AR-TV and AR-Diànshì: Cultural Differences in Users' Preferences for Augmented Reality Television

Irina Popovici
MintViz Lab, MANSiD Research Center
Ştefan cel Mare University of Suceava
Suceava, Romania
i.popovici.irina@gmail.com

Radu-Daniel Vatavu
MintViz Lab, MANSiD Research Center
Ştefan cel Mare University of Suceava
Suceava, Romania
radu.vatavu@usm.ro

Pu Feng
NLSDE Lab, Beihang University
Beijing, China
fengpu@buaa.edu.cn

Wenjun Wu
NLSDE Lab, Beihang University
Beijing, China
wwj@nlsde.buaa.edu.cn

ABSTRACT

As Augmented Reality television gains momentum, it is important to understand whether cultural differences among viewers favor different expectations and preferences for immersion in such new television environments. A previous study documented the preferences of 172 participants from various European countries for twenty application scenarios for ARTV, such as virtual objects coming out of the TV screen into the room. In this work, we conduct an empirical generalization of this previous study to understand potential cultural differences in users' preferences for and expectations of ARTV. To this end, we report insights from data collected from a sample of 147 participants from China, which we compare against the preferences expressed by the participants from Europe from the original study. Our findings reveal similarities, but also differences in terms of expectations of ARTV across the two cultural groups. We draw implications for future research on culturally-aware augmentations of the television watching experience.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Empirical studies in HCI**.

KEYWORDS

Augmented Reality; Mixed Reality; Television; Smart TV; Cultural differences; Study.

ACM Reference Format:
Irina Popovici, Radu-Daniel Vatavu, Pu Feng, and Wenjun Wu. 2021. AR-TV and AR-Diànshì: Cultural Differences in Users' Preferences for Augmented Reality Television. In *ACM International Conference on Interactive Media Experiences (IMX '21), June 21–23, 2021, Virtual Event, NY, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3452918.3458801

IMX '21, June 21–23, 2021, Virtual Event, NY, USA

ACM ISBN 978-1-4503-8389-9/21/06…$15.00
https://doi.org/10.1145/3452918.3458801

1 INTRODUCTION

Augmented Reality (AR) technology provides new ways to enhance the television watching experience with new content, input and output modalities, interaction techniques, and interactive devices toward Augmented Reality TV [28,35,43], or ARTV for short. Examples of such new experiences include extending the field of view beyond the TV screen [14,15,35,40], overlaying content on top of the TV broadcast [41], telepresence for social television watching [25,41], auditory enhancements [21], new ways to interact with the TV [26,29], assistance for sign language interpretation [44], and using new devices to consume TV content [3,4].

As research and development of technology advance in the area of ARTV, it is essential to understand users' expectations and preferences for ARTV experiences in which they would like to immerse. In 2019, Popovici and Vatavu [28] published the results of the first exploratory user study in this direction by surveying a number of 172 participants from several European countries and documenting their preferences for ARTV. Their results highlighted high perceived value of wall-sized and room-sized video projections in the living room, and preferences for watching movie scenes from multiple perspectives as well as for virtual objects coming out of the TV screen into the room; see Figure 1, on the next page, for illustrations of some of these scenarios and their positioning in the ARTV conceptual space [43]. Other scenarios, such as a live video stream of friends watching the same television broadcast or displaying multiple virtual TV channels around the physical TV screen were rated less valuable by the participants from that study.

Although these findings are useful for the design of ARTV informed by users' preferences and expectations of such technology, it is unclear how they generalize across different cultures. It is a known fact that the culture in which one was born and raised determines different beliefs and perceptions regarding life in general as well as preferences about the use of technology and interactions with technology in particular; see Kyriakoullis and Zaphiris [18] for a survey of "Culture and HCI" and their discussion about culture intertwining interaction design. For example, in drone computing, E *et al.*'s [9] study reported cultural differences between users from the USA and China determining different preferences for gestures to interact with personal drones compared to the status-quo [5].

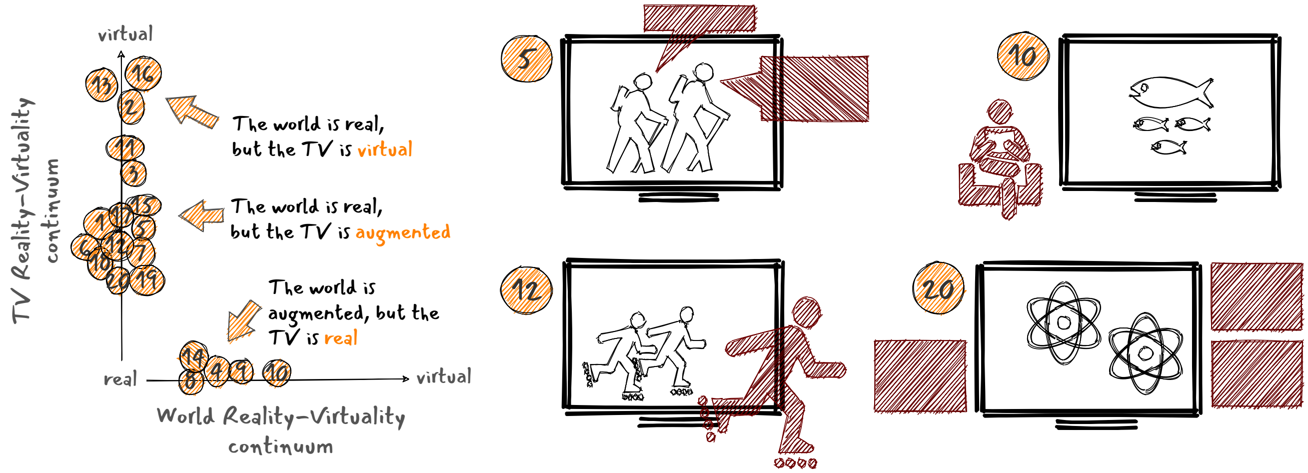

Figure 1: Illustrations of several scenarios for ARTV [28] examined in this work: (5) additional content, such as characters' names, displayed next to the TV set, (10) live 3D representation of remote viewers watching the same TV show, (12) real characters coming out of the TV screen, and (20) virtual TV channels projected next to the TV set; see Table 7 for descriptions of all of the twenty scenarios examined in our work. On the left, these scenarios are positioned in the ARTV space of Vatavu *et al.* [43].

Another example is Wu *et al.* [49] that reported different preferences of users to interact inside the vehicle, in VR, and with the TV set, according to their different cultural backgrounds.

By examining and documenting cultural differences to understand users' preferences regarding interactive systems, we can foster better design that is favorably received by end users, culturally meaningful, and capable of delivering richer and more rewarding user experiences. Unfortunately, the effect of culture on the perceptions of and expectations from AR technology applied to television has not been examined so far, although it may prove key to the success of ARTV adoption in various regions of the world. In this context, several research questions emerge: *Do people with different cultural backgrounds expect the same kind of immersive experiences from AR content and applications for television? Or do their preferences and expectations differ, depending on the characteristics and values of the cultures in which they were born and raised? How are the expectations of enhancing social television watching with immersive technology influenced by being raised in a specific culture? Is AR for television perceived more useful, valuable, informative, or fun by some cultures, but not others?*, and other questions could easily follow. In this paper, we seek to provide preliminary answers in this direction with the following practical contributions:

(1) We conduct an exploratory study with 147 participants from China to elicit their understanding of AR technology in general and preferences for ARTV scenarios in particular, such as application scenarios in which virtual objects or movie characters come out of the TV screen into the room. Our work represents an *empirical generalization*[1] of Popovici and Vatavu's [28] study that documented the preferences for ARTV of 172 participants from Europe.

(2) We compare our results to those reported by Popovici and Vatavu [28], and highlight similarities and differences between the expectations and preferences of participants from China and Europe with respect to applications combining AR and television. Our findings show that ARTV is perceived equally valuable by the participants from both cultural groups, but we also report a significant interaction effect between culture and ARTV scenario, which enables us to discuss different culturally-aware preferences for immersive ARTV experiences. Based on our findings, we advocate for designing ARTV content and applications that are culturally meaningful to the intended audiences, and we discuss opportunities for future research in ARTV.

[1]See Tsang and Kwan [37] for different types of replication studies in the context of research reproducibility. An "empirical generalization" means reconducting a previous study with the same research question and employing the same methods of analysis, but with new participants from a new population.

2 RELATED WORK

We overview prior work on AR systems and applications designed for television as well as work that examined cultural aspects for designing interactive systems. We also provide a brief description of Popovici and Vatavu's [28] study.

2.1 Augmented Reality Television

Vatavu *et al.* [43] proposed the first conceptualization of ARTV by building on Milgram *et al.*'s [22,23] Reality-Virtuality continuum, which they adapted for the two dimensions, *world* and *television*, of their ARTV conceptual space; see Figure 1, left. They also identified a number of possibilities for ARTV systems and applications, such as conventional world/virtual TV or augmented world/conventional TV, and described connections between ARTV, as a distinct area of research, and 3D TV, ambient media, and multimedia alternate realities. For instance, 3D TV technology is relevant to support developments in ARTV since 3D registration is one of the three characteristics of any AR system [2]. One conclusion of Vatavu *et al.* [43] was that *"ARTV reveals itself as a specific type of experience, immersion, media, service, and gateway for televised content"* (p. 9).

AR for interactive television has been explored with a variety of technologies and in a variety of contexts of use. For example, the space around the TV screen has been appropriated by Jones *et al.* [15] with the IllumiRoom system by employing video projections on the wall behind the TV screen. Consequently, IllumiRoom created a more immersive and captivating gaming experience compared to conventional video games displayed on the TV screen alone. The Around-TV [40] prototype utilized the space around the TV screen to display additional TV channels and user interface widgets and controls for the conventional, physical TV set, and explored users' preferences for interacting with such UI elements. The RoomAlive [14] concept and system employed video projections that covered the walls of an entire room to create a physical-virtual environment where users submerge more fully into the video game.

Besides video projections, mobile and wearable devices have been leveraged to augment viewers' experience of watching television. For example, Baillard *et al.* [3] developed a prototype for headsets and smartphones that presented viewers with more content when looking in the direction of the TV set. Vinayagamoorthy *et al.* [44] developed a HoloLens application that displayed a sign language interpreter next to the TV screen. Other systems focused on augmenting content displayed on the TV screen or by using second-screen devices. For example, Vatavu's [41] "audience silhouettes" concept and system displayed the body contours of other viewers at the bottom of the TV screen. Bibiloni *et al.* [4] used second-screen devices to provide more content related to the broadcast. To enable social interaction, Pejsa *et al.* [25] developed Room2Room, a system that displayed the image of a remote interlocutor to recreate the experience of face-to-face conversations.

Interacting with content displayed on or around the TV screen has been addressed with new interaction techniques, input devices, and dedicated user interfaces. For example, in Hover [26], viewers could define mid-air shortcuts to access their preferred TV channels. The PalmRC system of Dezfuli *et al.* [8] enabled viewers to control the TV set by touching specific areas on their palms. Popovici *et al.* [31] integrated "smart pockets" [42] with smart TVs, enabling users to take out digital content from their pockets, e.g., a movie file, and display it on the TV screen.

2.2 Culture and Designing Interactive Systems

Understanding how cultural differences impact the adoption and use of interactive technology can lead to better designs of interactive systems that users find relevant, meaningful, and intuitive according to their cultural backgrounds. For example, the importance of users' expectations of AR technology was revealed by Sarkar *et al.* [36], who documented Indian students' outlook towards technology. Their study comprised of interviews that resulted in a set of twelve expectations about AR technology. However, the authors acknowledged that their results, obtained in one Indian school, might change when extending the study to other schools in India or to other countries with a different culture. Jung *et al.* [16] examined cross-cultural differences in using mobile AR applications by visitors from South Korea and Ireland at two cultural heritage sites. Their findings showed that the participants from South Korea, living in a culture characterized by high-power distance, collectivism, and high uncertainty avoidance (see Hofstede's [11] dimensions of national cultures) perceived stronger dependence on social influence and the hedonic characteristics of AR. The collectivistic nature characteristic of the Chinese culture was observed by Prophet *et al.* [32] for an AR game where users were in charge of nurturing a virtual tree. The participants from that study proposed sharing their virtual trees with their friends through social media applications, such as WeChat,[2] while also expressed the desire to help their friends in attending to their virtual trees.

Cultural studies are thus important in HCI to guide the process of designing meaningful and intuitive interactions between humans and computers [18]. For example, Cauchard *et al.* [5] conducted a gesture elicitation study to understand users' preferences for intuitive gesture commands to interact with personal drones, and identified a set of user-defined gestures with high agreement among participants from the USA. Their work was expanded by E *et al.* [9] that employed the same experimental procedure, but elicited gestures from participants from China. The new results showed that some gestures were proposed by participants from both cultures, but the participants from China used more multimodal interactions to control the drone compared to the participants from the USA. Wu *et al.* [49] examined the impact of different cultural backgrounds, American and Chinese, on user-defined freehand gestures for interacting in VR, inside the vehicle, and with the TV set, and highlighted the impact of culture on users' preferences for gesture commands. In fact, the scientific literature on gesture elicitation has regularly pointed out that empirical findings about users' preferences for gesture-based interactions are likely to not generalize well across cultures. For example, in the first study applying the method to unveil users' preferences for hand gesture input, Wobbrock *et al.* [48] noted: *"Application context could also impact users' choice of gestures, as could the larger contexts of organization and culture. Our participants were all non-technical literate American adults; undoubtedly, children, Eastern, or uneducated participants would behave differently. These issues are worthy of investigation..."* (p. 1091). Other examples that pointed to the need of examining cultural-specific interactions are Ruiz *et al.* [34] regarding input on smartphones, Mauney *et al.* [20] for touchscreen UIs, Kistler and André [17] for interactive storytelling, and Vatavu [39] for free-hand TV control.

2.3 Users' Preferences for ARTV

In this work, we examine the effect of culture, Chinese and European, on users' preferences for potential scenarios of ARTV, such as those illustrated in Figure 1. To this end, we conduct an empirical generalization of Popovici and Vatavu's [28] study by employing participants from China. It is thus useful at this point to briefly overview the original study. Popovici and Vatavu [28] surveyed 172 participants from several countries in Europe with an online questionnaire that, among other aspects, elicited preferences for twenty scenarios for ARTV, such as virtual objects or movie characters breaking out of the TV screen. Results revealed high perceived value for scenarios involving interactive content, wall-sized and room-sized video projections, and presenting multiple perspectives of the same movie scene. Other scenarios, such as live video of remote friends watching the same broadcast, were rated less valuable. These results were revisited in a follow-up work [27]

[2] https://www.wechat.com/en/

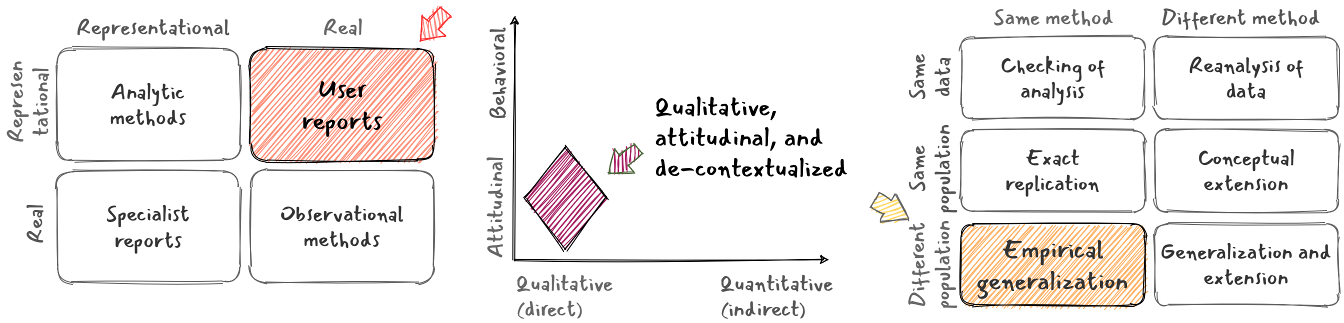

Figure 2: Characteristics of our study expressed using Whitefield *et al.*'s [46] framework for the evaluation of human factors (left), Rohrer's [33] types of UX methods (middle), and Tsang and Kwan's [37] models for research reproducibility (right).

from the perspective of the research agenda [30] proposed for the visual augmentation of the television watching experience. One conclusion [27] (p. 73) regarding the first item of the agenda, i.e., understanding users' preferences, was that cultural effects need to be examined in ARTV research.

3 STUDY

We implemented the procedure from Popovici and Vatavu [28] with a new population, Chinese young adults, to understand whether their expectations of and preferences for ARTV are similar to those of the European participants from the original study [28].

3.1 Study Scope and Characteristics

Our study, following [28], surveys potential end users about their attitudes and preferences for AR technology applied to television in a de-contextualized setting by employing an online questionnaire. The specifics of our study can be precisely characterized in the context of the scientific literature on user research methods as follows. According to Whitefield *et al.*'s [46] framework of evaluation methods of human factors, our study implements a *User Report*, where *real* users are presented with a *representational* interactive system,[3] where the term representational refers to specifications, notational models, or users' mental representations of the system and no physical artifact is available during the study. According to Rohrer's [33] framework of user research methods, our study falls under the *attitudinal*, *qualitative*, and *de-contextualized* categories by documenting "what people say" with direct methods of observation/self-reporting, before actually using the product. Also, our study implements a replication of Popovici and Vatavu [28] using the same research question and methods of analysis, but with a new population. From this perspective, we conduct an *empirical generalization* [37], where we want to understand whether the results of the original study transfer to other populations, i.e., participants from China in our case. Figure 2 illustrates the scope and characteristics of our study according to these frameworks.

3.2 Participants

A number of 206 participants from China volunteered for our study, which was advertised in an academic environment through mailing lists and word of mouth. Participants' educational and professional backgrounds were diverse, including Computer Science, Robotics, Automation, Aeronautics, Chemistry, Biology, Medicine, and Finance. When analyzing the data, we decided to discard 59 of the responses because of uncommitted participation,[4] and validated N=147 respondents. Of these, 128 were male participants with ages between 19 and 25 years old (M=20.6), and 19 were female between 18 and 22 years old (M=19.3). The original study [28] reported results from 109 male (M=25.8 years) and 63 female participants (M=26.8 years). Just like in [28], our sample contains young adults, more likely to adopt new technology.

3.3 Task

Participants were asked to fill in a Wjx questionnaire.[5] Questions were presented in English, but participants were allowed to enter their answers in English or Chinese, whichever they felt was more meaningful to them to express their understanding of AR and preferences for ARTV. Just like in Popovici and Vatavu's [28] study, we primed participants with the same YouTube videos[6] depicting a representative perspective of ARTV as well as visions of this application domain. We collected our data at approximately 1-1.5 years after Popovici and Vatavu [28]. Since our task replicates [28], we refer readers to that paper for further details.

3.4 Measures

We kept the same structure of the questionnaire from [28], which we summarize below for convenience:

(1) **Demographical information:** participants entered details about their age, gender, occupation, and smart devices used on a regular basis (by choosing device types from a list).
(2) **Television watching behavior** with with the following measures: an estimation of the number of hours spent daily

[3]Computer in Whitefield *et al.*'s [46] original formulation.

[4]There are inherent challenges in conducting research with data collected via online questionnaires; see, for example Lefever *et al.* [19] that reported how respondents lose their concentration or interest before finishing the survey if they feel that the survey is too long. Participants were excluded from our study if they entered the same rating for half of the scenarios. (Popovici and Vatavu [28] also discarded 26 of their participants' responses for the same reason of uncommitted participation.)

[5]The original study [28] was implemented with Google Forms, which was difficult to access by the participants from China, so we used https://www.wjx.cn.

[6]https://youtu.be/s8lb3JCNqcc; https://youtu.be/XQ96C2Zolh0; https://youtu.be/re1EatGRV0w; https://youtu.be/lpj90PkfPJ8

on watching TV, most watched genres, and use of second-screen devices while watching television.

(3) **Experience with AR:** a definition of what AR means, a rating of the importance of AR for the respondents on a 5-point Likert scale from 1 (not important) to 5 (very important), and participants' self-reported experience with AR from 1 (not at all experienced) to 5 (experienced to a great extent).
(4) **Perception of ARTV:** participants rated eight characteristics of ARTV using 5-point Likert scales by entering their degree of agreement with the following statements: ARTV is *useful*, *informative*, *desirable*, *fun*, *futuristic*, *discomforting*, demands *cognitive effort*, and enables *social interaction*.
(5) **Preferences for ARTV scenarios:** participants rated the perceived value of twenty ARTV scenarios defined in the original study [28] (see Table 7) using 5-point Likert scales with values from 1 (not valuable) to 5 (very valuable).

3.5 Statistical Analysis

We employ the chi-square test of independence to compare frequency data, e.g., the self-reported use of smart devices, preferences for TV genres, etc. We also use a between-by-within ANOVA test based on ranks—the Brunner, Domhof, and Langer rank-based method described in Wilcox [47] (p. 554)—to analyze participants' ratings of ARTV scenarios, a test that we prefer to Friedman's and Kruskal-Wallis's since it enables us to consider the effect of two variables, including their interaction, in our analysis. (Moreover, this method was designed to perform well, in terms of controlling the probability of Type I errors, under tied values [47].) For multiple comparisons, we use Bonferroni corrections to interpret the results. However, when the number of post-hoc comparisons becomes too large (there are twenty conditions for the ARTV scenarios), we use the Brunner-Munzel test [47] to report a probabilistic characterization of how the two cultural groups of participants differ, and we complement these results with Spearman's correlation coefficient computed on the preference ranks regarding ARTV scenarios.

4 RESULTS

We report in this section preferences for ARTV expressed by our participants from China, which we compare to those of the participants from Europe from the original study [28]. We also compare the use of smart devices and the experience with and perceptions of AR technology of the two cultural groups.

4.1 Use of Smart Devices

Table 1 presents the frequency of use of several categories of smart devices, such as smartphones, smartwatches, smartglasses, smart TVs, etc., reported by the participants from China, next to which we pasted the data from [28] to facilitate a direct comparison with the participants from Europe. As argued by Popovici and Vatavu [28], these devices are relevant for ARTV since they can render AR content, enable the possibility to interact with AR, or both. The smartphone was the most frequently used device reported by both cultural groups (85.7% and 93.6%, respectively), which was an expected result due to the prevalence of such devices worldwide. Less expected was the fact that smartphone use reported by the participants from China was lower compared to the European participants, but the difference was not statistically significant under multiple comparisons (p=.019 with the Bonferroni correction of α=.05/9=.0056). We found significant differences for *wireless earbuds* (42.1% vs. 11.0%, p<.001, V=.357) and *smart rings* (6.8% vs. 0%, p<.001, V=.195) with participants from China reporting higher usage of these devices. Also, while just one in six participants from China reported using *smart TVs* on a regular basis, the responses of the participants from Europe indicated that one in three were regularly using smart TVs (15.6% vs. 33.1%, p<.001, V=.201).

Table 1: Smart devices reported by the participants from China and Europe as being used on a regular basis.

Device category	China		Europe		Chi-square test†			
	N	%	N	%	$\chi^2_{(1,N=319)}$	p	-	V
Smartphone	126	85.7% ▼	161	93.6% ▲	5.468	.019	*n.s.*	.131
Wireless earbuds	62	42.1% ▲	19	11.0% ▼	40.546	**<.001**	***	**.357**
Smart fitness band	24	16.3% ▲	18	10.4% ▼	2.382	.123	*n.s.*	.086
Smart TV	23	15.6% ▼	57	33.1% ▲	12.909	**.001**	***	**.201**
Tablet	22	14.9% ▼	33	19.1% ▲	0.989	. 320	*n.s.*	.056
Video game console	16	10.8% ▼	21	12.2% ▲	0.136	.713	*n.s.*	.021
Smartwatch	13	8.8% ▼	26	15.1% ▲	2.906	.088	*n.s.*	.095
Smart ring	10	6.8% ▲	0	0% ▼	12.079	**.001**	***	**.195**
AR glasses	3	2.0% ▲	3	1.7% ▼	0.038	.846	*n.s.*	.011

†Significant differences (Bonferroni corrections applied at α=.05/9=.0056) are highlighted in light orange; V is Cramer's measure of association.

Table 2: Self-reported experience with AR technology by the participants from China and Europe.

Level of experience with AR	China		Europe		Chi-square test
	N	%	N	%	
1. *Not experienced at all*	44	29.9% ▼	58	33.7% ▲	$\chi^2_{(4,N=319)}$ = 9.433, p=.051, V=.172
2. *Very little experienced*	59	40.1% ▲	67	39.0% ▼	
3. *Somewhat experienced*	36	24.5% ▲	36	20.9% ▼	
4. *Good experience*	2	1.4% ▼	10	5.8% ▲	
5. *Experienced to a great extent*	6	4.1% ▲	1	0.6% ▼	
Total	**147**	**100%**	**172**	**100%**	

4.2 Experience with AR and Perceptions of AR Technology in General

Table 2 shows the levels of experience with AR technology reported by the participants from China, next to which we pasted the data from [28]. Overall, most participants from both cultural groups were either *not experienced* or *very little experienced* with AR (29.9% and 40.1% for China and 33.7% and 39.0% for Europe, respectively), and approximately one in three participants (30% and 27.3%) was at least *somewhat experienced* with AR technology. A chi-square test showed a marginally significant difference (p=.051) with a weak effect size (V=.172) regarding the experience with AR of the two groups. The slight larger level of experience self-reported by the participants from China can be explained by the fact that popular smartphone apps in China, such as Alipay,[7] QQ,[8] and WeChat,[9]

[7] https://intl.alipay.com/
[8] https://en.mail.qq.com/
[9] https://qpsoftware.net/blog/wechat-vr-ar-pokemon-go

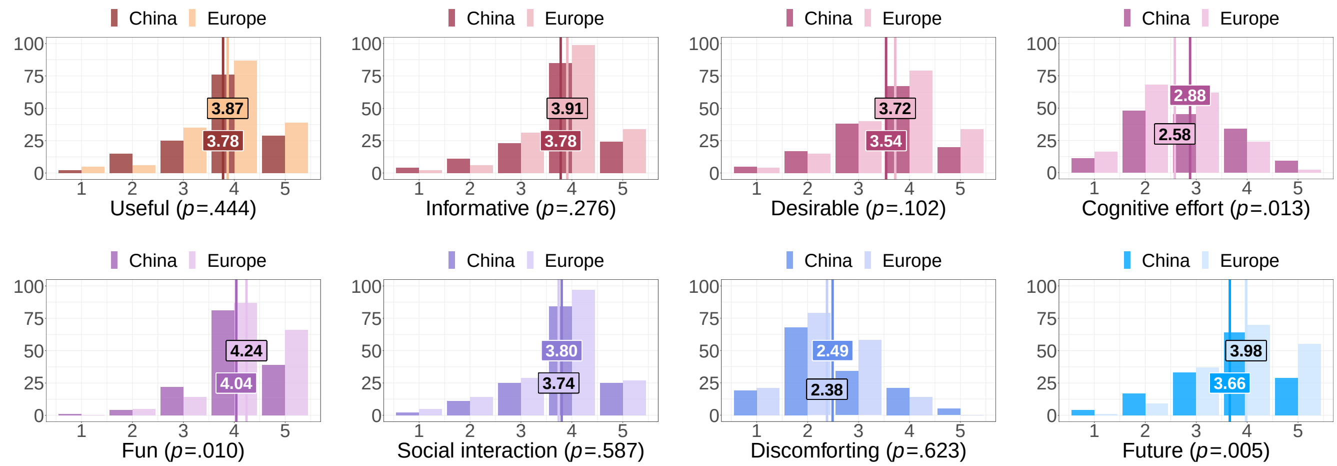

Figure 3: Perception of ARTV. ***Notes:*** **a higher rating means that participants agreed more with the statements presented to them, e.g., ARTV enables social interaction; the absolute number of participants is shown on the vertical axes.**

Table 3: Perceived importance of AR technology reported by the participants from China and Europe.

Level of importance	China N	China %	Europe N	Europe %	Chi-square test
1. *Not important*	19	12.9% ▲	4	2.3% ▼	$\chi^2_{(4,N=319)}$ = 16.953, p=.002, V=.231
2. *Slightly important*	30	20.4% ▲	26	15.1% ▼	
3. *Moderately important*	41	27.9% ▼	53	30.8% ▲	
4. *Important*	39	26.5% ▼	64	37.2% ▲	
5. *Very important*	18	12.3% ▼	25	14.5% ▲	
Total	**147**	**100%**	**172**	**100%**	

incorporate AR features, some of which were mentioned by our participants when asked if they had used AR before. However, the participants from China considered AR technology less important to them (M=3.05 on a scale from 1 to 5) than did the participants from Europe (M=3.47); see Table 3. We connect this result to an observation from Popovici and Vatavu [28], according to which AR products have not yet achieved mainstream [1] and, thus, their potential benefits to end users may be perceived at a low level.

We analyzed participants' responses about their understanding of AR technology and extracted the most frequent words that appeared in their definitions explaining what AR is; see Table 4 for the top 10 words used by the participants from China according to their self-declared level of experience with AR. Overall, we identified the same vocabulary patterns reported by Popovici and Vatavu [28] at the participants from Europe. For example, our participants who were *not experienced at all* provided a variety of definitions for AR, ranging from a good understanding of the concept (e.g., *"Augmented reality (AR) is a technology that integrates virtual information and the real world"* and *"Virtual 3D moving objects that can interact with reality"*) to descriptions that were rather vague and very general (e.g., *"next generation UI"* and *"magic"*). In the second group (participants who self-reported as *very little experienced* with AR), we found characterizations of AR as overlaying of virtual objects onto the physical reality (e.g., *"It's a kind of technology combining virtual information and reality"*). For participants that were *somewhat experienced*, words such as "combine" and "enhance" were more frequently used (e.g., *"[AR] combines reality with the virtual world"* and *"[AR can] enhance my experience while watching or playing games"*), whereas the last two groups (*good experience* and *experienced to a great extent*) provided definitions that confirmed a sound understanding of AR technology (e.g., *"3D pictures with interaction"* and *"AR is a technology that intelligently fuses virtual information with the real world"*)—two definitions that identify all the three characteristics listed by Azuma [2] for AR systems.

4.3 Television Watching Behavior

The participants from China reported spending on average 2.5 hours (SD=1.9) per day watching television and Internet videos, a result similar to that from the original study [28] with European participants (M=2.7, SD=1.7 hours). The participants from China also reported using second-screen devices more often than participants from Europe (p=.023); see Table 6. Regarding TV genres, we noticed several differences as well. The most favorite genres reported by the participants from China were *cartoons*, *movies*, and *news*, as opposed to *movies*, *science and technology*, and *documentaries* preferred by the European participants [28]. Regarding the least preferred genres, both groups rated *reality TV* and *religious TV* in the last of their preferences. Table 5 presents the results for seventeen distinct television genres and highlights statistically significant differences between the two cultural groups.

4.4 Perception of the Characteristics of ARTV

We asked participants to enter their perceptions of ARTV with respect to eight characteristics (e.g., *useful*, *desirable*, *fun*, etc.; see Figure 3), and we compared their ratings with those reported by the participants from Europe. Both cultural groups considered that ARTV was *fun* (4.04 and 4.24 out of a maximum of 5), enabling *social interaction* (3.80 and 3.74), *informative* (3.78 and 3.91) and *desirable* (3.54 and 3.72, respectively)—all average ratings above 3.

Table 4: Top 10 words used by the participants from China to describe their understanding of AR technology. *Note:* words that also made the top-10 list for the European participants from [28] are highlighted in light orange.

	1. Not experienced at all			2. Very little experienced			3. Somewhat experienced			4. Good experience			5. To a great extent		
	Word	N	%	Word	N	%	Word	N	%	Word	N	%	Word	N	%
1	**reality**	13	8.7%	**reality**	15	5.7%	**world**	11	8%	**real**	2	14.2%	information	2	12.5%
2	**virtual**	7	4.6%	**real**	14	5.3%	**reality**	9	6.6%	**reality**	2	14.2%	reality	2	12.5%
3	AR	6	4%	**world**	12	4.5%	**real**	8	5.8%	augmented	1	7.1%	augmented	1	6.25%
4	information	6	4%	ar	10	3.8%	combine	6	4.4%	image	1	7.1%	different	1	6.25%
5	**technology**	6	4%	information	10	3.8%	technology	5	3.6%	interact	1	7.1%	display	1	6.25%
6	image	5	3.3%	make	8	3%	**virtual**	5	3.6%	objects	1	7.1%	experience	1	6.25%
7	make	5	3.3%	technology	8	3%	ar	4	2.9%	sight	1	7.1%	means	1	6.25%
8	**life**	4	2.6%	**virtual**	8	3%	experience	4	2.9%	technology	1	7.1%	method	1	6.25%
9	**real**	4	2.6%	way	8	3%	**augmented**	3	2.2%	unreal	1	7.1%	mixture	1	6.25%
10	3D	3	2%	something	7	2.6%	**enhance**	3	2.2%	use	1	7.1%	now	1	6.25%

Table 5: TV genres reported by the participants from China and Europe.

Television genres	China		Europe		Chi-square test†			
	N	%	N	%	$\chi^2_{(1,N=319)}$	p	-	V
Cartoons	79	53.7% ▲	55	31.9% ▼	15.412	**<.001**	***	**.220**
Movies	79	53.7% ▼	124	72.0% ▲	11.535	**.001**	***	**.190**
News	73	49.6% ▲	68	39.5% ▼	3.294	.070	*n.s.*	.102
Comedy / sitcoms	64	43.5% ▼	88	51.1% ▲	1.848	.174	*n.s.*	.076
Game shows	64	43.5% ▲	48	27.9% ▼	8.499	.004	*n.s.*	.163
Educational shows	56	38.0% ▼	97	56.3% ▲	10.635	**.001**	***	**.183**
Science & tech.	54	36.7% ▼	116	67.4% ▲	30.025	**<.001**	***	**.307**
Music television	47	31.9% ▼	77	44.7% ▲	5.461	.019	*n.s.*	.131
Sports	44	29.9% ▲	43	25.0% ▼	0.972	.324	*n.s.*	.055
Food and drink TV	36	24.4% ▲	39	22.6% ▼	0.145	.703	*n.s.*	.021
Variety shows	36	24.4% ▲	25	14.5% ▼	5.079	.024	*n.s.*	.126
Documentary	31	21.0% ▼	109	63.3% ▲	57.544	**<.001**	***	**.425**
Talk shows	21	14.2% ▼	28	16.2% ▲	0.242	.623	*n.s.*	.028
Soap operas	11	7.4% ▲	4	2.3% ▼	4.705	.030	*n.s.*	.121
Travel television	8	5.4% ▼	35	20.3% ▲	15.101	**<.001**	***	**.218**
Reality television	8	5.4% ▼	15	8.7% ▲	1.274	.259	*n.s.*	.063
Religious television	1	0.6% ▼	13	7.5% ▲	8.935	**.003**	**	**.167**

†Significant differences (Bonferroni corrections applied at α=.05/17=.0029) are highlighted in light orange; V is Cramer's measure of association.

Table 6: Second-screen use during television watching reported by the participants from China and Europe.

Use second-screen device	China		Europe		Chi-square test
	N	%	N	%	
1. *Never*	11	7.5% ▼	19	11.0% ▲	$\chi^2_{(4,N=319)}$
2. *Rarely*	23	15.6% ▼	49	28.5% ▲	= 11.304,
3. *Sometimes*	58	39.5% ▲	60	34.9% ▼	p=.023,
4. *Very often*	38	25.8% ▲	33	19.2% ▼	V=.188
5. *Always*	17	11.6% ▲	11	6.4% ▼	
Total	**147**	**100%**	**172**	**100%**	

Both the *cognitive effort* and *discomforting* descriptions of ARTV were rated below 3, showing that participants rather disagreed with such characterizations. Moreover, the average ratings computed from both cultural groups were highly positively correlated, as indicated by a Pearson's test ($r_{(N=6)}$=.977, R^2=.955, p<.001).

4.5 Preferences for ARTV Scenarios

The last part of the questionnaire consisted of descriptions of twenty scenarios for ARTV that participants rated from 1 (not valuable) to 5 (very valuable) using 5-point Likert scales. The average rating computed for the participants from China was above the neutral threshold 3 and close to the average rating of the participants from Europe: 3.51 (SD=0.55) vs. 3.49 (SD=0.62), respectively. A between-by-within ANOVA based on ranks (see Section 3.5) did not find a significant effect of the group, China or Europe, on participants' perceived value of ARTV scenarios (F=.243, p=.622, *n.s.*), but revealed a significant effect of the scenario (F=15.177, p<.001) as well as a significant interaction between group and scenario (F=3.840, p<.001). Individual results for each scenario are shown in Table 7, next to which we pasted the data from [28] representing the ratings of the participants from Europe. To provide a complete picture of these data, we report the median rating (*Mdn*), the mean (*M*), and the rank order of each scenario. The column $\hat{p}$ reports the Brunner–Munzel [47] estimation of the probability that the rating of a randomly drawn participant from the China group is less than the rating of a randomly drawn participant from the Europe group for the respective ARTV scenario, e.g., 58.5% for scenario 1 and 41.6% for scenario 8. The results show that, on average, the same perceived value was found across the two groups for the ARTV scenarios examined in our work. However, the statistically significant interaction (p<.001) detected between the group and scenario variables is interesting to examine further to understand differences in the *relative* preferences of ARTV scenarios within each cultural group. In the next section, we discuss these differences in detail.

5 DISCUSSION

We discuss similarities and dissimilarities between the preferences of the participants from China and Europe for ARTV. We point to limitations of our study and present opportunities for future work.

5.1 Cultural Similarities and Dissimilarities in the Expectations and Preferences for ARTV

Our analysis of the ratings entered by the participants from China and Europe regarding their perceived value of ARTV highlights similarities between the two cultural groups with average ratings

Table 7: Preferences for ARTV scenarios elicited from 147 participants from China and 172 from Europe. *Note:* scenarios were rated from 1 (not valuable) to 5 (very valuable); differences in ranks larger than 5 are highlighted with ✻ in the last column.

Rank	ARTV scenario (correspondence with Figures 1 and 4)	China			Europe			$\hat{p}$	Rank in [28]	Rank diff.
		Mdn	M	SD	Mdn	M	SD			
1	I would like to be able to control and interact with AR content displayed around or in front of the TV set	4.00	3.79 ▼	0.86	4.00	4.01 ▲	1.03	.585	1	-
2	TV channels displayed next to physical objects in the room, such as weather channel next to the window, documentary channel next to the bookshelf	4.00	3.66 ▲	1.09	4.00	3.62 ▼	1.12	.488	8	+6 ✻
3	Different perspectives of the TV broadcast, such as a movie or show filmed from different angles, displayed next to the TV set	4.00	3.64 ▼	0.92	4.00	3.66 ▲	1.02	.512	6	+3
4	Links to other content, such as movies or photos, displayed around the TV set	4.00	3.63 ▲	1.00	3.50	3.40 ▼	1.02	.431	13	+9 ✻
5	Additional content, such as character names or details about the transmission, displayed next to the TV set	4.00	3.61 ▼	0.90	4.00	3.72 ▲	1.02	.539	2	-3
6	User interface controls, such as capture snapshot or record video buttons, displayed next to the TV set	4.00	3.56 ▼	1.01	4.00	3.63 ▲	0.94	.513	7	+1
7	Real objects coming out of the TV set	4.00	3.56 ▼	1.09	4.00	3.66 ▲	1.16	.531	5	-2
8	Additional information, such as description of buttons, displayed near the buttons of the TV remote control	4.00	3.56 ▲	0.99	3.00	3.28 ▼	0.98	.416	16	+8 ✻
9	Additional content from my smartphone or tablet, such as contacts, notifications, or photos, shown next to the TV set	4.00	3.55 ▲	0.91	3.00	3.31 ▼	1.07	.436	14	+5 ✻
10	Live video or 3D representations of my friends, who are not in the same room with me, but who are watching the same TV show as I do	4.00	3.54 ▲	1.07	3.00	3.27 ▼	1.31	.446	17	+7 ✻
11	A larger field of view created with video projections on the wall behind the TV	4.00	3.53 ▼	0.90	4.00	3.70 ▲	0.99	.557	4	-7 ✻
12	Real characters coming out of the TV set	4.00	3.53 ▲	0.98	4.00	3.46 ▼	1.16	.488	12	-
13	TV content displayed close to me so that I can touch it	4.00	3.52 ▲	1.06	4.00	3.50 ▼	1.05	.501	11	-2
14	Additional content, such as channel info or channel preview, on top or around the TV remote control	4.00	3.47 ▲	0.91	3.00	3.31 ▼	1.00	.447	15	+1
15	Additional content, such as character names or details about the transmission, displayed in front of the TV set	4.00	3.47 ▼	0.89	4.00	3.58 ▲	1.00	.533	10	-5 ✻
16	A very large field of view using video projections in the entire room	4.00	3.44 ▼	0.94	4.00	3.70 ▲	1.03	.578	3	-13 ✻
17	Movie subtitles shown outside the TV set	4.00	3.36 ▲	0.87	3.00	3.25 ▼	1.20	.477	18	+1
18	Menus displayed next to the TV set	3.00	3.40 ▼	0.91	4.00	3.59 ▲	0.98	.553	9	-9 ✻
19	TV channels shown in front of the TV, between the viewer and the TV set	3.00	3.23 ▲	0.97	3.00	2.94 ▼	1.09	.430	20	-
20	TV channels shown next to the TV set	3.00	3.18 ▲	0.97	3.00	3.15 ▼	1.17	.496	19	-

very close together; see Table 7. However, the significant interaction between the *group* and *scenario* variables suggests that the relative orderings of the perceived value of ARTV scenarios are different within each cultural group. To highlight these aspects, we plotted the twenty ARTV scenarios according to the mean ratings corresponding to the two groups; see Figure 4. We found a statistically significant correlation between these ratings that were positively linearly related (Pearson's $r_{(N=20)}$=.710, R^2=.504, p=.01) as well as between their ranks, showing a positive monotonic relationship (Spearman's $\rho_{(N=20)}$=.543, p=.05).

5.1.1 Cultural Similarities. The highest rated ARTV scenario by both groups, China and Europe, was scenario 1 featuring controlling and interacting with AR content displayed around or in front of the TV screen (M=3.79 and M=4.01, respectively). This result strengthens one conclusion from [28] regarding a strong interest of TV viewers to "lean forward" instead of "leaning back" when interactive content is available on TV. The next highest rated scenario by the participants from China was 2 with TV channels displayed next to physical objects in the room, such as the weather channel next to the window, for which the average rating was very close to that of the participants from Europe (3.66 and 3.62, respectively). Along with the other scenarios that feature displaying content near the TV or next to other objects, this finding connects well with the second-screen usage reported by the participants from China, i.e., 76.7% stated that they were *always*, *very often* or *sometimes* using a smart device, such as the smartphone, when watching TV. Scenario 3, watching different perspectives of the same movie scene filmed from different angles, received equally close ratings (3.64 and 3.66) as well as did scenario 13, depicting the situation where TV content

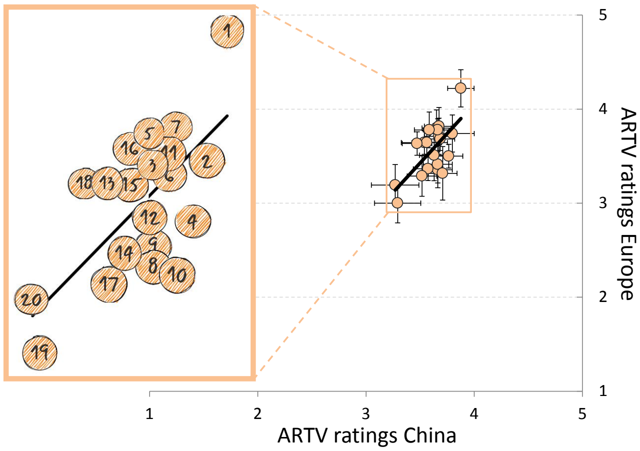

Figure 4: Mean preference ratings and 95% CIs for the twenty ARTV scenarios; correspondences with Figure 1 and Table 7. *Note:* since the rating values were not normally distributed, we computed and plotted 20%-trimmed means and 95% CIs using Wilcox's [47] trimci(...) R function.

is displayed so close to the viewer that they could touch it (mean ratings of 3.52 and 3.50, respectively). However, even though these scenarios elicited the same preference ratings from the two cultural groups of participants, the actual content for which multiple perspectives would be provided 3 or the content that could be touched 13 is likely to be different for each group, given that we found different preferences for TV genres expressed by the two cultural groups (Table 5). At the opposite end of the list, scenarios 19 and 20 describing multiple virtual TV channels shown in front of or next to the conventional TV set were rated the least preferred by both groups; see Table 7.

An interesting result that we replicated with the participants from China regards the scenarios where virtual content, either objects 7 or characters 12, comes out of the TV screen and enters the room. Popovici and Vatavu [28] noted the different nuances of these two scenarios, where virtual characters were less preferred by the participants from Europe, which ranked these scenarios on places 5 and 12, respectively.

5.1.2 Cultural Dissimilarities. We found several ARTV scenarios with mean preference ratings departing from the regression line shown in Figure 4. Overall, 9 of the 20 scenarios (45%) were ranked at a difference of at least five positions compared to the rank ordering from the original study [28]; see the last column from Table 7. For example, the participants from China showed less interest in ARTV scenarios involving larger fields of view created with video projections, either on the wall behind the TV set 11 or in the entire room 16. In contrast, these scenarios occupied positions 3 and 4 in the preferences of the participants from Europe [28]. Another example is scenario 10, depicting the situation where a live video or a 3D representation of one's friends, not in the same room but who are watching the same TV transmission, would be represented in AR. This scenario was ranked 17 in [28] but 10 in our replication study. This result connects to the collectivistic characteristic of the Chinese culture, accentuated by Kyriakoullis and Zaphiris [18] in their overview of culture and HCI, as well as to the low individualism[10] score (20 out of 100) observed for the Chinese culture, according to Hofstede's [11] dimensions. Another example is scenario 4, depicting the situation where links to additional content, such as photos, are available around the TV screen, that was ranked on the 13 position by the participants from Europe [28]. We connect this finding to the significant difference observed in the use of second-screen devices during television watching between the two groups of participants.

5.2 Implications for ARTV

Our findings show that application scenarios that combine AR and television are perceived as valuable across cultures (mean ratings were above 3, see Table 7), but differences are present in how these scenarios are ranked *within* each culture. These results have implications on the future research and practice in the area of ARTV. For instance, practitioners are encouraged to elicit early feedback about their AR application prototypes and AR content from users with cultural backgrounds representative of the target audience of those applications. User-centered [45] and participatory design [24] approaches are recommendable to identify, as early as possible in the design stage, culture-dependent differences in the perception of and preferences for AR content for television in order to make ARTV applications and products culturally meaningful for the intended audiences. For example, user-centered design would favor implementation of scenario 7, where virtual objects break out of the TV screen into the room, but not of scenario 12. Also, culturally-aware design would consider displaying large video projections around the TV set for users from one culture, but not when the product is shipped in a different region of the world, as revealed by the mixed preferences found for scenarios 11 and 16. Moreover, where the audience is known to second-screen during television watching, scenarios in which additional content is displayed both on the second-screen device but also around the TV set, such as 4, might be interesting to investigate further.

Regarding future research opportunities in ARTV, our findings suggest that the *audience* is an important dimension to consider for the design of content, devices, and applications for AR and television. The recent formalization of ARTV [43] could use the incorporation of the audience dimension to highlight different nuances in the perception of possible combinations between physical and virtual *world* and *TV*. The audience dimension should reflect not only cultural aspects, as addressed in this work, but also various abilities to interact [6,38] with television and AR content.[11] Also, users' preferences for interaction techniques and devices to support visual augmentation for television and consumption of ARTV content as well as for elaborating designs where smart TVs integrate smart environments represent interesting avenues for future research. We refer interested readers to the research agenda [30] for the visual augmentation of the television experience for more directions of work for which the culture dimension is relevant. Next,

[10] Hofstede Insights, https://www.hofstede-insights.com/country-comparison/china/
[11] See https://xraccess.org/, a community committed to making virtual, augmented, and mixed reality (XR) accessible to people with disabilities.

we present two more opportunities for future work in ARTV by capitalizing on replications and research reproducibility [10].

5.3 Study Limitations

Our study has a few limitations. For instance, to foster a large number of responses, we conducted our study online but online studies, despite their advantages, are affected by specific problems, such as uncommitted participation; see Lefever *et al.* [19] for details. Also, our sample of participants included young adults only (young adults are more likely to adopt new technology compared to elderly people), but future work should explore the preferences of users of other age categories as well. Moreover, our cultural study has examined Chinese and European[12] cultures only, but whether and how these findings generalize to other cultures needs further investigation as well. Also, media exposure [7] in different cultures may represent another important factor to understand expectations from ARTV. We expect to see more such studies in the measure that ARTV technology, applications, and services become more available for smart TVs and home environments. For now, our results show that, on average, the perceived value of ARTV seems to be similar across cultures, but culture-specific differences manifest in the *relative* value of ARTV scenarios. In the support of future studies that may consider useful comparing their findings to our own, we release our dataset at the web address http://www.eed.usv.ro/~vatavu.

5.4 Replication Studies in the IMX Community

Our study represents a specific instance of a replication, namely an empirical generalization [37], where previous findings are verified with new data collected from a new population (see Figure 2 for our scope and methods). We take this opportunity to highlight the importance of conducting replication studies in HCI, for which the amount of replications has been very small—about 3%, according to an estimate from Hornbæk *et al.* [13]. In this context, more replications are needed to confirm (or contrast) previous results in HCI research [12] in order to consolidate knowledge, and the same is true for IMX. At the intersection of culture, HCI, and AR applied to television, this means conducting more replication studies in the future with new populations of different cultural backgrounds. A caveat, however, is represented by the fact that users' preferences may change quickly in a fast moving technology landscape, such as AR, MR, and VR technology and, thus, findings from replication studies need to be interpreted in this light.

6 CONCLUSION

We examined in this paper the impact of culture on users' expectations and perceptions of ARTV by examining the responses of 147 participants from China and comparing them to the preferences expressed by 172 participants from Europe. We found that, although ARTV was perceived equally valuable on average by the participants from both cultural groups, a significant interaction between culture and ARTV scenario unveiled specific preferences within each culture for immersive ARTV experiences. We hope that our investigation on culture and ARTV will be the starting point for more studies towards culturally-aware design of AR experiences for television. By systematically documenting cultural differences, expectations, and preferences for AR applied to television, we can foster better designs of ARTV systems and applications that are favorably received by end users, culturally meaningful, and that deliver rich and rewarding user experiences.

[12] We treated European culture as unitary but Europe's cultural heritage is rich and diverse; see https://ec.europa.eu/culture/policies/selected-themes/cultural-heritage. And again, so is China's, http://en.chinaculture.org/culturea-z.

ACKNOWLEDGMENTS

This work was supported by a grant of the Romanian Ministry of Education and Research, CNCS-UEFISCDI, project number PN-III-P4-ID-PCE-2020-0434 (PCE29/2021), within PNCDI III.

REFERENCES

[1] A.J. Agrawal. 2018. Three reasons augmented reality hasn't achieved widespread adoption. https://thenextweb.com/contributors/2018/02/16/3-reasons-augmented-reality-hasnt-achieved-widespread-adoption/

[2] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoper. Virtual Environ.* 6, 4 (Aug. 1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

[3] Caroline Baillard, Matthieu Fradet, Vincent Alleaume, Pierrick Jouet, and Anthony Laurent. 2017. Multi-Device Mixed Reality TV: A Collaborative Experience with Joint Use of a Tablet and a Headset. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. ACM, New York, NY, USA, Article 67, 2 pages. https://doi.org/10.1145/3139131.3141196

[4] Toni Bibiloni, Miquel Mascaro, Pere Palmer, and Antoni Oliver. 2015. A Second-Screen Meets Hypervideo, Delivering Content Through HbbTV. In *Proc. of the ACM Int. Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 131–136. https://doi.org/10.1145/2745197.2755513

[5] Jessica R. Cauchard, Jane L. E, Kevin Y. Zhai, and James A. Landay. 2015. Drone & Me: An Exploration into Natural Human-Drone Interaction. In *Proc. of the 2015 ACM Int. Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 361–365. https://doi.org/10.1145/2750858.2805823

[6] Daniel Costa and Carlos Duarte. 2019. Personalized and Accessible TV Interaction for People with Visual Impairments. In *Proc. of the 16th International Web for All Conference (W4A '19)*. ACM, New York, NY, USA, Article 24, 4 pages. https://doi.org/10.1145/3315002.3317566

[7] Claes H. de Vreese and Peter Neijens. 2016. Measuring Media Exposure in a Changing Communications Environment. *Communication Methods and Measures* 10, 2-3 (2016), 69–80. https://doi.org/10.1080/19312458.2016.1150441

[8] Niloofar Dezfuli, Mohammadreza Khalilbeigi, Jochen Huber, Florian Müller, and Max Mühlhäuser. 2012. PalmRC: Imaginary Palm-Based Remote Control for Eyes-Free Television Interaction. In *Proc. of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. ACM, New York, NY, USA, 27–34. https://doi.org/10.1145/2325616.2325623

[9] Jane L. E, Ilene L. E, James A. Landay, and Jessica R. Cauchard. 2017. Drone & Wo: Cultural Influences on Human-Drone Interaction Techniques. In *Proc. of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6794–6799. https://doi.org/10.1145/3025453.3025755

[10] Steven N. Goodman, Daniele Fanelli, and John P.A. Ioannidis. 2016. What Does Research Reproducibility Mean? *Science Translational Medicine* 8, 341 (June 2016), 341ps12. http://dx.doi.org/10.1126/scitranslmed.aaf5027

[11] G. Hofstede. 2011. Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture* 2, 1 (2011), 23 pages. https://doi.org/10.9707/2307-0919.1014

[12] Kasper Hornbæk. 2015. We Must Be More Wrong in HCI Research. *Interactions* 22, 6 (Oct. 2015), 20–21. https://doi.org/10.1145/2833093

[13] Kasper Hornbæk, Søren S. Sander, Javier Andrés Bargas-Avila, and Jakob Grue Simonsen. 2014. Is Once Enough? On the Extent and Content of Replications in Human-Computer Interaction. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3523–3532. https://doi.org/10.1145/2556288.2557004

[14] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proc. of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383

[15] Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 869–878. https://doi.org/10.1145/2470654.2466112

[16] T.H. Jung, H. Lee, N. Chung, and M.C. tom Dieck. 2018. Cross-cultural differences in adopting mobile augmented reality at cultural heritage tourism sites. *International Journal of Contemporary Hospitality Management* 30, 3 (2018), 1621–1645.

https://doi.org/10.1108/IJCHM-02-2017-0084

[17] Felix Kistler and Elisabeth André. 2013. User-Defined Body Gestures for an Interactive Storytelling Scenario. In *Human-Computer Interaction (INTERACT 2013)*, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 264–281. https://doi.org/10.1007/978-3-642-40480-1_17

[18] Leantros Kyriakoullis and Panayiotis Zaphiris. 2016. Culture and HCI: A Review of Recent Cultural Studies in HCI and Social Networks. *Univers. Access Inf. Soc.* 15, 4 (Nov. 2016), 629–642. https://doi.org/10.1007/s10209-015-0445-9

[19] Samuel Lefever, Michael Dal, and Ásrún Matthíasdóttir. 2006. Online data collection in academic research: Advantages and limitations. *British Journal of Educational Technology* 38 (06 2006), 574 – 582. https://doi.org/10.1111/j.1467-8535.2006.00638.x

[20] Dan Mauney, Jonathan Howarth, Andrew Wirtanen, and Miranda Capra. 2010. Cultural Similarities and Differences in User-Defined Gestures for Touchscreen User Interfaces. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 4015–4020. https://doi.org/10.1145/1753846.1754095

[21] Mark McGill, Florian Mathis, Mohamed Khamis, and Julie Williamson. 2020. Augmenting TV Viewing Using Acoustically Transparent Auditory Headsets. In *ACM International Conference on Interactive Media Experiences (IMX '20)*. ACM, New York, NY, USA, 34–44. https://doi.org/10.1145/3391614.3393650

[22] Paul Milgram and Herman Colquhoun Jr. 1999. A Taxonomy of Real and Virtual World Display Integration. In *Mixed Reality: Merging Real and Virtual Worlds*, Yuichi Ohta and Hideyuki Tamura (Eds.). Springer-Verlag, Berlin, Heidelberg. https://www.researchgate.net/publication/2440732_A_Taxonomy_of_Real_and_Virtual_World_Display_Integration

[23] Paul Milgram and Fumio Kishino. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems* E77-D, 12 (December 1994), 1321–1329. https://search.ieice.org/bin/summary.php?id=e77-d_12_1321

[24] Michael J. Muller and Sarah Kuhn. 1993. Participatory Design. *Commun. ACM* 36, 6 (June 1993), 24–28. https://doi.org/10.1145/153571.255960

[25] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. 2016. Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1716–1725. https://doi.org/10.1145/2818048.2819965

[26] Irina Popovici, Ovidiu-Andrei Schipor, and Radu-Daniel Vatavu. 2019. Hover: Exploring Cognitive Maps and Mid-Air Pointing for Television Control. *International Journal of Human-Computer Studies* 129 (2019), 95–107. https://doi.org/10.1016/j.ijhcs.2019.03.012

[27] I. Popovici and R. Vatavu. 2019. Consolidating the Research Agenda of Augmented Reality Television with Insights from Potential End-Users. In *Proc. of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, New York, NY, USA, 73–74. https://doi.org/10.1109/ISMAR-Adjunct.2019.00033

[28] I. Popovici and R. Vatavu. 2019. Understanding Users' Preferences for Augmented Reality Television. In *Proc. of the 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR '19)*. IEEE, New York, NY, USA, 269–278. https://doi.org/10.1109/ISMAR.2019.00024

[29] I. Popovici and Radu-Daniel Vatavu. 2018. Perceived Usability, Desirability, and Workload of Mid-Air Gesture Control for Smart TVs. In *Proc. of the 15th International Conference on Human Computer Interaction (RoCHI '18)*. MatrixRom, Bucharest, Romania, 91–98. https://dblp.org/rec/conf/rochi/PopoviciV18

[30] Irina Popovici and Radu-Daniel Vatavu. 2019. Towards Visual Augmentation of the Television Watching Experience: Manifesto and Agenda. In *Proc. of the 2019 ACM Int. Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 199–204. https://doi.org/10.1145/3317697.3325121

[31] Irina Popovici, Radu-Daniel Vatavu, and Wenjun Wu. 2019. TV Channels in Your Pocket! Linking Smart Pockets to Smart TVs. In *Proc. of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 193–198. https://doi.org/10.1145/3317697.3325119

[32] Jane Prophet, Yong Ming Kow, and Mark Hurry. 2018. Small Trees, Big Data: Augmented Reality Model of Air Quality Data via the Chinese Art of "Artificial" Tray Planting. In *ACM SIGGRAPH 2018 Posters (SIGGRAPH '18)*. ACM, New York, NY, USA, Article 16, 2 pages. https://doi.org/10.1145/3230744.3230753

[33] Christian Rohrer. 2014. When to Use Which User-Experience Research Methods. https://www.nngroup.com/articles/which-ux-research-methods/

[34] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-Defined Motion Gestures for Mobile Interaction. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[35] Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. 2019. Augmenting Television With Augmented Reality. In *Proc. of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 255–261. https://doi.org/10.1145/3317697.3325129

[36] Pratiti Sarkar and Jayesh S. Pillai. 2019. User Expectations of Augmented Reality Experience in Indian School Education. In *Research into Design for a Connected World*, Amaresh Chakrabarti (Ed.). Springer Singapore, Singapore, 745–755. http://dx.doi.org/10.1007/978-981-13-5977-4_63

[37] Eric W.K. Tsang and Kai-Man Kwan. 1999. Replication and Theory Development in Organizational Science: A Critical Realist Perspective. *The Academy of Management Review* 24, 4 (Oct. 1999), 759–780. http://dx.doi.org/10.2307/259353

[38] Ovidiu-Ciprian Ungurean and Radu-Daniel Vatavu. 2021. Coping, Hacking, and DIY: Reframing the Accessibility of Interactions with Television for People with Motor Impairments. In *Proc. of the ACM International Conference on Interactive Media Experiences (IMX '21)*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3452918.3458802

[39] Radu-Daniel Vatavu. 2012. User-Defined Gestures for Free-Hand TV Control. In *Proc. of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. ACM, New York, NY, USA, 45–48. https://doi.org/10.1145/2325616.2325626

[40] Radu-Daniel Vatavu. 2013. There's a World Outside Your TV: Exploring Interactions beyond the Physical TV Screen. In *Proc. of the 11th European Conference on Interactive TV and Video (EuroITV '13)*. ACM, New York, NY, USA, 143–152. https://doi.org/10.1145/2465958.2465972

[41] Radu-Daniel Vatavu. 2015. Audience Silhouettes: Peripheral Awareness of Synchronous Audience Kinesics for Social Television. In *Proc. of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 13–22. https://doi.org/10.1145/2745197.2745207

[42] Radu-Daniel Vatavu. 2017. Smart-Pockets: Body-deictic gestures for fast access to personal data during ambient interactions. *International Journal of Human-Computer Studies* 103 (2017), 1–21. https://doi.org/10.1016/j.ijhcs.2017.01.005

[43] Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. 2020. Conceptualizing Augmented Reality Television for the Living Room. In *Proc. of the ACM Int. Conf. on Interactive Media Experiences (IMX '20)*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3391614.3393660

[44] Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 2019. Personalising the TV Experience Using Augmented Reality: An Exploratory Study on Delivering Synchronised Sign Language Interpretation. In *Proc. of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300762

[45] Karel Vredenburg, Ji-Ye Mao, Paul W. Smith, and Tom Carey. 2002. A Survey of User-Centered Design Practice. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 471–478. https://doi.org/10.1145/503376.503460

[46] Andy Whitefield, Frank Wilson, and John Dowell. 1991. A Framework for Human Factors Evaluation. *Behavior & Information Technology* 10, 1 (1991), 65–79. https://www.tandfonline.com/doi/abs/10.1080/01449299108924272

[47] Rand Wilcox. 2012. *Modern Statistics for the Social and Behavioral Sciences: A Practical Introduction*. CRC Press, Boca Raton, FL, USA.

[48] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[49] Huiyue Wu, Jinxuan Gai, Yu Wang, Jiayi Liu, Jiali Qiu, Jianmin Wang, and Xiaolong (Luke) Zhang. 2020. Influence of Cultural Factors on Freehand Gesture Design. *International Journal of Human-Computer Studies* 143 (2020), 102502. https://doi.org/10.1016/j.ijhcs.2020.102502