Suh & Prophet, 2018), telepresence (Lee & Chung, 2008; Steuer, 1992), and the sense of “being there” (Heeter, 1992; IJsselsteijn & Riva, 2003). As a further extension, when combining AR and VR together, Augmented Virtuality (AV) can be constructed for a more hybrid experience. There are high expectations towards AR and VR on creating interesting, novel and playful experiences (Lin & Yeh, 2019; Raptis et al., 2018), however, an increasing number of studies have shown that the use of AR and VR in activities and completing tasks requires various resources and costs.

2.3 Workload in extended realities

The current study assumes that both AR and VR technology will lead to a high overall workload compared with non-technology mediated reality, but will influence different sub-dimensions of workload. In the following, we describe how AR and VR are expected to influence the particular dimensions of workload, and derive them according to hypotheses.

2.3.1 Mental demand

On one hand, XR technologies have been found to offer perceptual and cognitive benefits such as supporting cognitive processing, mental elaboration and imagery, by providing visual cues that are lacking from the physical environment (e.g., Bogicevic et al., 2019; Fan et al., 2020; Heller et al., 2019; Park & Yoo, 2020), as well as by offering simulating environments and situations that would otherwise be costly or difficult to produce in the real world (e.g., Barré et al., 2019; Clifford et al., 2018). On the other hand, however, it seems that these benefits often come at the cost of additional mental effort. For example, previous studies point out that using AR can result in a higher cognitive load and mental strain (e.g., Dunleavy et al., 2009; Tarafdar et al., 2019; Weidinger et al., 2018). It seems that especially in AR, individuals may be confronted with a more challenging perceptual task in processing augmented reality simultaneously with the physical reality, as the different visual sources may demand more mental effort from individuals. According to multiple resource theory, the mental workload is affected by humans’ limited capacity to deal with several sources of attention simultaneously (Wickens, 2008). It is likely that in AR, individuals expend mental resources towards interpreting both reality-based and augmented reality-based spatial information (and possibly auditory or other multimodal information), while also having to deal with potential extraneous distractions.

In comparison, we can expect that in a pure VR setting, individuals only have to mentally process the virtual reality, and hence it should be similar to exclusively processing the perceived current reality. Individuals may perceive their surroundings through the VR interface as being natural, and therefore they do not have to exert additional cognitive effort (Wickens, 1992). Moreover, VR has the ability to immerse users into the mediated environment so that they feel “present” in the virtual world (Schuemie et al., 2001; Steuer, 1992). Through this sense of presence in the virtual environment, extraneous cognitive load and distractions may be blocked, and consequently, there is no divided attention between the real world and the virtual world. Ultimately, it seems that when compared to the perceived current reality and VR, individuals have face greater perceptual challenges in AR since the visual cues coming from multiple resources (i.e., the perceived current reality and the augmented reality) have to be combined. Thus, we expect that for AR, the level of mental demand will be higher compared to purely physical or VR environments, since in the latter conditions, individuals only have to process a single reality and are not confronted with various sources of multimodal information. This also implies that there will be no significantly different experiences in the mental workload between fully virtual reality (as well as the interaction of AR and

VR) and the perceived current reality. Such assumptions are in line with prior studies which indicate that VR does not cause additional mental workload compared to purely physical situations (without VR) (e.g., Chao et al., 2017), and that AR entails more mental workload in comparison to situations without AR. Therefore, we propose the following hypothesis:

H1: a) AR has a positive effect, while b) VR has no effect on mental demand.

2.3.2 Physical demand

In addition to mental demand, prior studies postulate that using XR technologies can result in higher degrees of physical exertion (e.g., Chao et al., 2017; Dickinson et al., 2020; Madathil & Greenstein, 2017; Millais et al., 2018). This can mostly be explained by the fact that when individuals perform tasks in the perceived current reality, they are used to the involved movements and can execute them naturally and effortlessly, whereas in AR and VR, the same tasks often require more demanding physical effort because individuals have to use specific hardware and controls generally involve movement-based mechanisms. The physical workload is caused by movements such as pulling, pushing, pressing, reaching, holding, and so on (Hart et al., 1981), many of which are movement patterns required to control XR technologies. There seems to be ample support for the notion that XR can increase physical effort, with previous studies proposing that XR requires more vigorous interactions compared to other less interactive technologies, and that the controls such as movement, pointing and selecting can result in higher degrees of physical burden (Chao et al., 2017; Madathil & Greenstein, 2017; Millais et al., 2018). Given the above evidence, the second hypothesis states that:

H2: a) AR has a positive effect and b) VR has a positive effect on physical demand.

2.3.3 Temporal demand

Temporal demand is related to perceived time pressure, which is the perception that there is an inadequate amount of time available to carry out the given tasks (Guo et al., 2020). Time pressure has been considered as the need for time allocation, i.e. allocating time between competing options (Guo et al., 2020). Prior knowledge on the psychology of time pressure underlines the influence of increased work intensity and multitasking (Szollos, 2009). We presume that by being exposed to multimodal features and also handling different realities at once, individuals may be more inclined to feel agitated and rushed, experiencing higher temporal demand through the intensified work demands. For example, users in

augmented reality have to synchronously process information from different channels (e.g., digital information being displayed on an AR device and information coming from the physical environment). Therefore, in AR, users may experience a higher temporal demand compared to the non-AR enabled reality.

While VR users might feel stressed from unnaturally interacting with digital objects and the environment with devices while performing tasks, the perception of time pressure might be the same as in an environment without VR (e.g., the physical world). In the purely virtual world (see “magic circle”: Castronova, 2008), users have opportunities to conduct activities that are typically either not done well or safely in real life without concerns for the consequences involved (Schultheis & Rizzo, 2001). They can further have a higher self-efficacy to complete tasks (Shu et al., 2019) without following the rules, which reduces the pressure of time allocation. For instance, in some studies related to VR shopping tasks, users neither need to take virtual products back to shelves nor consider whether the product could be damaged without careful interaction. Therefore, they are more likely to feel that they have enough time to complete the given tasks. However, an “offset effect” might occur when considering the additional tasks brought on by manipulating in virtual reality due to the limitations of VR technology, and less task-related demands from “magic circle” nature of VR, which together lead to the same temporal demands as seen in environments without VR. Given the above, we propose that:

H3: a) AR has a positive effect, while b) VR has no effect on temporal demand.

2.3.4 Effort, performance and frustration

Based on the assumption that XR can generate both higher mental and physical demands, and considering the notion that perceived effort is caused by the level of work (e.g., mental and physical) that an individual has to expend in order to reach a certain level of performance (Hart, 2006), it seems a valid assumption that engagement with XR will also necessitate greater effort compared to task performance in non-XR conditions. Moreover, compared to the smooth interactions that we can experience with physical environments and physical objects, the interaction with virtual environments and objects in XR can be perceived as being more strenuous and frustrating. Reasons for this may include technological insufficiencies such as an unsatisfactory mapping quality of virtual content (Poushneh, 2018), outages, or other malfunctions which can increase frustration (Ahsen et al., 2019). A lack of intuitiveness and deficiencies with the modalities for interacting with virtual objects can also lead to increased levels of frustration

Table 2 The 2 (VR: with vs. without)×2 (AR: with vs. without) Between-Subject Experiment Design

No.	VR	AR	Value*	Name	Reality	Information	Device
1	Without	Without	0 0	Non-XR shop	Physical	Traditional	N/A
2	Without	With	0 1	AR shop	Physical	Superimposed	HoloLens
3	With	Without	1 0	VR shop	Virtual	Traditional (digital)	Valve index & controllers
4	With	With	1 1	AV shop	Virtual	Superimposed	Valve index & controllers

* Operationalization as dummy variable for VR and AR

(Whitlock et al., 2018). Generally speaking, prior literature indicates that there can be certain detriments involved with using XR technologies, such as difficulties in controlling virtual objects as effortlessly as in a physical environment, as well as increased perceptions of stress, discomfort, disorientation, and especially in VR, the problem of cybersickness (Munafò et al., 2017; Sharples et al., 2008; Somrak et al., 2019; Weech et al., 2019). These impediments (both technological and in terms of the user experience) can increase the required effort and frustration levels of individuals when using XR technology. In line with the understanding of the effort and frustration dimensions stemming from workload theory which entail higher degrees of expended work to achieve performance levels, as well as higher levels of experienced stress, irritation, annoyance, and discouragement during task completion (Hart 2006), we hypothesize that both AR and VR can lead to higher perceptions of effort (H4a/H4b) as well as frustration (H6a/H6b).

It should be noted that this does not necessarily mean that individuals' performance will suffer. Rather, users may have to expend more effort to achieve the same level of performance in XR compared to the physical reality. According to theories of workload, performance may stay consistent even if individuals have to expend higher mental or physical effort, as long as the task demands do not reach a critical limit (i.e., the maximum capacity of workload) (Grier et al., 2008; Kantowitz, 1987). Prior studies also indicate that XR can in fact increase performance (e.g., Chang & Hwang, 2018; Kim et al., 2018). However, this seems more likely if the XR environment provides some added value (e.g., in terms of extra information, content, better interaction, enhanced interface features, and a multi-modal experience). If no such added value is provided, the nature of VR only substitutes the perceived current reality and the essence of AR only changes the way information is presented. Thus, by default, the augmentations do not provide any enhanced experience beyond what the physical reality has to offer, and it seems unlikely that performance gains can be achieved. In line with these arguments, we presume that there will be no significant difference in performance between executing tasks within XR and physical realities. Thus, we derive the following hypotheses:

H4: a) AR has a positive effect and b) VR has a positive effect on effort.

H5: a) AR has no effect, and b) VR has no effect on performance.

H6: a) AR has a positive effect, and b) VR has a positive effect on frustration.

3 Research method

3.1 Research Design

To investigate how the six dimensions of workload differ in different extended realities, we conducted a 2 (VR: with vs. without) × 2 (AR: with vs. without) between-subject experiment with a psychometric survey (see Table 2). The shopping task featured is usually involved with a series of holistic mental, cognitive and behavioral costs such as information searching, processing, interactive and movement-based behavior, time pressure and decision making, and thus perfectly matches the goal of a workload-related study. Thus, a shopping scenario was designed as the experiment context and participants were required to complete a 10-minute task. Each participant was given a 10 euros gift card for buying products in the shop we designed. The level of compensation depended upon the products each participant selected (they could keep the selected products). The purpose of such a compensation scheme was to garner a realistic context of decision making in the shop as the compensation depended upon the rational choice of the participant. The study design and procedure adhered to the Finnish National Board on Research Integrity TENK Guidelines 2019.

3.2 Participants

From September to November 2019, a total of 165 student participants were recruited from a Finnish university to participate in a shopping-related experiment (see section 3.4 Procedure). Participants were randomly assigned to one of four groups. Two participants were omitted due to their inability to use the VR headset to complete the experiment²

² One participant wore a head garment that prevented her wearing a headset, and one participant wore eyeglasses that were incompatible with the VR HMD lenses.

Table 3 Participant Characteristics (Mean and Standard Deviation)

Measure	Group 1 (<i>n</i> = 40)	Group 2 (<i>n</i> = 41)	Group 3 (<i>n</i> = 40)	Group 4 (<i>n</i> = 36)	<i>F</i> (3, 153)	<i>p</i>
Gender	67.50 %	46.34 %	50 %	52.78 %	1.396	.246
Age	3.58 (1.40)	3.24 (0.80)	3.33 (1.02)	3.11 (0.67)	1.426	.237
Education	1.63 (0.67)	1.44 (0.59)	1.55 (0.64)	1.36 (0.54)	1.388	.248
Income	1.87 (1.34)	1.68 (1.00)	2.05 (1.53)	1.32 (0.83)	2.220	.088
Height	5.93 (1.59)	5.51 (1.75)	5.30 (2.16)	5.47 (1.65)	0.859	.464
VR experience	1.98 (1.10)	1.95 (1.00)	2.10 (1.01)	1.61 (0.84)	1.651	.180

Gender was considered as a binary variable; the percentage of males in each group is presented in the table; Age was measured from 1 = less than 15 years old, 2 = 15-19, 3 = 20-24 to 11 = 60 or older; Education was measured from 1 = bachelor student, 2 = master student and 3 = PhD student; Income (pre-tax) was measured from 1 = 0-499 euros, 2 = 500-999 euros to 9 = 4000 euros or more; Height was measured from 1 = 150 cm or less, 2 = 151-155 cm, 3 = 156-160 cm, to 8 = Above 180 cm; VR experience was measured by the frequency of using VR devices, ranging from 1 = never, 4 = sometimes, to 7 = every day

and one disqualified participant was also dropped. Table 3 provides the detailed demographic information of participants including gender, age, education, study area, income, and nationality. The final sample size was 157 (54.14% male, 56.05% 20-24 years old, 56.69% undergraduate students, and 56.69% less than 499 euros monthly income)³. The participants had diverse nationalities, coming from Finland, China, Germany, Vietnam, Spain, Russia, India, France, and elsewhere.

3.3 Materials

3.3.1 Shop

A physical LP record shop (4.24m×5.09m = 21.58m²) was built on the university campus, and functioned as the experimental setup for each condition. The shop floor plan, size, decoration and layout were identical in each condition. Fifty-four LP records were displayed on three different walls of the room, and each wall was equipped with three layers of shelves. For the Group 1 which is designed to control other groups, the shop functioned as a common style physical shop and the product information was presented on printed papers attached to the shelves. For Group 2, the shop functioned similarly except that the product information was displayed through an AR headset display (Microsoft HoloLens: see section 3.3 *Apparatus*). For Group 3, the same room was used. The physical shop was scanned for 3D reconstruction by using an active scanning method, with a LiDAR sensor.

Then the shop was modeled and refined in Blender with 1:1:1 scale and then imported to Unity. Thus, the shopping environment used in the “control” condition was reconstructed in virtual reality (e.g., textures, lighting conditions, geometry and “attached” product information). For Group 4, the environment combined the condition in Group 3 with the superimposed product information as in the condition of Group 2. See Fig. 1.

3.3.2 Product and its information

Purchasing second-hand English LP records (0.314m×0.314m, price categories of 3, 6 and 9 euros) was deemed as a suitable activity for the study as it represents a human activity that requires multimodal, multi-sensory interaction and effort. Moreover, students are the main consumers of music-related products. The average expenditure on a music hobby was 81.81 euros per year among the 157 participants in the current study. However, we wished to minimize any participant familiarity with specific artists/bands on display among the records in the shop. Therefore, all records in the shop were acquired second-hand and produced before the 1990s. To confirm our expectations of previous product knowledge, in the post-survey, we measured previous product knowledge by seven items based on 7-point Likert scale that were adapted from Awasthy et al. (2012), and the results indicate that the participants were not familiar with the products featured in the shop ($M = 2.58$, $SD = 1.13$). In addition, LP-record packaging is efficient to be modeled to match the physical versions of the products and the interactive quality for a computer-based environment. In the VR shop and AV shop, the front and back covers of the LP records were scanned in high quality and added as 3D objects into the shopping programs.

The product information was gathered from the website Discogs. The general information (e.g., label,

³ Five outliers were identified and dropped for MANOVA and ANOVAs.

Fig. 1 *Design of Four Shops in The Study. Note.* In the Group 2, Microsoft HoloLens 1 was used to present “pop-up” information; Valve Index VR headset was used in both Group 3 and 4. See open-access video link <https://cutt.ly/XR-shopping>

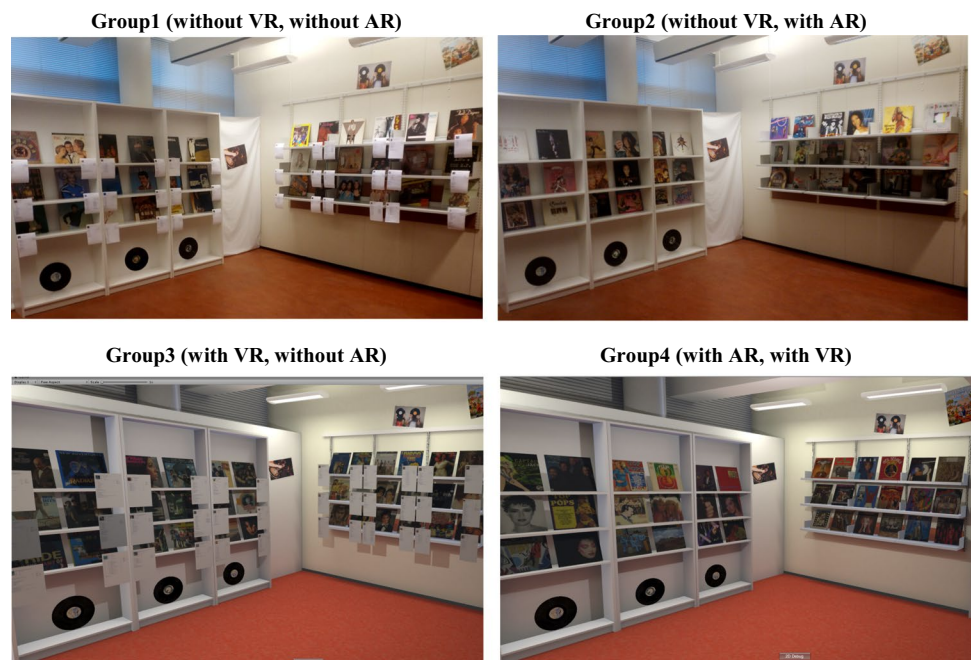
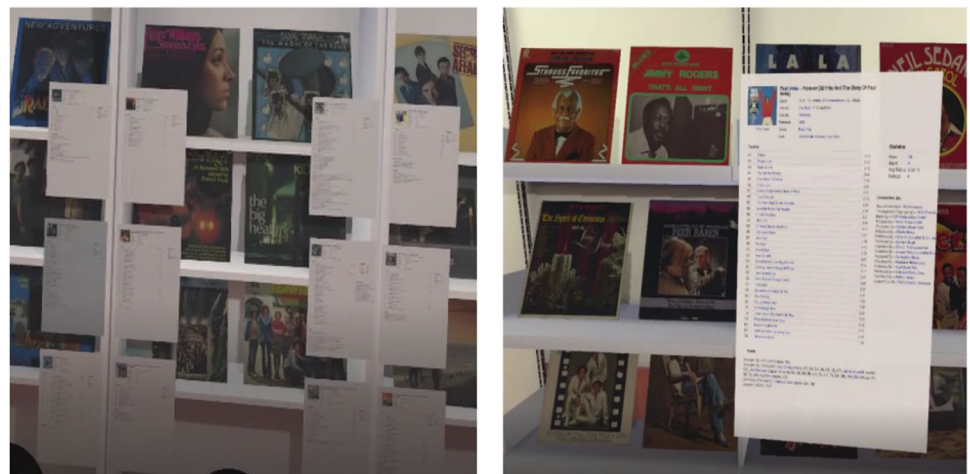


Fig. 2 *Non-augmented (Group 1 & 3) vs Augmented (Group 2 & 4): Participant's View*



format, country, released year, genre and style), track list, statistics, companies, and credits information were selected by the researchers. In the conditions without AR, all of the information pages were pasted either physically (Group 1) or virtually (Group 3) onto the edge of the shelf next to each record. In the conditions with AR (Groups 2 and 4), the corresponding information about the LP record was displayed superimposed on the environment based on the image recognition algorithm from the Vuforia Engine. In Group 4, similar to the shop for Group 2 but without using image recognition, the head position of the participant was tracked to predict which record the participant was looking at (see Fig. 2).

3.3.3 Apparatus

Microsoft HoloLens In order to superimpose a computer-generated image with the LP record information on the users' view of the real world and ensure flexible interactions with the products by hand (as in other treatments), the Microsoft HoloLens version 1 AR headset was used for Group 2. Microsoft HoloLens is equipped with a 1280×720 display resolution (per eye) and weighs 579 grams. The digital information only “popped up” whenever the participant was looking directly at the same LP record (front cover) for more than 2 seconds. The straight-line distance between the participant's head position and the targeted record was between 1 to 1.5 meters.

Valve Index For Group 3 and 4, the participants used the Valve Index headset and its controllers. The Valve Index headset is a fully-immersive headset weighing 809g and is equipped with dual 1440×1600 RGB LCDs, runs at 120Hz, and covers around a 130 degree field-of-view. The Valve Index is accompanied by two controllers (one for each hand) that enable the haptic interface to manipulate records naturally (as in real life) via 87 sensors used to track hand and finger positions.

3.4 Measures

In the present study, the independent variables were composed of the dummy-coded variables of AR (without = 0, with = 1) and VR (without = 0, with = 1). To examine the workload during shopping in extended realities, we employed the NASA-TLX (Task Load Index) (Hart & Staveland, 1988). The NASA-TLX has been used in many empirical studies for subjective workload assessment (see Longo, 2018; Dang et al., 2020). Initially developed at NASA where it was tested and refined in more than forty simulations, it is regarded as a valid and highly reliable source for measuring workload. It has been continually employed and further validated in different contexts for more than thirty years. Due to its high validity, we adopted the NASA-TLX to assess workload, and the procedure is as follows. The first step is to divide each dimension of workload from 0 to 100 in increments of 5. Performance should be given a score between 0 to 100 from good to poor while the other five aspects should be given a score between 0 to 100 from low to high (Hart & Staveland, 1988). Specific to this study, participants were required to rate the six factors of the workload of their shopping separately, namely, *mental demand* (how mentally demanding was the shopping task e.g., thinking, deciding, calculating, remembering, looking, searching?), *physical demand* (how physically demanding was the shopping task e.g., walking, picking, turning, controlling?), *temporal demand* (how hurried or rushed was the pace of the shopping task?), *performance* (how successful were you in accomplishing what you were asked to do in the shopping task?), *effort* (how hard did you have to work mentally and physically to accomplish your level of performance in the shopping task?), and *frustration* (how insecure, discouraged, irritated, stressed and annoyed were you in the shopping task?), after the shopping task was complete. The second step is to obtain a weighted score for each workload factor for calculating the overall workload. Instead of pairwise comparisons, in the current study, participants proceeded to rank the six factors from 0 (least important) to 5 (most important), considering their contribution to the total load of their shopping experience. Thus, for each participant, the weighted score for each dimension was obtained by multiplying the rating score by the corresponding weight. The

sum of the weighted ratings was then divided by 15 (the sum of the weights) for an overall workload score.

3.5 Procedures

3.5.1 Recruitment

We described the research purpose as “a study of shopping experience” without disclosing any experiment details in the online and offline advertisements posted in a higher education university in Finland. From September to November in 2019, 265 students successfully entered our recruitment system by scanning a QR code or opening the link presented on our advertisement page or flyer (see *Appendix A*). They were then linked to the online self-booking system on Doodle. A total of 165 students participated in the final experiment.

3.5.2 Pre-survey

All participants were randomly assigned to join one of the four groups. Before starting the shopping task, researchers first introduced the entire experiment procedure and guided participants to read the consent form for the study. Then, participants filled out a pre-survey related to the relevant information about music products.

3.5.3 Tutorial

After the pre-survey, participants got to practice before they started shopping. *Experimenter 1* guided the participant to the shop room and introduced the experiment procedure step by step, according to the instruction page (see *Appendix B*). For tutorial purposes, the same sample LP record and product information were presented to participants in all shopping conditions.

For Group 2, the experimenters introduced the Microsoft HoloLens and guided participants on how to wear and use the headset to recognize the physical sample LP record displayed outside of the shop room. Two tutorial shopping programs were developed for both Group 3 and Group 4 respectively, without revealing any details of the shops. *Experimenter 2* was responsible for the tutorial program, and *Experimenter 1* guided the participant on how to wear a VR headset, move in the virtual space, read information, and interact with the record. The interpupillary distance (IPD) and eye relief for each participant were adjusted carefully during the tutorial.

Each participant was also told that if he/she needed any help or felt uncomfortable during shopping, a short break could be offered during the experiment, and they were welcome to request additional breaks. After ensuring the participants understood the shopping tasks, *Experimenter 1*

Table 4 Descriptive Statistics of Overall Workload in Each Group

Group	VR	AR	N	M	SD	Shapiro-Wilk	
						Statistic	p
1	Without	Without	40	26.308	17.689	0.939	.031
2	Without	With	41	40.081	20.861	0.973	.423
3	With	Without	40	34.150	16.076	0.973	.453
4	With	With	36	34.417	15.238	0.964	.288

Table 5 ANOVA with Overall Workload As Dependent Variable, AR and VR as Independent Variables

Independent variables	SS	df	MS	F	p	Partial η^2
VR	46.389	1	46.389	.148	.701	.001
AR	1929.286	1	1929.286	6.173	.014	.039
VR * AR	1785.492	1	1785.492	5.713	.018	.036

Type III Sum of Squares; SS = Sum of Squares; MS = Mean Square

instructed participants to a) enter the shop room (Group 1 and 2) or b) wait for the program to start (Group 3 and 4), and then start shopping.

3.5.4 Shopping task

A pilot study ($N = 20$) was conducted for all conditions to test the measurement items in the pre-survey and post-survey, as well as the experimental procedure, instruction, shopping program apparatus, and methods. In the actual experiment, participants were asked to spend 10 minutes in the shop and make their purchase decision independently and in accordance with their own preferences. The experimenter played the role of “cashier” in the four shopping conditions. When the shopping time had run out, the “cashier” counted the price of each record product on the check-out table and asked participants to pay with the gift card they had been given. When the shopping was completed and after making sure that the subjects had no physical problems, participants were guided into another room and filled out the post-survey. The 1-3 records that were “bought” by the participants were then restocked with new records of the same price category in the shop.

4 Data analyses and results

First, we calculated the overall workload by combining weights and ranks of the sub-dimensions according to the NASA-TLX scale (see section 3.3.1 *Measures*). Given that AR (with = 1, without = 0) and VR (with = 1, without = 0) were considered as the two categorical independent variables and overall workload as the (continuous) dependent variable, a between-subjects two-way analysis of variance (ANOVA) was employed. We then conducted a more

granular analysis on the effects of AR and VR on each sub-dimension of workload for testing the proposed hypotheses. A two-way multivariate analysis of variance (MANOVA) was conducted to help protect against inflating the Type I error rate in the follow-up two-way analysis of variance (ANOVAs) and post-hoc comparisons.

4.1 Overall workload

The descriptive statistics related to overall workload among the four groups are reported in Table 4 ($N = 157$). It can be seen that the overall workload in the XR-mediated groups (Group 2&3&4) were higher than in the non-XR group (Group1). In addition, Group 2 had the highest overall workload ($M = 40.081$) and Group 1 had the lowest overall workload ($M = 26.308$). In order to test whether AR and VR have an effect on workload, a two-way ANOVA was employed. Even though the normality in one group cannot be assumed according to Shapiro-Wilk results, the two-way ANOVA is reasonably robust to violations of normality when group sizes are similar. Furthermore, the assumption of a homogeneity of variances was tested and satisfied based on Levene’s F test, $F(3, 153) = 1.576, p = .197$.

The results reveal that the groups with AR ($M = 37.249, SD = 2.019$) had a higher overall workload compared to those groups without AR ($M = 30.229, SD = 1.976, F(1, 153) = 6.173, p < .05$), and AR explains 3.9% of the variance in the dependent variable. On the other hand, the groups with VR ($M = 34.283, SD = 2.031$) did not reveal any significantly higher workload in comparison to those groups without VR ($M = 33.195, SD = 1.96, F(1, 153) = 0.148, p > .05$), and VR only accounts for 0.1% of the variance in workload. The results also show that there was a significant interaction between AR and VR ($F(1, 153) = 5.713, p = .018$) and the interaction explains 3.6% of the

Table 6 Pairwise Comparisons on Overall Workload

	(I)	(J)	MD (I-J)	SE	p	95% CI	
						Lower	Upper
Without AR	Without VR	VR	-7.842*	3.953	.049	-15.651	-.032
AR	Without VR	VR	5.665	4.038	.163	-2.312	13.642
Without VR	Without AR	AR	-13.773**	3.929	.001	-21.535	-6.011
VR	Without AR	AR	-.267	4.061	.948	-8.290	7.757

* $p < .05$, ** $p < .01$, *** $p < .001$; MD = Mean Difference; SE = Std. Error

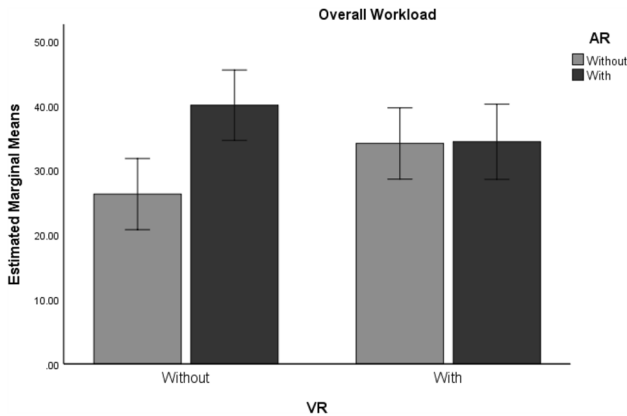


Fig. 3 Interaction Effect of AR and VR on Overall Workload

variance (see Table 5). To interpret the interaction effect, a pairwise comparison was conducted (Bonferroni) using the EMMEANS syntax command within SPSS (see Table 6 and Fig. 3). The results show that only when participants were in the groups without AR, did VR ($M = 34.150$) lead to significantly higher workload than without VR ($M = 26.308$, $p = .049$). Similarly, when participants were in the groups without VR, AR ($M = 40.081$) led to significantly higher workload than those without AR ($M = 26.308$, $p = .001$).

4.2 Sub-dimensions of workload

A series of Pearson correlations were performed among the six dependent variables in each group for testing the MANOVA assumption. The results reveal no concern for multicollinearity, although a few variables were not highly correlated. In addition, the *Box's M* value of 91.254 was associated with a p-value of .035, which was interpreted as non-significant based on the Hahs-Vaughn (2016) guideline ($p < .001$). The covariance matrices between the groups were assumed to be equal for the purpose of MANOVA. In addition, the absence of multivariate outliers was checked by assessing the Mahalanobis Distances among the participants (the highest value was 22.031 and less than 22.46). A two-way MANOVA was conducted to test the hypothesis that AR and VR have different effects on the various sub-dimensions of workload. A statistically

significant MANOVA effect was obtained for the interaction between AR and VR, where Pillai's Trace = 0.097, $F(6,148) = 2.643$, $p = .018$. The multivariate effect size was estimated at .097, which implies that 9.7% of the variance in the dependent variables was accounted for by the interaction between AR and VR.

4.2.1 Hypothesis test

Prior to conducting the ANOVA tests, the homogeneity of variance assumption was tested for all six dependent variables. Based on a series of Levene's F tests, the homogeneity of variance was considered as satisfied, as all six Levene's F tests were insignificant ($p > .05$) - see Table 7. A series of two-way ANOVA's on each of the six dependent variables was conducted as follow-up tests to the MANOVA. As can be seen in Table 7, AR had a significant effect on mental demand ($F(1, 153) = 5.084$, $p = 0.026$, Mean difference = 9.597) and effort ($F(1, 153) = 6.500$, $p = 0.012$, Mean difference = 9.619), while all of the ANOVA's for VR were statistically insignificant (all p -values $> .05$). Accordingly, regarding the effects of AR on each dimension of workload, $H1a$, $H4a$ and $H5a$ were supported, and $H2a$, $H3a$ and $H6a$ were rejected. Regarding the effects of VR on each dimension of workload, $H1b$, $H3b$ and $H5b$ were supported while $H2b$, $H4b$ and $H6b$ were rejected. In addition, the interaction effect between AR and VR significantly influenced physical demand, $F(1, 153) = 11.405$, $p = .001$ and effort, $F(1, 153) = 5.647$, $p = .019$.

4.2.2 Additional analysis

Finally, a series of post-hoc analyses were performed to examine individual mean difference comparisons across different levels of AR and VR and the two subscales of workload (see Fig. 4 and Table 8). According to the pairwise comparisons (Bonferroni) for groups without AR, VR led to a significant increase in physical demand. In conditions with AR, VR led to a significant decrease in physical demand. Similarly, in the groups without VR, AR significantly led to high physical demand and in the conditions with VR, AR decreased physical demand, albeit insignificantly ($p = .052$).

Table 7 Two-Way ANOVAs with Workload Subscales As Dependent Variables, AR and VR As Independent Variables

	Levene's		AR			VR			AR*VR		
	<i>F</i> (3, 153)	<i>p</i>	<i>F</i> (1, 153)	<i>p</i>	Partial η^2	<i>F</i> (1, 153)	<i>p</i>	Partial η^2	<i>F</i> (1, 153)	<i>p</i>	Partial η^2
Mental demand	0.891	.447	5.084	.026	.032	0.022	.883	.000	0.583	.446	.004
Physical demand	2.562	.057	209.950	.310	.578	0.005	.945	.000	11.405	.001	.069
Temporal demand	1.014	.388	2.264	.134	.015	0.205	.652	.001	2.588	.110	.017
Effort	2.670	.050	6.500	.012	.041	0.139	.709	.001	5.647	.019	.036
Performance	2.567	.057	1.098	.296	.007	0.503	.479	.003	0.469	.495	.003
Frustration	2.087	.104	3.659	.058	.023	0.161	.689	.001	2.029	.156	.013

In addition, only in the groups without VR did AR lead to a significant increase in effort compared to those without AR.

5 Discussion

Overall, users experienced higher workload in all three XR-mediated realities than non-XR. To be more specific, AR was significantly associated with overall workload,

especially its sub-dimensions of mental demand and effort. In this study, AR was operationalized as “superimposed information” (see Table 2). It is reasonable that when simultaneously processing information from multiple channels and different realities, the cognitive resources users usually spend is higher (Xi et al., 2021). For example, participants in Group 2 (with AR, without VR) had to seek for and process both the digital information presented by the AR headset, and also the information coming from physical products and shopping environment. This high cognitive load was associated with high mental demand and effort.

Additionally, the results also show that using VR did not cause increased perceptions of workload for any of the sub-dimensions, which served to reject the proposed hypotheses that VR has a positive effect on physical demand, effort and frustration. These unexpected results might be explained by the “non-consequentiality” feature of VR. VR has been believed to provide opportunities for a variety of activities that are typically either not done well or safely in real life (Schultheis & Rizzo, 2001; Reid, 2004). Therefore, “non-consequentiality” refers to the feature that the VR-mediated environment could be seen as free of many of the constraints and consequences that govern activities in the real world, and so transforms the environment into playgrounds of free experimentation (Xi & Hamari, 2021). Thus, users can experience a relatively high degree of safety and freedom without concerning any adverse consequences such as physical and psycho-social risks when moving and operating with objects in virtual reality. Such a high degree of freedom without the necessity of concern of consequences might be associated with a low perception of workload. On the other hand, it can also be assumed that there might be yet unmeasured psychological variables or other unexplored features of VR that influence the assessment of difficulties related to using VR for completing certain tasks such as physical demand, effort and frustration.

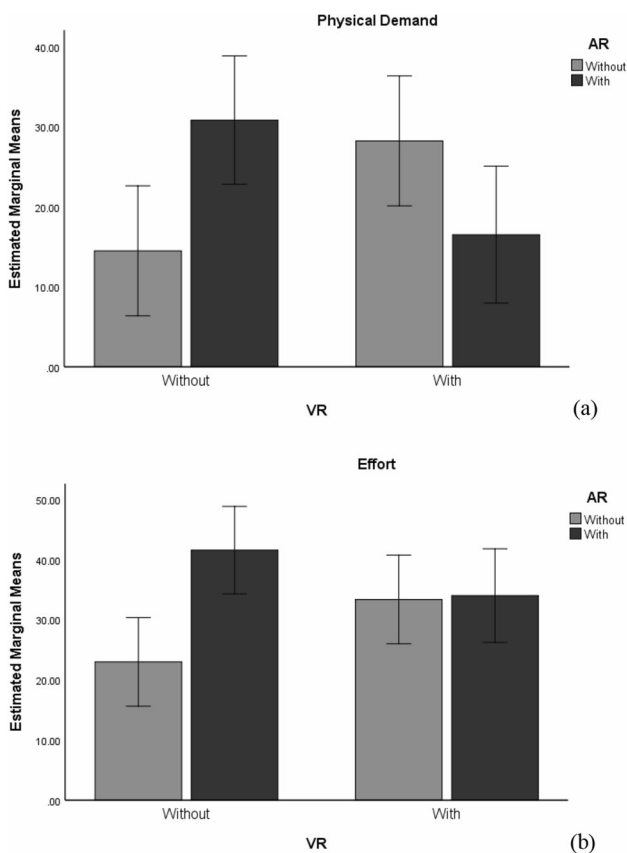


Fig. 4 Interaction Effect of AR and VR on Physical Demand and Effort

Table 8 Pairwise Comparisons—Physical Demand and Effort

		(I)	(J)	MD (I-J)	SE	p	95% CI	
							Lower	Upper
Subscale: <i>Physical demand</i>								
Without AR	Without VR	VR		-13.750*	5.816	.019	-25.240	-2.260
AR	Without VR	VR		14.326*	5.941	.017	2.590	26.062
Without VR	Without AR	AR		-16.354**	5.780	.005	-27.773	-4.934
VR	Without AR	AR		11.722	5.975	.052	-.082	23.527
Subscale: <i>Effort</i>								
Without AR	Without VR	VR		-10.375	5.279	.051	-20.804	.054
AR	Without VR	VR		7.558	5.392	.163	-3.095	18.210
Without VR	Without AR	AR		-18.585**	5.247	.001	-28.951	-8.220
VR	Without AR	AR		-.653	5.424	.904	-11.367	10.062

* $p < .05$, ** $p < .01$, *** $p < .001$; MD = Mean Difference; SE = Std. Error

The interesting findings of the study can be observed in the interaction effects between AR and VR on physical demand, effort and overall workload. Specifically, in terms of physical demand, the combination of AR and VR had a lower physical demand compared to using AR or VR alone. In the present study, the two XR technologies might be associated with physical demand in different aspects: information detection in AR vs. information reading⁴ in VR (low readability - see Peukert et al., 2019); while such negative aspects of AR and VR might be eliminated when the two technologies were combined. One possible reason is that the advantage of AR-based information detection could decrease the difficulty of reading information in virtual reality. In addition, users might have to spend considerable physical resources (head movement, hand movement and general movement) to interact with augmented information as well as content separate in the physical environment, which is not as natural and easy as in the condition combining AR and VR. A similar pattern of interaction effect between AR and VR can be observed on effort and overall workload, and when using AR and VR together, users would not experience higher degrees of effort and overall workload compared with using either VR or AR alone.

⁴ In this study, *information detection* refers to such features of AR that information available in physical reality such as 2D images (LP record covers) can be detected and tracked, triggering digital content to appear (digital product information page); *information reading* in VR refers to the function of VR that allow users to be able to browse and read digital information.

6 Conclusion

With the development of information technologies, human-computer interfaces and computing power, our daily lives are increasingly lead in technology-mediated environments. Therefore, one of the key research questions is how the workload for carrying out activities as mediated by extended reality technologies may differ both quantitatively and qualitatively from carrying them out without them, as well as how the usability can be improved in the future. Thus, the present study investigated how XR (AR and VR) influence the six dimensions of workload (NASA Task Load Index: mental demand, physical demand, temporal demand, performance, effort, and frustration) and overall workload based on a 2 by 2 between-subject experiment in the retail context. The results from granular analysis indicate that AR was significantly associated with overall workload, especially mental demand and effort, while VR had no significant effect on any of the dimensions of workload. Additionally, the results of the interaction effects show that a combination of AR and VR compared to a single technology would not increase the difficulty of completing tasks (e.g. overall workload and effort), and may even decrease the difficulty (e.g. physical demand). By filling the current gap in research on XR of information systems, human computer interaction and management science, the current study made a considerable research contribution as well as practical guidance for XR designers, developers and practitioners.

6.1 Research implications

The analysis of workload in the present study deepens the understanding of what factors affect usability and efficiency, and what factors are important to users in extended realities, contributing to the corpus of information system and technology-oriented information

science. There is a lack of comprehensive understanding and accurate measurement relating to system efficiency and usage barriers in the extant XR-related literature. For a long time, usability has been considered as one of the determinant factors of efficiency in a system (Bangor et al., 2008; Bevan & Macleod, 1994). This construct has mostly been investigated as one single variable in previous studies (system usability score (SUS): see e.g. Sauro, 2011), or measured by perceived ease of use and usefulness under other research frameworks (e.g. the technology acceptance model (TAM): see Davis, 1989). Instead of directly measuring usability as a whole, our study adopted the NASA Task Load Index (NASA-TLX). Even though such a workload-related measure scale is less seen in information systems, it provides strong theoretical support for examining the different dimensions influencing usability when processing information in shopping environments enabled by XR technologies. Beyond offering new insights on how AR and VR independently and interactively influence workload, this study enriches the IS theory by building bridges between fields in the frontier between adoption and ease of use research in relation to new technologies.

In addition, our study provides valuable hands-on research guidance of an experimental design for use in future studies related to the application of XR technologies. We constructed three different XR-mediated shopping environments based on a physical store. Instead of adding interesting interactive interfaces and vivid content in the XR-mediated environment, we aimed at understanding the essence and nature of XR technologies as display methods rather than their information richness (quantity and quality). In order to minimize the potential impact of all possible external factors on the study outcomes, we looked to ensure that all behaviors were to the greatest extent similar in the four shopping conditions by simplifying the shopping programs as close to non-XR shopping conditions as possible, and creating natural interactive ways to shop in each condition. For example, AR was simplified as a technology used to present information (the product information page) in an augmented way. Regarding the devices, given the consideration that consumers needed to interact with products by hand, head-mounted displays were used rather than hand-held devices (e.g., mobile phones or tablets). In addition, we also took the naturalness of interaction into consideration, and instead of using triggers or buttons as controls, two controllers that enabled haptic interfaces for simulating handling products naturally were used in the VR-mediated environments. The rigorous research design aims at ensuring the internal validity of the experiment results.

6.2 Practical implications

The study findings have valuable implications, mainly for XR designers and programmers, and also for practitioners from different areas, especially those in the retail sector. Due to a lack of theoretical guidance and granular analysis of perceived workload, many program developers and designers face the challenge of finding good solutions of how to improve the perceived usability, ease of use, usefulness and efficiency extended towards XR technologies and systems (Gabbard & Swan, 2008; Kaufmann & Dünser, 2007; Shin et al., 2014; Virvou & Katsionis, 2008). The results indicate that the use of AR causes relatively high effort and mental demand compared to conditions without AR. Thus, designers should consider factors such as ease of use, wearability, mobility and interactivity for reducing the required effort in completing tasks with AR, and be more receptive to the degrees of information complexity and richness that might lead to high mental demand, especially when the AR application targets specific groups such as the elderly, children, patients, and the disabled. In terms of VR, similar experiences to those in real life can be emulated without necessarily increasing the difficulty experienced. The results of this study may help reduce the concerns of designers and developers to using VR to create digital experiences that look to replace physical reality. Given that the pure “reconstruction” setting did not increase the difficulty of conducting activities or completing certain tasks, there seems to be an opportunity for adding more vivid content and comprehensive interaction mechanisms into virtual reality, with a relatively low increased challenge of use.

Furthermore, the findings of this study can be applied to retail, and other contexts such as education, training, healthcare, and entertainment. Since the research context is shopping and retail, retail practitioners can directly benefit from the findings of the present study to develop XR retail strategies. One of the main findings is that VR did not increase the challenges for use or offer inferior experiences towards shopping when compared to a situation without VR (e.g., physical reality). Therefore, we encourage retailers to bravely attempt to use VR technology, for example to replace online and offline shopping environments, given that VR can provide a possibility for consumers to shop online at any time and in any place (Bonetti et al., 2018; Rosedale, 2016), and at the same time, immerse consumers in an enhanced artificial reality (such as the 3D environment, a rich interface and natural interactions) that mimics a physical store. Retailers are also encouraged to consider embedding elements of an augmented interface and content into the virtual reality, since AR might bring additional shopping experiences to consumers such as enjoyment,

playfulness, fun, etc. (Huang & Liao, 2015; Rese et al., 2017; Spreer & Kallweit, 2014). In addition, practitioners can refer to the findings of the present study and improve the efficiency of XR applications in fields such as learning, working, rehabilitation, and entertainment.

7 Limitations

The shopping environment was carefully considered and designed for investigating the matter of perceived workload related to XR, and selected as the research context in the current study. Shopping tasks usually require a lot of physical and cognitive resources, and have been used in many XR-related studies (see e.g. studies on user experience, Speicher et al., 2017; the memory of spatial knowledge, Liang et al., 2019, Man et al., 2011; shopper behavior, Meißner et al., 2019). However, there is still a possibility that VR and AR may have different impacts on the six dimensions of workload in different information systems and application environments. For example, in learning and training, students and trainees usually receive information passively and require more cognitive resources. However, for a technology-related large-scale laboratory experiment, it is unrealistic to design multiple research contexts within a short time. Shopping therefore seems to be the best research context, although we have to admit that one of the limitations of the current study is that the generalization and robustness of the results may not be readily or freely transferable to other sectors.

A second limitation of the research design in our study is the lack of multi-sensory experience in virtual reality, which is associated with user experience. Even though the shops we modeled in virtual reality (Groups 3 and 4) were highly realistic compared to a physical shop, some participants reported that their shopping experiences were a bit “odd” due to the lack of sound and touch experience when moving and interacting with the products. Given that visual experience has been considered as the most important dimension of human sensory experience, we expect that the visual experience itself is sufficient to create the perception of being “there”. Nevertheless, future studies should provide a more multi-dimensional sensory experience in their research design (Xi & Hamari, 2021).

Another limitation is related to the XR devices that were used in the experiments. On one hand, head-mounted displays have limitations in use and are unsuited for certain users. In the current study, two participants were omitted due to a difficulty in wearing / using headsets. Even though the omissions did not influence the results of the experiment, we can still say that there are a certain requirements for XR users which slightly

weaken the broad generalization of the findings. On the other hand, two different head-mounted displays were selected to either provide augmenting information in physical reality (Microsoft HoloLens) or to create a virtual reality (Valve Index). Even though participants used these head-based devices and interacted with products in a similar way, there is still a possibility that the statistical differences seen in some areas of the results may be influenced to some degree by the different technical capabilities of the equipment, such as resolution, detection system, field of view, comfort, response time, etc. As pioneering research, the current study made a great effort to investigate comparisons between AR and VR, based on a rigorous experimental method. However, due to relatively irrevocable technical limitations, these potential technical factors might not have been fully controlled in this experiment. Therefore, we encourage future researchers to adopt more improved XR devices to control and reduce any meaningful influence brought by technological difference.

In addition, the underlying mechanism and boundary conditions of how AR and VR influence different dimensions of workload could be investigated by future researchers. For example, our study did not reveal why AR would lead to high effort and mental demand and overall workload (compared to conditions without AR) while VR did not seem to affect any dimensions of workload. Therefore, there might be some important mediating variables (e.g., presence - see Lackey et al., 2016; Lum et al., 2018; flow - see Lackey et al., 2016; emotion - see Cai & Lin, 2011; Lum et al., 2018; Truschzinski et al., 2018) related to the experiential values of AR and VR. More importantly, demographic information was only presented in a descriptive way in the current study, and factors such as gender, age, nationality and income might play moderating roles in differently influencing the difficulty levels of using AR and VR.

As a final consideration, student participants were recruited in the current study for the reason of ready availability, less foreseeable difference within the sample, and an anticipated acceptance towards novel technologies. Therefore, the results might be influenced by the singularity of the sample. Particularly, the younger age groups have been seen to have a higher degree of flexibility, mobility, more tolerance of discomfort, and less difficulty in operating systems and devices. As for most younger users, wearing headsets and using controllers do not seem to restrict their movement and interaction. However, for specific groups such as elder adults, patients, children, or the disabled, there might be more challenges in operating extended realities. Therefore, future studies could investigate the workload of using AR and VR by recruiting participants from other demographic groups.

Appendix A

Examples of advertisements

**Tampereen yliopisto
Tampere University**

Gamification Group
GLIHP1...GG1

CALL FOR PARTICIPANTS

Contact Us
nannan.xi@tuni.fi
shuo.yang@tuni.fi

We are looking for volunteers (STUDENTS ONLY) to take part in a study of [shopping experience]

As a participant in this study, you would be asked to spend [the gift card in the shop and fill out the computer-based survey]

Your participation is **entirely voluntary** and would take up approximately [1 hour] of your time. The experiment will be conducted in Tampere University, Hervanta campus. In appreciation for your time, you will receive [the products selected in the shop as your own they are worth roughly 10 Euro]

OPEN
Yes WE ARE

YOU CAN SCAN THE QR CODE NOW OR GET THE LINK TO FILL OUT YOUR BASIC INFORMATION ONLINE.

WE WILL CONTACT YOU AS SOON AS POSSIBLE.

<https://cutt.ly/gamification>

(Tri-fold poster)

Appendix B

Experiment instruction



Scenario:

While you are passing by a second-hand LP record shop, you suddenly realize that you have a **10 euro** gift card given by your friend last week. You find out that the expiry date of the gift card is today, which means you have to use it as soon as possible. Thus, you decide to use this gift card to get records for yourself before the shop closes. Remember the shop will close in 10 minutes.

Gift card:

This gift card has a 10 euro value. You can use it to buy any records in the shop. Please try to make the best purchasing decision because you can get the records in the end and get them back home. Remember you cannot get any amount of the gift card and you have to use it completely.

LP record:

You can pick up records, turn around and read information. Each record has its own price tag on the back. There are **54 records** in the shop in total. Just remember don't open the cover to avoid scratching the record and each hand can only hold one LP record.

Extra information: We provide extra information for each record which can help you to make a better purchase decision. You can find e.g., the artist's name, album title, released year, style, track-list, company, credit, and social-related information.

Time: You need to spend a **full 10 minutes** in the shop. The researcher will knock the door when the timer ends. You are not allowed to use the phone or watch during shopping.

Purchase decision: During the shopping time, you can put the selected records on the cashier table and change your selection at any time. We will only ask you to pay the LP records

on the cashier table when the time ends. Thus, make sure the total amount of the selected products does not exceed **10 euro**.

Notice: If you need any help or feel uncomfortable during shopping, you will be offered to take a short break during the experiment and are welcome to request additional breaks.

Acknowledgements This work has been supported by KAUTE Foundation under Grant No. 20190003 and No. 20200531; OP Ryhmän Tutkimussäätiö under Grant No. 20200040; Liikesivistysrahasto under Grant No. 14-7798; Business Finland under Grant No. 5654/31/2018 and Grant No. 4708/31/2019 and Academy of Finland under Grant No. 327241 and Grant No. 337653. This work was carried out with the support of Centre for Immersive Visual Technologies (CIVIT) research infrastructure in Tampere University, Finland. Special thanks to Henry Korkeila, Henrietta Jylhä and Shuo Yang's contribution to the project. The authors would also like to thank all student participants from Tampere University.

CrediT author statement Nannan Xi: Conceptualization, Methodology, Validation, Investigation, Writing - Original Draft & Review & Editing, Project administration, Funding acquisition

Juan Chen: Investigation, Data Curation, Formal analysis, Writing - Review & Editing

Filipe Gama: Software, Resources, Investigation, Writing - Original Draft, Visualization

Marc Riar: Writing - Original Draft & Review & Editing

Juho Hamari: Conceptualization, Methodology, Resources, Writing - Original Draft & Review & Editing, Supervision, Funding acquisition

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Afanasiev, S., Maykova, T., Kashuba, V., Carp, I., Afanasieva, A., & Futorny, S. (2018). Ways of improving the efficiency of physical rehabilitation means for patients with generative-distrophic diseases of adolescents. *Journal of Physical Education and Sport.*, 18, 1945–1949. <https://doi.org/10.7752/jpes.2018.s4287>
- Ahsen, T., Dogar, F. R., & Gardony, A. L., 2019. Exploring the Impact of Network Impairments on Remote Collaborative Augmented Reality Applications. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, Paper LBW2619, 1–6. <https://doi.org/10.1145/3290607.3312774>.

- Azuma, R.T., 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4), 355–385. <https://doi.org/10.1162/pres.1997.6.4.355>.
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594. <https://doi.org/10.1080/10447310802205776>
- Barré, J., Michelet, D., Truchot, J., Jolivet, E., Recanzone, T., Stiti, S., Tesnière, A., & Pourcher, G. (2019). Virtual reality single-port sleeve gastrectomy training decreases physical and mental workload in novice surgeons: An exploratory study. *Obesity Surgery*, 29(4), 1309–1316. <https://doi.org/10.1007/s11695-018-03680-9>
- Bednar, P. M., & Welch, C. (2020). Socio-technical perspectives on smart working: Creating meaningful and sustainable systems. *Information Systems Frontiers*, 22(2), 281–298. <https://doi.org/10.1007/s10796-019-09921-1>
- Bevan, N., & Macleod, M. (1994). Usability measurement in context. *Behaviour & Information Technology*, 13(1–2), 132–145. <https://doi.org/10.1080/01449299408914592>
- Bogicevic, V., Seo, S., Kandampully, J. A., Liu, S. Q., & Rudd, N. A. (2019). Virtual reality presence as a preamble of tourism experience: The role of mental imagery. *Tourism Management*, 74, 55–64. <https://doi.org/10.1016/j.tourman.2019.02.009>
- Bonetti, F., Warnaby, G., Quinn, L., 2018. Augmented reality and virtual reality in physical and online retailing: A review, synthesis and research agenda, in: Jung T., tom Dieck M. (Eds.), *Augmented Reality and Virtual Reality*. Springer, Cham, pp. 119–132. https://doi.org/10.1007/978-3-319-64027-3_9.
- Boyd, D. E., & Koles, B. (2019). An Introduction to the Special Issue “Virtual Reality in Marketing”: Definition. *Theory and Practice. Journal of Business Research*, 100, 441–444. <https://doi.org/10.1016/j.jbusres.2019.04.023>
- Cai, H., & Lin, Y. (2011). Modeling of operators’ emotion and task performance in a virtual driving environment. *International Journal of Human-Computer Studies*, 69(9), 571–586. <https://doi.org/10.1016/j.ijhcs.2011.05.003>
- Cain, B. (2007). *A review of the mental workload literature*. Defence Research And Development Toronto.
- Caria, M., Todde, G., Sara, G., Piras, M., & Pazzona, A. (2020). Performance and usability of smartglasses for augmented reality in precision livestock farming operations. *Applied Sciences*, 10(7), 2318. <https://doi.org/10.3390/app10072318>
- Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1), 341–377. <https://doi.org/10.1007/s11042-010-0660-6>
- Castronova, E. (2008). *Synthetic Worlds: The Business and Culture of Online Games*. University of Chicago press.
- Chang, S. C., & Hwang, G. J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students’ scientific project performance and perceptions. *Computers & Education*, 125, 226–239. <https://doi.org/10.1016/j.compedu.2018.06.007>
- Chao, C. J., Wu, S. Y., Yau, Y. J., Feng, W. Y., & Tseng, F. Y. (2017). Effects of three-dimensional virtual reality and traditional training methods on mental workload and training performance. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 27(4), 187–196. <https://doi.org/10.1002/hfm.20702>
- Chihara, T., & Seo, A. (2018). Evaluation of physical workload affected by mass and center of mass of head-mounted display. *Applied Ergonomics*, 68, 204–212. <https://doi.org/10.1016/j.apergo.2017.11.016>
- Clifford, R.M., Khan, H., Hoermann, S., Billinghamurst, M., Lindeman, R.W., 2018. Development of a multi-sensory virtual reality training simulator for airborne firefighters supervising aerial wildfire suppression, in: *2018 IEEE Workshop on Augmented and Virtual Realities for Good (VAR4Good)*, IEEE, pp. 1–5. <https://doi.org/10.1109/VAR4GOOD.2018.8576892>.
- Dang, Y. M., Zhang, Y. G., Brown, S. A., & Chen, H. (2020). Examining the impacts of mental workload and task-technology fit on user acceptance of the social media search system. *Information Systems Frontiers*, 22(3), 697–718. <https://doi.org/10.1007/s10796-018-9879-y>
- Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340. <https://doi.org/10.2307/249008>.
- Dickinson, P., Gerling, K., Wilson, L., & Parke, A. (2020). Virtual reality as a platform for research in gambling behaviour. *Computers in Human Behavior*, 107, 106293. <https://doi.org/10.1016/j.chb.2020.106293>
- Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, 18(1), 7–22. <https://doi.org/10.1007/s10956-008-9119-1>
- Fan, X., Chai, Z., Deng, N., & Dong, X. (2020). Adoption of augmented reality in online retailing and consumers’ product attitude: A cognitive perspective. *Journal of Retailing and Consumer Services*, 53, 101986. <https://doi.org/10.1016/j.jretconser.2019.101986>
- Fast-Berglund, Å., Gong, L., & Li, D. (2018). Testing and validating Extended Reality (xR) technologies in manufacturing. *Procedia Manufacturing*, 25, 31–38. <https://doi.org/10.1016/j.promfg.2018.06.054>
- Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, 100, 547–560. <https://doi.org/10.1016/j.jbusres.2018.10.050>
- Fromm, J., Eyilmez, K., Babfeld, M., Majchrzak, T. A., & Stieglitz, S., 2021. Social media data in an augmented reality system for situation awareness support in emergency control rooms. *Information Systems Frontiers*, 1–24. <https://doi.org/10.1007/s10796-020-10101-9>.
- Gabbard, J. L., & Swan II, J. E. (2008). Usability engineering for augmented reality: Employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics*, 14(3), 513–525. <https://doi.org/10.1109/TVCG.2008.24>
- Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 213–236. <https://doi.org/10.2307/249689>
- Grier, R., Wickens, C., Kaber, D., Strayer, D., Boehm-Davis, D., Trafton, J.G., St. John, M., 2008. The red-line of workload: Theory, research, and design, in: *Proceedings of The Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 18), Sage Publications, Los Angeles, pp. 1204–1208. 10.1177/154193120805201811.
- Guo, Y., Lu, Z., Kuang, H., & Wang, C. (2020). Information avoidance behavior on social network sites: Information irrelevance, overload, and the moderating role of time pressure. *International Journal of Information Management*, 52, 102067. <https://doi.org/10.1016/j.ijinfomgt.2020.102067>
- Hahs-Vaughn, D. (2016). *Applied Multivariate Statistical Concepts*. Taylor & Francis.
- Harley, D., Verni, A., Willis, M., Ng, A., Bozzo, L., & Mazalek, A. (2018). *Sensory vr: Smelling, touching, and eating virtual reality*, in: *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 386–397). Association for Computing Machinery. <https://doi.org/10.1145/3173225.3173241>
- Hart, S.G., 2006. NASA-task load index (NASA-TLX); 20 years later, in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, No. 9), Sage Publications,

- Los Angeles, pp. 904–908. <https://doi.org/10.1177/154193120605000909>.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hart, S.G., Childress, M.E., Bortolussi, M., 1981. Defining the subjective experience of workload, in: *Proceedings of the Human Factors Society Annual Meeting* (Vol. 25, No. 1), Sage Publications, Los Angeles, pp. 527–531. <https://doi.org/10.1177/1071181381025001135>.
- Hart, S.G., 1982. Theoretical basis for workload assessment research at NASA-Ames research center. In Frazier, M.L., Crombie, R.B., (Eds.), *Proceedings of the workshop on flight testing to identify pilot workload and pilot dynamics*. Edward Air Force Base, CA, AFFTC-TR-82-5.
- Hart, S. G., & Wickens, C. D. (1990). Workload Assessment and Prediction. In H. R. Booher (Ed.), *Manprint*. Springer. https://doi.org/10.1007/978-94-009-0437-8_9
- Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2), 262–271. <https://doi.org/10.1162/pres.1992.1.2.262>
- Heller, J., Chylinski, M., de Ruyter, K., Mahr, D., & Keeling, D. I. (2019). Let me imagine that for you: Transforming the retail frontline through augmenting customer mental imagery ability. *Journal of Retailing*, 95(2), 94–114. <https://doi.org/10.1016/j.jretai.2019.03.005>
- Hoehle, H., Venkatesh, V., 2015. Mobile application usability: conceptualization and instrument development. *MIS Quarterly*, 39(2), 435–472. <https://doi.org/10.25300/MISQ/2015/39.2.08>.
- Huang, T. L., & Liao, S. (2015). A model of acceptance of augmented-reality interactive technology: The moderating role of cognitive innovativeness. *Electronic Commerce Research*, 15(2), 269–295. <https://doi.org/10.1007/s10660-014-9163-2>
- Huang, Y. C., Backman, K. F., Backman, S. J., & Chang, L. L. (2016). Exploring the implications of virtual reality technology in tourism marketing: An integrated research framework. *International Journal of Tourism Research*, 18(2), 116–128. <https://doi.org/10.1002/jtr.2038>
- IJsselstein, W., Riva, G. 2003. Being there: The experience of presence in mediated environments. In Riva, G., Davide, F., IJsselstein, W.A., (Eds.), *Studies in new technologies and practices in communication. Being there: Concepts, effects and measurements of user presence in synthetic environments*, IOS Press, Amsterdam, pp. 3–16.
- Jost, P., Cobb, S., & Hämmerle, I. (2020). Reality-based interaction affecting mental workload in virtual reality mental arithmetic training. *Behaviour & Information Technology*, 39(10), 1062–1078. <https://doi.org/10.1080/0144929X.2019.1641228>
- Jung, H. S., & Jung, H. S. (2001). Establishment of overall workload assessment technique for various tasks and workplaces. *International Journal of Industrial Ergonomics*, 28(6), 341–353. [https://doi.org/10.1016/S0169-8141\(01\)00040-3](https://doi.org/10.1016/S0169-8141(01)00040-3)
- Kantowitz, B. H. (1987). 3. Mental workload. *Advances in Psychology*, 47, 81–121. [https://doi.org/10.1016/S0166-4115\(08\)62307-9](https://doi.org/10.1016/S0166-4115(08)62307-9)
- Kaufmann, H., Dünser, A., 2007. Summary of usability evaluations of an educational augmented reality application, in: Shumaker, R., (Ed.), *ICVR 2007: Virtual Reality. Lecture Notes in Computer Science (vol 4563)*. Springer, Berlin, Heidelberg, pp. 660–669. https://doi.org/10.1007/978-3-540-73335-5_71.
- Kim, J., & Forsythe, S. (2008). Adoption of Virtual Try-on technology for online apparel shopping. *Journal of Interactive Marketing*, 22(2), 45–59. <https://doi.org/10.1002/dir.20113>
- Kim, K., Billingham, M., Bruder, G., Duh, H. B. L., & Welch, G. F. (2018). Revisiting trends in augmented reality research: A review of the 2nd decade of ISMAR (2008–2017). *IEEE Transactions on Visualization and Computer Graphics*, 24(11), 2947–2962. <https://doi.org/10.1109/TVCG.2018.2868591>
- Kim, M. J., & Hall, C. M. (2019). A hedonic motivation model in virtual reality tourism: Comparing visitors and non-visitors. *International Journal of Information Management*, 46, 236–249. <https://doi.org/10.1016/j.ijinfomgt.2018.11.016>
- Klinker, K., Wiesche, M., & Krcmar, H. (2020). Digital transformation in health care: Augmented reality for hands-free service innovation. *Information Systems Frontiers*, 22(6), 1419–1431. <https://doi.org/10.1007/s10796-019-09937-7>
- Kwok, A. O., & Koh, S. G. (2021). COVID-19 and extended reality (XR). *Current Issues in Tourism*, 24(14), 1935–1940. <https://doi.org/10.1080/13683500.2020.1798896>
- Lackey, S. J., Salcedo, J. N., Szalma, J. L., & Hancock, P. A. (2016). The stress and workload of virtual reality training: The effects of presence, immersion and flow. *Ergonomics*, 59(8), 1060–1072. <https://doi.org/10.1080/00140139.2015.1122234>
- Lal, B., Dwivedi, Y. K., & Haag, M. (2021). Working from home during Covid-19: Doing and managing technology-enabled social interaction with colleagues at a distance. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-021-10182-0>
- Lah, U., Lewis, J. R., & Šumak, B. (2020). Perceived usability and the modified technology acceptance model. *International Journal of Human-Computer Interaction*, 36(13), 1216–1230. <https://doi.org/10.1080/10447318.2020.1727262>
- Lee, K. (2012). Augmented reality in education and training. *TechTrends*, 56(2), 13–21. <https://doi.org/10.1007/s11528-012-0559-3>
- Lee, K. C., & Chung, N. (2008). Empirical analysis of consumer reaction to the virtual reality shopping mall. *Computers in Human Behavior*, 24(1), 88–104. <https://doi.org/10.1016/j.chb.2007.01.018>
- Lewis, J. R. (2014). Usability: Lessons learned... and yet to be learned. *International Journal of Human-Computer Interaction*, 30(9), 663–684. <https://doi.org/10.1080/10447318.2014.930311>
- Li, W. C., Horn, A., Sun, Z., Zhang, J., & Braithwaite, G. (2020). Augmented visualization cues on primary flight display facilitating pilot's monitoring performance. *International Journal of Human-Computer Studies*, 135, 102377. <https://doi.org/10.1016/j.ijhcs.2019.102377>
- Lin, P. H., & Yeh, S. C. (2019). How motion-control influences a VR-supported technology for mental rotation learning: From the perspectives of playfulness, gender difference and technology acceptance model. *International Journal of Human-Computer Interaction*, 35(18), 1736–1746. <https://doi.org/10.1080/10447318.2019.1571784>
- Liang, H. N., Lu, F., Shi, Y., Nanjappan, V., & Papangelis, K. (2019). Evaluating the effects of collaboration and competition in navigation tasks and spatial knowledge acquisition within virtual reality environments. *Future Generation Computer Systems*, 95, 855–866. <https://doi.org/10.1016/j.future.2018.02.029>
- Longo, L. 2018. Experienced mental workload, perception of usability, their interaction and impact on task performance. *PloS One*, 13(8), e0199661. [10.1371/journal.pone.0199661](https://doi.org/10.1371/journal.pone.0199661).
- Lum, H.C., Greatbatch, R., Waldfofle, G., Benedict, J., 2018. How immersion, presence, emotion, & workload differ in virtual reality and traditional game mediums. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1), SAGE Publications, Los Angeles, pp. 1474–1478. <https://doi.org/10.1177/1541931218621334>.
- MacDonald, W. (2003). The impact of job demands and workload on stress and fatigue. *Australian Psychologist*, 38(2), 102–117.
- Madathil, K. C., & Greenstein, J. S. (2017). An investigation of the efficacy of collaborative virtual reality systems for moderated remote usability testing. *Applied Ergonomics*, 65, 501–514. <https://doi.org/10.1016/j.apergo.2017.02.011>

- Man, D. W., Chung, J. C., & Lee, G. Y. (2012). Evaluation of a virtual reality-based memory training programme for Hong Kong Chinese older adults with questionable dementia: A pilot study. *International Journal of Geriatric Psychiatry*, 27(5), 513–520. <https://doi.org/10.1002/gps.2746>
- Manis, K. T., & Choi, D. (2019). The virtual reality hardware acceptance model (VR-HAM): Extending and individuating the technology acceptance model (TAM) for virtual reality hardware. *Journal of Business Research*, 100, 503–513. <https://doi.org/10.1016/j.jbusres.2018.10.021>
- Mathieson, K. (1991). Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research*, 2(3), 173–191. <https://doi.org/10.1287/isre.2.3.173>
- McLean, G., & Wilson, A. (2019). Shopping in the digital world: Examining customer engagement through augmented reality mobile applications. *Computers in Human Behavior*, 101, 210–224. <https://doi.org/10.1016/j.chb.2019.07.002>
- Meißner, M., Pfeiffer, J., Pfeiffer, T., & Oppewal, H. (2019). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research*, 100, 445–458. <https://doi.org/10.1016/j.jbusres.2017.09.028>
- Meshkati, N. (1988). Toward development of a cohesive model of workload. *Advances in Psychology*, 52, 305–314. [https://doi.org/10.1016/S0166-4115\(08\)62394-8](https://doi.org/10.1016/S0166-4115(08)62394-8)
- Millais, P., Jones, S. L., & Kelly, R. (2018). *Exploring data in virtual reality: Comparisons with 2d data visualizations*, in: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–6). ACM. <https://doi.org/10.1145/3170427.3188537>
- Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental Brain Research*, 235(3), 889–901. <https://doi.org/10.1007/s00221-016-4846-7>
- Park, M., & Yoo, J. (2020). Effects of perceived interactivity of augmented reality on consumer responses: A mental imagery perspective. *Journal of Retailing and Consumer Services*, 52, 101912. <https://doi.org/10.1016/j.jretconser.2019.101912>
- Peukert, C., Pfeiffer, J., Meißner, M., Pfeiffer, T., & Weinhardt, C. (2019). Shopping in virtual reality stores: The influence of immersion on system adoption. *Journal of Management Information Systems*, 36(3), 755–788. <https://doi.org/10.1080/07421222.2019.1628889>
- Pfeiffer, J., Pfeiffer, T., Meißner, M., & Weiß, E. (2020). Eye-tracking-based classification of information search behavior using machine learning: Evidence from experiments in physical shops and virtual reality shopping environments. *Information Systems Research*, 31(3), 675–691. <https://doi.org/10.1287/isre.2019.0907>
- Poushneh, A., & Vasquez-Parraga, A. Z. (2017). Discernible impact of augmented reality on retail customer's experience, satisfaction and willingness to buy. *Journal of Retailing and Consumer Services*, 34, 229–234. <https://doi.org/10.1016/j.jretconser.2016.10.005>
- Poushneh, A. (2018). Augmented reality in retail: A trade-off between user's control of access to personal information and augmentation quality. *Journal of Retailing and Consumer Services*, 41, 169–176. <https://doi.org/10.1016/j.jretconser.2017.12.010>
- Raptis, G. E., Fidas, C., & Avouris, N. (2018). Effects of mixed-reality on players' behaviour and immersion in a cultural tourism game: A cognitive processing perspective. *International Journal of Human-Computer Studies*, 114, 69–79. <https://doi.org/10.1016/j.ijhcs.2018.02.003>
- Rauschnabel, P. A. (2021). Augmented reality is eating the real-world! The substitution of physical products by holograms. *International Journal of Information Management*, 57, 102279. <https://doi.org/10.1016/j.ijinfomgt.2020.102279>
- Reid, D. (2004). A model of playfulness and flow in virtual reality interactions. *Presence: Teleoperators & Virtual Environments*, 13(4), 451–462. <https://doi.org/10.1162/1054746041944777>
- Rese, A., Schreiber, S., & Baier, D. (2014). Technology acceptance modeling of augmented reality at the point of sale: Can surveys be replaced by an analysis of online reviews? *Journal of Retailing and Consumer Services*, 21(5), 869–876. <https://doi.org/10.1016/j.jretconser.2014.02.011>
- Rese, A., Baier, D., Geyer-Schulz, A., & Schreiber, S. (2017). How augmented reality apps are accepted by consumers: A comparative analysis using scales and opinions. *Technological Forecasting and Social Change*, 124, 306–319. <https://doi.org/10.1016/j.techfore.2016.10.010>
- Riar, M., Korbel, J., Xi, N., Zarnekow, R., Hamari, J., 2021. The use of augmented reality in retail: A review of literature. In *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*, Hawaii, USA, pp 638-647. <http://hdl.handle.net/10125/70689>
- Rosedale, P. (2016). Virtual reality: The next disruptor: A new kind of worldwide communication. *IEEE Consumer Electronics Magazine*, 6(1), 48–50. <https://doi.org/10.1109/MCE.2016.2614416>
- Sauro, J., 2011. Measuring usability with the System Usability Scale (SUS). Available from the following url: <http://www.measuringusability.com/sus.php>.
- Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, 4(2), 183–201. <https://doi.org/10.1089/109493101300117884>
- Schultheis, M., & Rizzo, S. (2001). The application of virtual reality technology in rehabilitation. *Rehabilitation Psychology*, 46, 296–311. <https://doi.org/10.1037/0090-5550.46.3.296>
- Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays*, 29(2), 58–69. <https://doi.org/10.1016/j.displa.2007.09.005>
- Shin, J. H., Ryu, H., & Jang, S. H. (2014). A task-specific interactive game-based virtual reality rehabilitation system for patients with stroke: A usability test and two clinical experiments. *Journal of Neuroengineering and Rehabilitation*, 11(1), 32. <https://doi.org/10.1186/1743-0003-11-32>
- Shu, Y., Huang, Y. Z., Chang, S. H., & Chen, M. Y. (2019). Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Reality*, 23(4), 437–446. <https://doi.org/10.1007/s10055-018-0376-x>
- Somrak, A., Humar, I., Hossain, M. S., Alhamid, M. F., Hossain, M. A., & Guna, J. (2019). Estimating VR Sickness and user experience using different HMD technologies: An evaluation study. *Future Generation Computer Systems*, 94, 302–316. <https://doi.org/10.1016/j.future.2018.11.041>
- Speicher, M., Cucerca, S. Krüger, A., 2017. VR Shop: A mobile interactive virtual reality shopping environment combining the benefits of on-and offline shopping, in: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (Vol. 13). ACM, New York, pp.1-31.
- Spreer, P., Kallweit, K., 2014. Augmented reality in retail: assessing the acceptance and potential for multimedia product presentation at the PoS. *Transactions on Marketing Research*. 1(1), 20-35. 10.15764/MR.2014.01002.
- Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of affordances for virtual reality and

- augmented reality. *Journal of Management Information Systems*, 36(3), 683–729. <https://doi.org/10.1080/07421222.2019.1628877>
- Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4), 73–93. <https://doi.org/10.1111/j.1460-2466.1992.tb00812.x>
- Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, 86, 77–90. <https://doi.org/10.1016/j.chb.2018.04.019>
- Szollos, A. (2009). Toward a psychology of chronic time pressure: Conceptual and methodological review. *Time & Society*, 18(2-3), 332–350. <https://doi.org/10.1177/0961463X09337847>
- Tarafdar, P., Leung, A., Yue, W.T., Bose, I., 2019. Impact of immersive interface design on consumer perceptions during online product, in: Proceedings of 2019 ICIS (International Conference of Information Systems): *Human Computer/Robot Interactions and Interfaces*. ACM, New York, article no. 4. <https://aisel.aisnet.org/icis2019/>.
- Taylor, S., & Todd, P. (2001). Understanding information technology usage: a test of coming models. *Information Systems Research*, 6, 144–176. <https://doi.org/10.1287/isre.6.2.144>
- Technavio, 2020. COVID-19 Impacts: Augmented Reality (AR) and Virtual Reality (VR) Market Will Accelerate at a CAGR of Over 35% Through 2020-2024 | The Increasing Demand for AR and VR Technology to Boost Growth | Technavio. <https://www.businesswire.com/news/home/20200903005356/en/COVID-19-Impacts-Augmented-Reality-AR-Virtual-Reality> (accessed 13 September 2020).
- Tremmel, C., Herff, C., Sato, T., Rechowicz, K., Yamani, Y., & Krusienski, D. J. (2019). Estimating cognitive workload in an interactive virtual reality environment using EEG. *Frontiers in Human Neuroscience*, 13, 401. <https://doi.org/10.3389/fnhum.2019.00401>
- Truschzinski, M., Betella, A., Brunnett, G., & Verschure, P. F. (2018). Emotional and cognitive influences in air traffic controller tasks: An investigation using a virtual environment? *Applied Ergonomics*, 69, 1–9. <https://doi.org/10.1016/j.apergo.2017.12.019>
- Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 243–268). John Wiley & Sons.
- Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education*, 50(1), 154–178. <https://doi.org/10.1016/j.compedu.2006.04.004>
- Wang, C. H., Tsai, N. H., Lu, J. M., & Wang, M. J. J. (2019). Usability evaluation of an instructional application based on Google Glass for mobile phone disassembly tasks. *Applied Ergonomics*, 77, 58–69. <https://doi.org/10.1016/j.apergo.2019.01.007>
- Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, 10, 158. <https://doi.org/10.3389/fpsyg.2019.00158>
- Weidinger, J., Schlauderer, S., & Overhage, S. (2018). Is the frontier shifting into the right direction?. A qualitative analysis of acceptance factors for novel firefighter information technologies. *Information Systems Frontiers*, 20, 669–692. <https://doi.org/10.1007/s10796-017-9785-8>
- Whitlock, M., Harnner, E., Brubaker, J. R., Kane, S., & Szafir, D. A. (2018). *Interacting with distant objects in augmented reality*, In *Proceedings of the 25th IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 41–48). Reutlingen.
- Wickens, C.D., 1992. Virtual reality and education. In *Proceedings of 1992 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, New York, pp. 842–847. <https://doi.org/10.1109/ICSMC.1992.271688>.
- Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, 50(3), 449–455. <https://doi.org/10.1518/001872008X288394>
- Xi, N., Chen, J., Gama, F., Korkeila, H., Hamari, J., 2021. The effect of operating in many realities on memory: An experiment on memory recognition in extended realities. In *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*, Hawaii, USA, pp 4486–4495.
- Xi, N., & Hamari, J. (2021). Shopping in virtual reality: A literature review and future agenda. *Journal of Business Research*, 134, 37–58. <https://doi.org/10.1016/j.jbusres.2021.04.075>
- Yim, M. Y.-C., Chu, S.-C., & Sauer, P. L. (2017). Is Augmented Reality Technology an Effective Tool for E-commerce?. An Interactivity and Vividness Perspective. *Journal of Interactive Marketing*, 39, 89–103. <https://doi.org/10.1016/j.intmar.2017.04.001>
- Zhang, X., Jiang, S., Ordóñez de Pablos, P., Lytras, M. D., & Sun, Y. (2017). How virtual reality affects perceived learning effectiveness: A task–technology fit perspective. *Behaviour & Information Technology*, 36(5), 548–556. <https://doi.org/10.1080/0144929X.2016.1268647>
- Zhao, X., Shi, C., You, X., & Zong, C. (2017). Analysis of mental workload in online shopping: Are augmented and virtual reality consistent? *Frontiers in Psychology*, 8, 71. <https://doi.org/10.3389/fpsyg.2017.00071>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dr. Nannan Xi is an assistant professor (tenure track) in the Faculty of Management and Business, at Tampere University, Finland as well as the senior research fellow at University of Vaasa, Finland. Her research is mainly focused on game-based approaches (gamification, XR/AR/VR, multimodality, wearable technology) in information systems, management, consumer psychology and organization studies towards responsible consumption, sustainable decision making, harmonious organizations and digital business transformation. Dr. Xi's research has been published in a variety of prestigious journals, such as *Journal of Business Research*, *International Journal of Human-Computer Studies*, *International Journal of Information Management*, and *Industrial Management & Data Systems*.

Dr. Juan Chen is a Ph.D., lecturer, currently working at the School of Business Administration, Anhui University of Finance and Economics, China. She obtained her doctoral degree (majoring in marketing management) from Zhongnan University of Economics and Law. Aiming at exploring the influences of emerging technologies on economy and marketing strategies, she focuses on researching the effects of extended reality technologies on retailing and tourism.

Filipe Gama is currently pursuing a Ph.D. degree in Computing and Electrical Engineering at the Faculty of Information Technology and Communication Sciences, Tampere University, Finland. He was one of the early-stage researchers of the Marie Skłodowska-Curie Innovative Training Network programme on Full-Parallax Imaging. His research interests include 3D imaging and 3D sensing, light-field imaging, XR/AR/VR - computer vision and computer graphics.

Marc Riar holds a M. Sc. in Business Informatics from the University of Mannheim. Currently he is a doctoral candidate at Technical University of Berlin. His research centers around behavioral perspectives in information systems with his latest projects involving the use and adoption of augmented reality and virtual reality, as well as the use of

gamification in cooperative settings. Prior to joining Technical University of Berlin, Mr. Riar worked as a research assistant in Karlsruhe and Munich. In addition, he worked at various companies, including SAP in Germany and Lufthansa German Airlines in New York.

Dr. Juho Hamari is a professor of gamification and the leader of the Gamification Group at the Faculty of Information Technology and Communication Sciences at Tampere University. Professor Hamari

is leading research on gamification under the Academy of Finland Flagship, Profiling, and Center of Excellence programs. His and his group's research covers several forms of information technologies such as games, motivational information systems, extended realities, new media, and peer-to-peer economies. Dr. Hamari has authored several seminal empirical, theoretical and meta-analytical scholarly articles on these topics from the perspective of consumer behavior, human-computer interaction, game studies and information systems science.