

The User Experience of Journeys in the Realm of Augmented Reality Television

Cristian Pamparău

MintViz Lab, MANSiD Research Center
Ștefan cel Mare University of Suceava
Suceava, Romania
cristian.pamparau@usm.ro

Radu-Daniel Vatavu

MintViz Lab, MANSiD Research Center
Ștefan cel Mare University of Suceava
Suceava, Romania
radu.vatavu@usm.ro

ABSTRACT

Augmented Reality Television (ARTV) can take many forms, from AR content displayed outside the TV frame to video-projected TV screens to social TV watching in VR to immersive holograms in the living room. While the user experience (UX) of individual forms of ARTV has been documented before, “journeys” as transitions between such forms have not. In this work, we examine the UX of watching TV when switching between various levels of augmentation. Our findings from an experiment with fourteen participants reveal an UX characterized by high perceived usability, captivation, and involvement with a low to medium workload and a moderate feeling of dissociation from the physical world. We interpret our results in the context of Garrett’s established five-plane model of UX—strategy, scope, structure, skeleton, and surface—and propose a sixth plane, “switch,” which separates conceptually the design of user journeys in ARTV from the specifics of the other UX planes.

CCS CONCEPTS

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

KEYWORDS

Augmented reality, interactive television, ARTV, user experience, UX, Garrett’s model of UX, experiment, HoloLens, HMDs

ACM Reference Format:

Cristian Pamparău and Radu-Daniel Vatavu. 2022. The User Experience of Journeys in the Realm of Augmented Reality Television. In *ACM International Conference on Interactive Media Experiences (IMX ’22)*, June 22–24, 2022, Aveiro, JB, Portugal. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3505284.3529969>

1 INTRODUCTION

Augmented Reality Television (ARTV) is the realm of computer engineering, interaction design, and media production where television meets Augmented Reality (AR), Virtual Reality (VR), and other types of computer-generated and mediated worlds [67,77]. At this confluence, ARTV delivers more than the sum of its parts by

hybridizing AR/VR and televised content in a process where each intensifies the other towards new entertainment experiences for the living room [23,63] and beyond [57,58]. In this process, ARTV can take many forms supported by a variety of computer technology, including personal devices for mobile AR [4,18,36], video projection systems for spatial AR [35,73], and head-mounted displays (HMDs) for immersive AR and VR [80,81], enabling a diversity of opportunities for augmented movies, sports, news, and other TV genres [63]. In fact, in their comprehensive examination of what is ARTV, Vatavu *et al.* [77] concluded: “We found that ARTV can be many things and that prior work has implemented it in various ways, from on-TV augmentations to off-TV content visualized via AR-enabled hand-held devices, HMDs, wall- and room-sized projections, and holograms. By drawing from the various perspectives and key properties that we examined, we can conclude that ARTV reveals itself as a specific type of experience, immersion, media, service, and gateway for televised content” [77, p. 9] (emphasis ours). The keyword from this quote, which encompasses the multitude of manifestations and nuances of ARTV, is “experience.”

The user experience (UX) of ARTV has been examined from diverse perspectives. Prior work has documented users’ preferences for ARTV scenarios [63], such as supplementary content displayed outside the TV frame, movie characters coming out of the TV into the room, and hologram displays of friends from remote locations watching the same TV broadcast. Other studies have elicited intuitive interactions for video-projected TV screens [73], documented user perception of AR holograms shown next to the TV [81], and reported levels of engagement with ARTV [68], among other UX constructs. In this context, knowledge about how to design the UX of ARTV has been growing steadily, but the prior work has focused on understanding the UX of specific forms of ARTV out of a multitude and, thus, design knowledge exists only with respect to specific UX out of a multitude of possible ARTV experiences.

Beyond the UX of specific forms of ARTV that have been examined in the scientific literature [35,63,64,68,74,81], the conceptual structure and medium of a continuum [77] invites to transitions between its points. By capitalizing on Milgram *et al.*’s [47,48] Reality-Virtuality axis, the ARTV continuum of Vatavu *et al.* [77] can support the design of such transitions as “UX journeys” that take users from one form of ARTV to another and, thus, engage them in new TV watching experiences. For example, consider a user watching a soccer game broadcast on the TV screen from their living room. The user activates the display of additional content in AR, such as information about the two teams or the silhouette of a live sports commentator shown next to the TV screen rendered with a video projection [42,73] or through a HMD [68,80]. Replays

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IMX ’22, June 22–24, 2022, Aveiro, JB, Portugal

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9212-9/22/06...\$15.00

<https://doi.org/10.1145/3505284.3529969>

of key moments from the game can be summoned in the form of immersive holograms in the living room, unfolding the action of the game from different viewpoints [14], e.g., of the striker, the goalkeeper, or the referee. After watching the replay, the user switches back to the conventional TV screen format. Another example is a user watching a mystery TV show, where reminders about the characters' pasts, relationships, or clues about the crime are available for inspection next to the TV screen. At some point, the user may wish to rewatch the context in which a specific clue was discovered, so they point to that clue to summon virtual screens that popup in mid-air showing the timeline of relevant events. The user may also wish to invite a friend, from a remote location, who is watching the same show [74], to discuss the turn of events created by a newly discovered clue. The friend joins through holoportation [55] and appears as if they were present in the same physical space.

Such transitions between different points of the ARTV continuum [77] offer flexible control of user immersion and engagement with content during TV watching with a distinctive UX that has not been examined before. In this work, we focus on understanding the UX of such journeys. Our practical contributions are as follows:

- (1) We report findings from a controlled experiment conducted to examine the UX of ARTV journeys represented by transitions among four forms of ARTV rendered with HoloLens: virtual TV screen, AR content shown outside the TV frame, AR content floating in the room between the TV screen and the user, and an augmented living room with immersive TV. Our results reveal an UX characterized by high perceived usability, low to medium workload, high captivation and comprehension, and high perceived involvement with the content, complemented by a moderate feeling of transportation and dissociation from the physical world.
- (2) We employ Garrett's [21] established five-plane model of UX to discuss our results, and propose a new plane called "switch" that conveniently separates at a conceptual level the UX of transitions in the ARTV continuum from the other planes of UX—strategy, scope, structure, skeleton, and surface.

2 RELATED WORK

We relate to prior work on ARTV, and focus on the UX reported by prior user studies examining specific forms of ARTV. We also present the ARTV continuum [77], on which we build in this work to evaluate the UX of journeys in this continuum.

2.1 Augmented Reality Television

ARTV has been implemented in various forms and with various supporting technology. Applications for mobile devices, such as smartphones and tablets, represent one of the most common implementation of ARTV, where users hold up their mobile device to see additional content superimposed on the video feed of the TV set. Examples include Augmented TV [36], Augmen.tv [18], Mixed Reality TV Mozaik [4,20], and Hypervideo [9]. In these implementations, the mobile device acts as a "window on the world" (WoW) display, in Milgram and Kishino's [48] sense. Other implementations use the TV screen to display additional content, such as InAir TV [32], LinkedCulture [53], and Audience Silhouettes [74]. Other scenarios do away with the conventional TV screen completely, and

the mobile device represents the window to content that appears to be present in the living room [85] or outdoors [57,58]. HMDs have also been employed to augment the TV watching experience. For example, the Mixed Reality TV Mozaik [4] runs on HoloLens as well as on a smartphone, TV+HoloLens [80] displays a sign language interpreter next to the TV, and Saeghe *et al.* [68] used HoloLens to synchronize holograms with televised content. Other implementations of ARTV have employed video projections, either on the wall behind the TV [35,42,72,73] or in the entire room [34]. Examples include Point & Click [72], a spatial AR system enabling users to position virtual TV screens on the wall behind the conventional TV set, IllumiRoom [35] and ExtVision [38] that expand the content shown on the TV with video projections on the wall behind it, Smart Wallpaper [7] for electronic wallpaper that can also show TV content, and the Around TV [73] spatial AR system with a graphical user interface composed of virtual TV screens, menus, and control widgets video projected on the wall behind the TV set. Some systems have combined multiple displays, such as screens, walls, and interactive tables [42,76], and everyday objects as new media [31] to enrich the TV watching experience. For more information, we refer to Saeghe *et al.*'s [67] review of ARTV themes.

2.2 Defining and Formalizing UX

UX has been defined generically as the experience encompassing the many aspects involving a specific product or service. For example, Norman and Nielsen [54] summarize UX as "all aspects of the end-user's interaction with the company, its services, and its products" and a broader concept than usability, and the ISO [33] standards define UX in the context of human-centered design for interactive systems as "user's perceptions and responses that result from the use and/or anticipated use of a system, product or service." UX process methodology has known variations, from requirements-based [21] to agile, outcome-based [24] to design sprints [40]. A classic formalization of UX comes from Garrett [21] that proposed a five-plane model. In this model, the five planes or elements of UX are: *strategy* (user needs and product objectives), *scope* (functional specifications and content requirements), *structure* (interaction design and information architecture), *skeleton* (interface design, navigation design, and information design), and *surface* (sensory design). Although Garrett's model was introduced in the context of web sites and software, it was also generically formulated to apply to products and services of many kinds [21, p. XIII] with the main stated goal to generate new product and value opportunities by capitalizing on understanding human behavior. In a 2021 article addressing the current state of UX practice, Garrett [22] reiterated on this goal: "The implicit promise of UX for many of us was a burgeoning philosophy of management by inquiry and insight, in which new creative explorations would lead to new questions about human behavior, which in turn would drive the definition of new product and value opportunities." To the best of our knowledge, Garrett's [21] model has not been employed to examine the UX of ARTV; see the next subsection.

2.3 The User Experience of ARTV

Prior work has reported the UX of ARTV from a diversity of perspectives and employed a wide range of UX constructs. For instance,

the UX of ARTV involving content displayed outside the TV frame was described as “useful [...], fun [...], desirable [...], informative [...] and with good opportunities to enable social interaction” [63, p. 276] and, in other study, “novel, stimulating, aesthetically appealing and perspicuous” [68, p. 260], while some types of AR content, such as virtual characters, were found “unsettling” [68, p. 259]. The UX of on-TV display of remote viewers in the form of audience silhouettes was perceived “creative [...], fun [...], friendly [...], entertaining [...], connected [...], innovative [...], attractive [...], and collaborative [...]” [74, p. 19], enabling an enjoyable social TV watching experience. The UX of an AR hologram of a sign language interpreter displayed outside the TV frame was “distracting,” but also “natural and afforded a sense of completeness” [81, p. 10]. Also, the UX of watching TV in VR revealed “significant improvements regarding participant’s media immersion, engagement, and enjoyment in a shared experience” [46, p. 33:50] compared to a conventional TV.

These findings show that different UX are delivered directly by different types of AR content designed for television, but also indirectly by the UX of the AR/VR technology to render that content. Regarding the former, Popovici and Vatavu [61,63] reported rankings of twenty ARTV scenarios by analyzing the preferences of 172 respondents. In a follow-up study [64], the effect of culture transpired in a different ranking of those scenarios with respondents from China, indicating the importance of acknowledging the particularities of individual cultures for successful UX of ARTV designs that are favorably received by end users. Also, different technology to render ARTV can lead to different UX. For example, Vinayagamoorthy *et al.* [81] reported potential adoption barriers for HMDs, including physical discomfort of wearing them. Holding a mobile device pointed at the TV [4,9,18] will eventually lead to experiencing arm fatigue, but if the fatigue aspects could be mitigated with appropriate interaction design, mobile ARTV applications are likely to integrate well the second-screen TV watching paradigm [16,43].

Prior work has employed a diversity of UX constructs and corresponding measures to document the UX of ARTV, including measures of attention [68,81], distractedness [74], perceived synchronization between the TV and AR content [81], aesthetic appeal [68], perceived usability [68,74], usefulness [63,64,74], engagement [46,68], immersion and presence [46], desirability [63,64,74], enjoyment [74], social interaction [46,63,64,74], and novelty [63,64,68]. These measures are useful to describe various facets of UX, a complex phenomenon encompassing many aspects [33,54].

2.4 The ARTV Continuum

Vatavu *et al.* [77] introduced the “ARTV continuum,” a conceptual space that describes possible augmentations of the TV (i.e., from a physical TV screen to a virtual one) and of the physical world (i.e., the living room, either physical, virtual, or mixed). The two axes of the ARTV space represent instances of Milgram *et al.*’s [47–49] Reality-Virtuality continuum applied to the TV and world, respectively. The origin of the continuum is the conventional TV set in a physical world with no augmentation, and the top right corner is watching TV in a VR environment. Between these two extremes, various forms of ARTV can be specified, characterized, and compared, including all of the systems discussed in Subsection 2.1; see Vatavu *et al.* [77] for more examples and an in-depth discussion.

By establishing its foundation on Milgram *et al.*’s [47–49] Reality-Virtuality axis, the ARTV continuum captures all possible ways to augment the TV watching experience, where the TV and world can be augmented separately, but also conjointly. This quality property enables a wide range of possibilities for practitioners to create diverse UX. The relation between the Reality-Virtuality axis and the experience of observing augmented content at various sensory levels was made explicit by Skarbez *et al.* [70] in a revised interpretation of Milgram *et al.*’s concept of mixed reality, from an environment in which “real world and virtual world objects are presented together within a single display” [48, p. 1322] to one in which “real world and virtual world objects and stimuli are presented together within a single *percept*” [70, p. 4]. According to this interpretation, the ARTV continuum equally specifies not just a multitude of possibilities of ARTV systems, but also a multitude of ARTV experiences resulting from the perception of integrated TV and AR content, among which the ones mentioned in Subsection 2.3. In this context, the ARTV continuum represents the basis for designing the UX of transitions between various forms of ARTV. Our experiment, presented next, examines such experiences.

3 EXPERIMENT

We conducted a controlled experiment to understand the UX of ARTV journeys. To this end, we employed the ARTV continuum [77] as the theoretical support and conceptual setting for specifying journeys in our experiment. In this continuum, augmentation of the physical world (e.g., the living room where the user is located) conjoins augmentation of the TV (the nature of the TV itself, from a physical electronic device to a virtual screen to screenless holograms) to generate a multitude of ARTV forms, from which we select several as our experimental conditions. We define a journey in the ARTV continuum as a transition between two points from the continuum implemented by the conditions of our experiment.

3.1 Participants

Fourteen volunteers (9 men, 5 women), representing young adults between 19 and 34 years ($M=25.9$, $SD=5.4$), participated in our study. Their self-reported average daily time dedicated to watching TV and video streaming platforms was between 1 and 7 hours ($M=3.1$, $SD=2.0$), and the most preferred TV genres were documentaries (71.4%), science & technology shows (64.2%), movies (57.1%), news (50%), and comedy/sitcoms (50%), respectively. Three participants reported having used HMDs before our study, and one participant reported having played Pokémon Go on their smartphone. More details about our sample of participants are presented in Figure 1.

3.2 Design

Our experiment was a within-subjects design with one factor, ARTV-SCENARIO, ordinal variable with five conditions representing increasingly progressing augmentations of the world with TV content, from no augmentation (scenario no. 0, our *control* condition) to fully immersive ARTV (scenario no. 4, the *augmented-living-room*):

- (0) The *control* condition is represented by watching a conventional TV screen with no AR content.
- (1) *Virtual-screen* reproduces the form factor of the TV screen from the *control* condition with a virtual TV placed in the

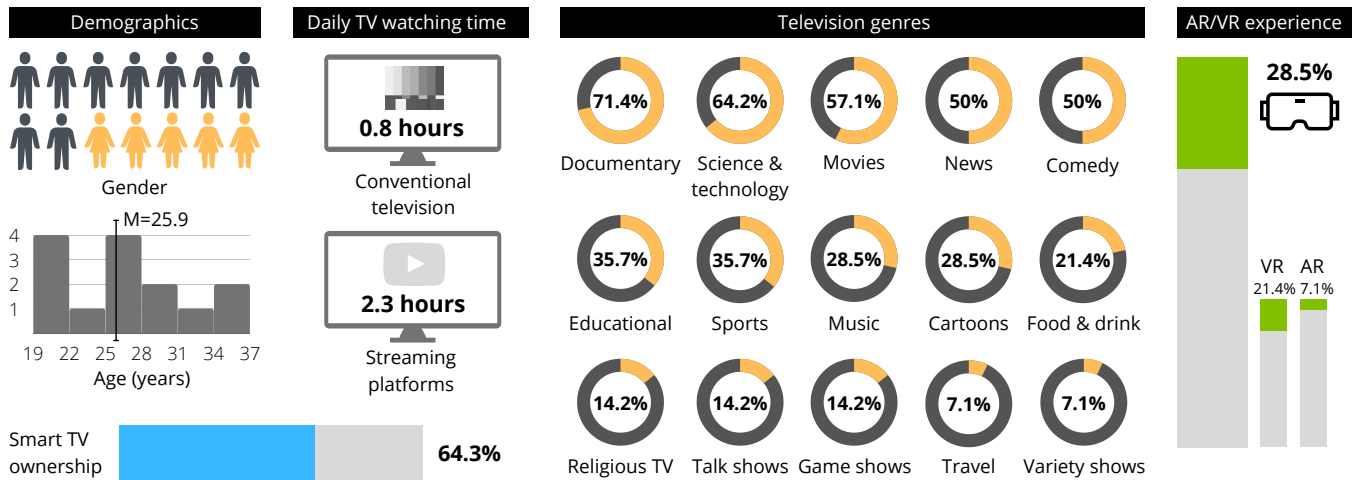


Figure 1: Participants' demographic details, TV watching profiles, and experience with AR/VR.

physical world and rendered by a HMD. The TV screen is the only virtual object from the AR scene.

- (2) *Off-screen-augmentation*, where some of the content is displayed around the virtual TV screen.
- (3) *Off-screen-augmentation-more*, where some of the content is displayed both around the virtual TV screen and in the room, between the TV and the user. More augmentations are present compared to condition (2), i.e., more characters and more of the scenery are displayed outside the TV frame.
- (4) In the *augmented-living-room* condition, the user is fully immersed in the AR content and there is no TV form factor.

Figure 2 positions these scenarios in the ARTV continuum, along with other research prototypes and commercial applications of ARTV, discussed in Section 2.1, included for context. The four ARTV scenarios that we chose for our experiment cover areas from the ARTV continuum that have been little examined before, such as the center (majoritarily targeted for immersive video game experiences [28,34] and little for TV) and the top part of the continuum involving virtual TVs with flexible form factors [72]. Condition no. 1, *virtual-screen*, implements a simple form of ARTV, i.e., in Mark Zuckerberg's words from the Facebook F8 Developer Conference of 2017, "You want to watch TV? We could put a digital TV on that wall and instead of being a piece of hardware, it's a \$1 app, instead of a \$500 piece of equipment" [71, minute 4:35]. Conditions no. 2 and 3, *off-screen-augmentation* and *off-screen-augmentation-more*, represent instances of Popovici *et al.*'s [63,64] scenario "I would like to be able to control and interact with AR content displayed around or in front of the TV set," which ranked first in a multi-cultural study with 319 participants. Condition no. 4, *augmented-living-room*, draws inspiration from highly immersive video games [28,34].

3.3 Apparatus

We developed a HoloLens application starting from an open-source Unity/Blender project¹ featuring a short animation in VR, to which we added a custom C# script to implement the four augmentations

¹https://gitlab.com/avinash-vadlamudi/Animation_Movie

represented by the conditions of the ARTV-SCENARIO independent variable. The topic of the VR movie was planet Earth in danger, while the human kind joined forces with an alien species to save it. The Unity scene consisted of a total number of thirty-eight GameObjects,² of which we employed three (7.9% of the content) for the *off-screen-augmentation* ARTV scenario, seven (18.4% of the content) for *off-screen-augmentation-more*, and all of the virtual objects for *augmented-living-room*. We implemented transitions between all of the ARTV scenarios, where the scenarios can be seen as the nodes of an undirected graph; see Figure 3 for a visual illustration. To transit from one form of ARTV to another, we employed the numerical keys "1" to "4" of a Bluetooth keypad connected to HoloLens. The experiment took place in a physical space of approximately 4m×4m in a controlled laboratory setting; see Figure 4.

3.4 Task

After signing the consent form, participants filled out a questionnaire with demographic information and reported their daily time spent watching TV and online streaming platforms (Figure 1). Before using our HoloLens application, participants were invited to sit on a chair placed at about 4m from a conventional TV screen (a 55-inch Samsung UE55D display), on which they watched a short, 90-second movie³ generated from our Unity VR scene representing the *control* condition of the ARTV-SCENARIO variable. Then, the participants were presented with the HoloLens HMD and our application. They were asked to rewatch the movie, which was augmented with new content represented by new scenery and characters, according to the various conditions of ARTV-SCENARIO, and to perform transitions between the different types of augmentations at will as long as they were paying attention to the action of the movie. Overall, the participants needed between 2.2 and 10.2

²Fundamental objects in Unity that represent characters, props, and scenery; see <https://docs.unity3d.com/Manual/GameObjects.html>.

³A duration that we chose in accordance to prior studies [68,74,81] that employed video clips of approximately 2, 3, and 5 minutes, respectively. According to Hawkins *et al.*'s [29] "hazard look" function, i.e., the probability that focused looks at television persisting a given length will terminate in the next half second, between 6 and 15 seconds are needed for engaged looks, while staring looks install after 15 seconds.

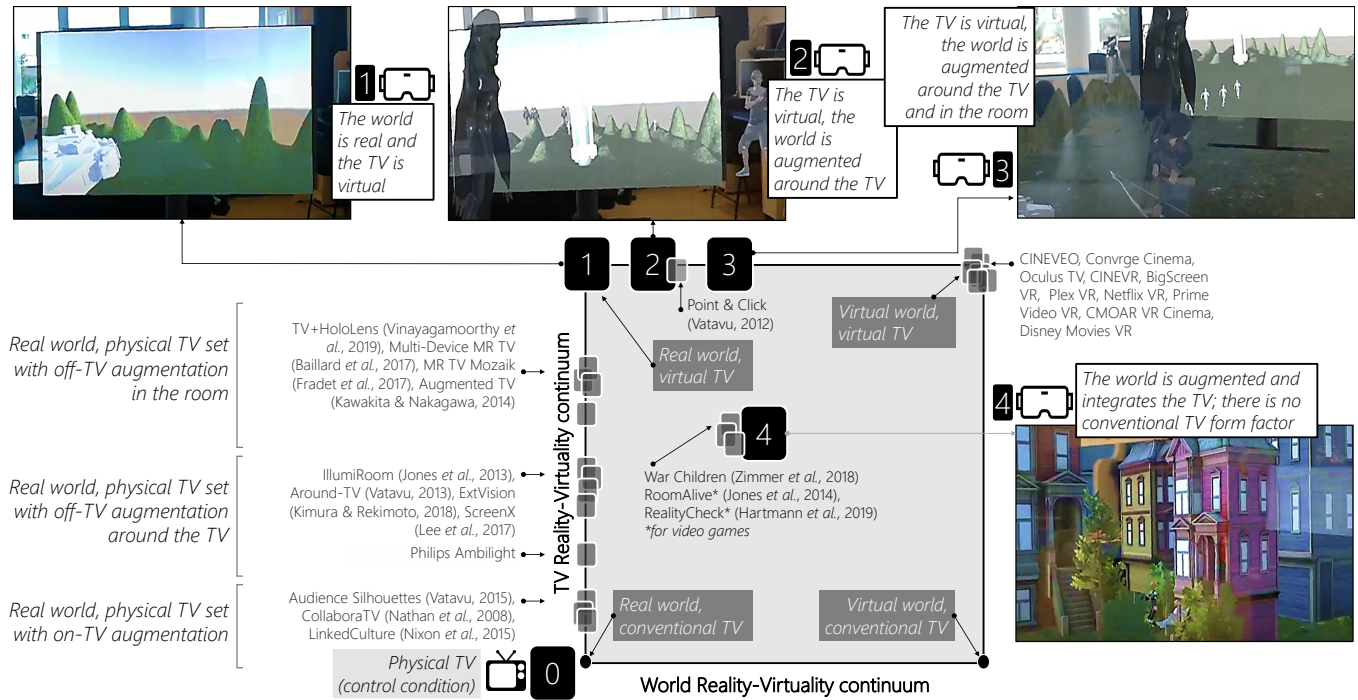


Figure 2: The conditions of our experiment positioned in the ARTV continuum [77]. Note how the left part of the continuum, involving conventional TV screens, is crowded with applications [4,20,35,36,41,51,53,72–74,81] as is the top-right corner representing watching TV/cinema in VR [2,10,15,17,19,30,50,52,59,60], but not the rest of the continuum. In this landscape of ARTV research prototypes and commercial applications, we target ARTV scenarios that have been little examined before, i.e., in the center of the ARTV continuum as well as augmentations involving virtual TV screens from the top of the continuum.

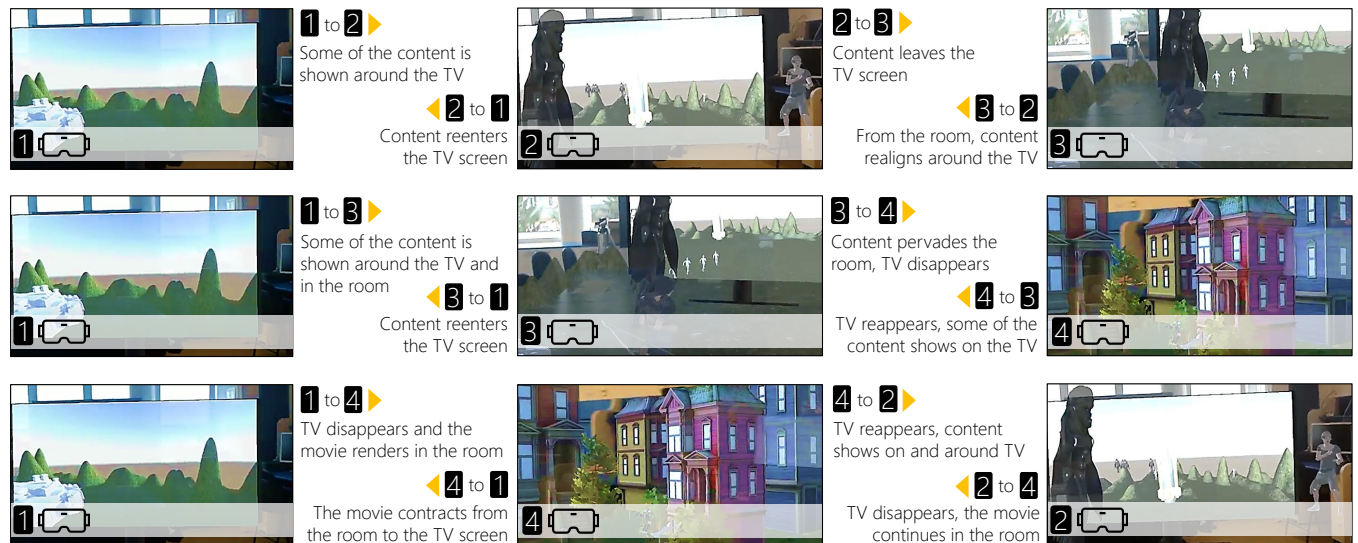


Figure 3: Twelve possibilities to transition between the four types of TV augmentations considered in our experiment, e.g., when moving from condition 1, virtual-screen to condition 3, off-screen-augmentation-more some of the characters and scenery exit from the virtual TV screen and are shown around the screen and in the room between the TV screen and the user.

minutes ($M=5.2$, $SD=2.3$) until they confirmed that they understood the action, at which point the experiment ended. A post-experiment

questionnaire collected measures of their experience; see next. On average, the experiment lasted 54 minutes ($SD=20.7$) per participant.



Figure 4: Photographs illustrating participants at various moments during the experiment, from lean back to lean forward engagement with the content, corresponding to the different conditions of the ARTV-SCENARIO independent variable.

3.5 Measures

We used a post-experiment questionnaire to collect various UX measures, representing the dependent variables in our experiment, for which we drew inspiration from prior studies about ARTV [35, 46, 63, 64, 68, 74, 81]. The measures capture diverse aspects of the benefits subjectively perceived by our participants regarding various forms of ARTV, e.g., ease of watching and the quality of ARTV being more fun compared to the *control* condition, among other UX constructs. We collected perceived benefits as agreement/disagreement reactions to the statement “Compared to the physical TV screen, the [name of the scenario] is [description of benefit],” as follows:

- **EASY-TO-WATCH**, binary variable, representing the reaction to the statement “Compared to the physical TV screen, the [name of the scenario] is easier to watch.”
- **INTERESTING**, binary variable, corresponding to the “more interesting” description over the *control* condition.
- **FUN**, corresponding to the “more fun” description.
- **INFORMATIVE**, corresponding to “more informative.”
- **ENGAGING**, corresponding to “more engaging with the TV.”

We also measured **ADOPTION-INTENTION**, a rating specifying the extent to which participants desired to have each of the ARTV-SCENARIOS at home, which we collected with a 5-point Likert scale with items ranging from 1 (“not at all”) to 5 (“very much”).

Besides the above measures that were collected for each ARTV-SCENARIO, we also employed the following tests to understand the overall experience of transitions between scenarios:

- **USABILITY**, measured with the System Usability Scale [13]. SUS consists of ten statements that elicit the degree of agreement using 5-point Likert scales with items from 1 (“strongly disagree”) to 5 (“strongly agree”). Answers are aggregated into a score between 0 (low usability) and 100 (perfect).
- **DESIRABILITY**, measured with the Microsoft Reaction Cards [8] method. Participants were asked to describe their experience with ARTV journeys using any number of words from a large list (N=118), such as “attractive,” “easy to use,” “desirable,” “exciting,” “confusing,” “overwhelming,” “distracting”; see [8].

- **TASK-LOAD**, measured with the NASA TLX test [27]⁴ to collect participants’ subjective ratings of perceived workload on six dimensions (mental demand, physical demand, temporal demand, performance, effort, and frustration). TLX returns a score from 0 (low) to 100 (high perceived workload).
- **IMMERSION**, measured with Rigby *et al.*’s [65] Immersive Experience Questionnaire for Film and TV (Film IEQ) for evaluating video viewing experiences in immersive environments. The test consists of 24 questions, e.g., “To what extent did the movie, TV show, or clip hold your attention?,” evaluated with 7-point Likert scales with items from 1 (“very little”) to 7 (“very much”). We adapted the questions to refer explicitly to ARTV and prefixed each question with the expression “Compared to a conventional TV,” e.g., “Compared to a conventional TV, to what extent did the movie watched in AR hold your attention?,” in order to measure the UX of ARTV journeys relative to the *control* condition. We normalized the total score by dividing it to 168 (=24 questions × 7, the maximum rating for a question), resulting an evaluation of IMMERSION from 0 (low) to 100 (high).
- **PRESENCE**, measured with Witmer and Singer’s [83] presence questionnaire (PQ v2.0) consisting of 32 questions, of which we used the 17 questions corresponding to the *involvement/control*, *natural*, and *interface quality* dimensions that were relevant to our scope.⁵ Just like for the IMMERSION test, we adapted the original questions to refer to ARTV and our *control* condition. For instance, the question “How much were you able to control events?” [83, p.232] was rephrased to “Compared to a conventional TV, how much were you able to control events in the AR movie?” We normalized the total score by dividing it to 119 (=17 questions × 7, the maximum rating of a question), resulting an evaluation of PRESENCE from 0 (low) to 100 (high).

We also logged all the transitions performed by our participants between various ARTV-SCENARIOS, and computed:

⁴We implemented the test using www.keithv.com/software/nasatlx/nasatlx.html.

⁵We removed the questions referring to the *haptic* and *auditory* dimensions.

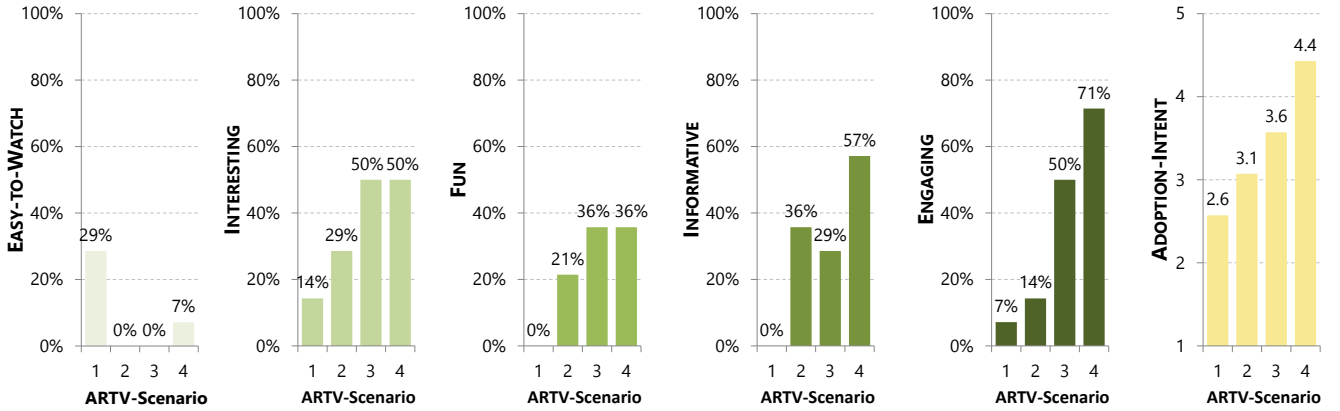


Figure 5: Perceived benefits of ARTV over the *control* condition (first five bar charts; higher percentages denote higher benefits) and the adoption intention for each ARTV scenario (bar chart from the right; higher values denote higher intentions).

- NUMBER-OF-TRANSITIONS, representing the total number of transitions between the four ARTV-SCENARIOS.
- TRANSITION-BEHAVIOR, representing observed user behavior during the transitions, for which we identified two patterns: (i) *linear*, i.e., scenarios are traversed in order from the most simple (1) to the most complex (4), and (ii) *principal*, i.e., frequent transitions are made to one of the four scenarios, to which a participant repeatedly returned.

4 RESULTS

We report our participants' experience with ARTV. We start by discussing the UX of each individual ARTV-SCENARIO and the perceived benefits over the conventional TV screen, and continue with the overall perception of transitions between ARTV scenarios.

4.1 The UX of Individual ARTV Scenarios

Figure 5 shows participants' characterizations of their experience with each of the four ARTV-SCENARIOS in terms of perceived benefits over the *control* condition represented by watching conventional TV. Overall, an increase in the perception of benefits delivered by ARTV can be identified as the level of augmentation increases, with the *augmented-living-room* condition scoring highest for the INTERESTING, FUN, INFORMATIVE, and ENGAGING measures. We found statistically significant effects of ARTV-SCENARIO on EASY-TO-WATCH (Cochran's $Q_{(3)}=9.923, p=.031$), FUN ($Q_{(3)}=8.739, p=.026$), INFORMATIVE ($Q_{(3)}=11.909, p=.007$), and ENGAGING ($Q_{(3)}=15.429, p=.001$), but not on the INTERESTING measure ($Q_{(3)}=7.200, p=.078$). Post-hoc Wilcoxon signed-rank tests conducted between pairs of progressively increasing augmentations (i.e., scenarios no. 1 and 2, 2 and 3, and 3 and 4, respectively) detected no statistically significant differences (Bonferroni-corrected level of significance $\alpha=.05/3=.017$), likely because of our small sample ($N=14$). Nevertheless, the main statistically significant effects are visible in Figure 5: more augmentations are perceived more interesting, fun, informative, and engaging compared to the *control* condition. Also, a Friedman test showed a statistically significant effect of ARTV-SCENARIO on ADOPTION-INTENTION ($\chi^2_{(3)}=21.939, p<.001$) with the *augmented-living-room*

scenario eliciting the most interest (Mdn=5, $M=4.40$, $SD=0.94$) and *virtual-screen* the least (Mdn=2.50, $M=2.57$, $SD=1.34$). Wilcoxon signed-rank tests (Bonferroni corrections at $\alpha=.05/3=.017$) revealed a significant difference between scenarios no. 3 and 4 ($p=.008$), but not between 1 and 2 or 2 and 3, respectively.

4.2 The UX of Journeys in the ARTV Continuum

Besides examining specific ARTV scenarios, we are interested in the UX of transitions between those scenarios. To this end, we report USABILITY, IMMERSION, PRESENCE, and TASK-LOAD; see Figure 6.

The average SUS score was 80.2 ($SD=12.1$), a high value falling under the fourth quartile, the "acceptable" range, and between "good" and "excellent" levels of usability, respectively, according to Bangor *et al.*'s [6] acceptability ranges and adjective ratings for interpreting SUS scores; see Figure 6, top left. Perceived TASK-LOAD was overall low ($M=37.2$, $SD=19.3$), with the lowest score obtained on the *frustration* subscale ($M=17.5$) and the highest on *mental demand* ($M=45.4$); see Figure 6, top right. To put the TLX results into perspective, our scores fall in the second quartile of a distribution of over one thousand TLX scores reported in academic publications [25], in the second quartile for computer activities, the first quartile for video game tasks, and the first quartile for cognitive tasks, respectively, denoting an overall low to medium perceived workload; see Grier's [25] meta-analysis of NASA TLX scores.

The average perceived IMMERSION was 67.9 ($SD=11.3$), situated in the second part of the measuring scale towards high immersion. The subscales of the IMMERSION experience, illustrated in Figure 6, bottom left, revealed high scores for *captivation* ($M=72.7$), representing viewer's enjoyment, how interested they were, and their motivation during ARTV journeys, and high *comprehension* ($M=71.9$), denoting how well the concepts and themes of the presented content were understood. The *real-world dissociation* and *transportation* subscales of the IMMERSION experience revealed lower scores ($M=55.4$ and 60.8 , respectively), indicating that viewers were aware of their real-world surroundings and felt they were located in the augmented world in a moderate manner. The average PRESENCE was 69.6 ($SD=10.1$), situated in the second part of the measuring scale towards high

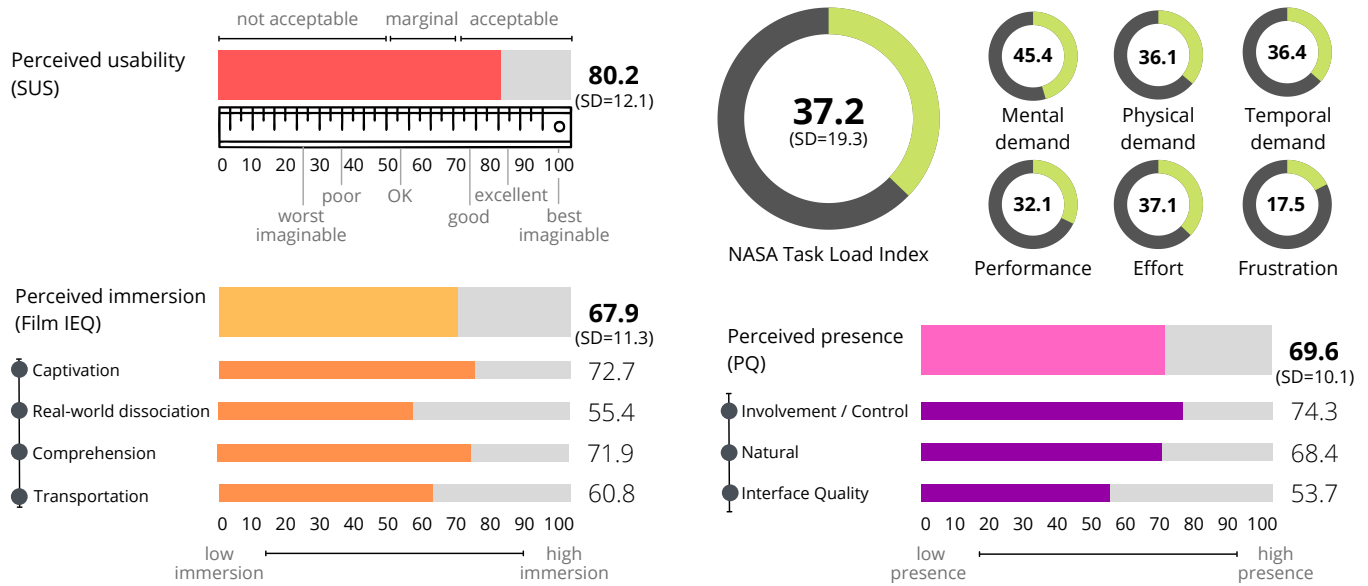


Figure 6: Perceived usability, workload, immersion, and presence for ARTV journeys compared to the conventional TV screen.



Figure 7: Word cloud describing the experience with ARTV journeys (N=264 words), generated using www.wordclouds.com.

levels of perceived presence; see Figure 6, bottom right. We found a relatively high perception of *involvement* with the virtual world ($M=74.3$) and of the extent to which interactions felt *natural* and the virtual world consistent with the physical reality ($M=68.4$). The *interface quality* subscale, representing the extent to which the participants felt able to concentrate on the task and were not distracted by controls/displays, indicated a moderate experience ($M=54.7$).

The positive UX of ARTV indicated by these measures is supported by the adjective descriptions obtained with the Reaction Cards test [8]. Figure 7 shows a word cloud generated from our participants' characterization of their experience with ARTV journeys. Out of a total of 264 words ($M=18.8$ words per participant), 93.9% had positive connotations, such as "attractive" (13 out of the 14 participants) "useful" (11/14) "creative" (9/14), "fun" (9/14), "exciting" (8/14), "comfortable" (7/14), or "impressive" (6/14), to name a few. A small number of 16 words (6.1%) described a more negative experience for some of the participants, such as "complex"

(5/14), “overwhelming” (2/14), “confusing,” “distracting,” and “time-consuming” (1/14), respectively. Overall, the experience with ARTV journeys was positively appreciated.

To complement these results obtained with self-reported measures, we looked at participants' transition behavior between the various ARTV-SCENARIOS (all the possible transitions are illustrated in Figure 3). Overall, our participants performed a total number of 158 transitions ($M=11.3$, $SD=5.8$). Most of the transitions occurred between scenarios no. 1 and 2, *virtual-screen* and *off-screen-augmentation* ($48/158=30.4\%$), followed by scenarios no. 2 and 3, *off-screen-augmentation* and *off-screen-augmentation-more* ($44/158=27.8\%$), and scenarios no. 3 and 4, *off-screen-augmentation-more* and *augmented-living-room* ($31/158=19.6\%$), respectively (both directions are counted, e.g., from scenario 3 to 4 and from 4 to 3); see Figure 8. This result suggests a tendency for a *linear* zapping, where participants transitioned most frequently between neighboring scenarios in terms of their level of augmentation, e.g., the sequence "1-2-3-4-2-3-1-2-3-4" for participant P₄ and "1-2-3-4-1-2-3-1-2-3-4-3-2-1-2-3-4" for P₉. We observed the linear zapping

	1	2	3	4	Transitions from
virtual screen 1	0	36	5	3	27.8%
off-screen augmentation 2	12	0	34	3	31.0%
off-screen augmentation-more 3	7	10	0	25	26.6%
augmented-living-room 4	13	4	6	0	14.6%
Transitions to	20.3%	31.6%	28.5%	19.6%	

Figure 8: Transitions between ARTV scenarios.

behavior for twelve participants (12/14=85.7%). A number of 35 transitions (35/158=22.2%) were not linear, such as between scenarios no. 1 and 4, *virtual-screen* and *augmented-living-room* (10.1%), scenarios no. 1 and 3, *virtual-screen* and *off-screen-augmentation-more* (7.6%), and scenarios no. 2 and 4, *off-screen-augmentation-more* and *augmented-living-room* (4.5%). We characterize these behaviors with the term *primary* in relation to the primary layout examined by Vatavu and Mancaş [75] for multi-screen TV, where one screen acts as the main screen catching most of the viewer’s attention, while the other screens are satellites to the primary screen. The main characteristic of the *primary* zapping behavior observed for ARTV journeys is returning repeatedly to one scenario, such as P₅ returned to scenario no. 2, *off-screen-augmentation* during their journey “1-2-1-2-3-2-4,” and P₁₀ returned repeatedly to the pair of scenarios 3 and 4, *off-screen-augmentation-more* and *augmented-living-room*, during their journey “1-3-4-1-2-3-4-3-4-2-1-4-3-2-1-2-3-4.”

5 DISCUSSION

Our findings revealed a UX of ARTV characterized by high perceived usability, low to medium workload, high captivation, comprehension, and involvement with the content that was complemented by a moderate feeling of transportation and dissociation from the physical world in accordance with the characteristic feature of AR to add to the physical world instead of completely immersing the user in a substitute reality [3,48]. In the following, we examine these findings with Garrett’s [21] model of UX; see Subsection 2.2 for a discussion of Garrett’s model in context.

5.1 ARTV and Garrett’s UX Five-Plane Model

From the most abstract to the most concrete, the five planes from Garrett’s [21] model of UX are *strategy*, *scope*, *structure*, *skeleton*, and *surface*, covering the UX process from user needs and organization goals to visual aspects of the finished product.

5.1.1 Strategy. At this level, user needs and product objectives are specified. The user profile that we addressed in this work was that of young adults, generally more willing to adopt and use new computer technology, including AR/VR, compared to other age groups. Although our findings are limited to this group alone, they show a distinctive experience of ARTV compared to the control condition of the conventional TV screen, and set the foundation for more investigations in this direction. Future explorations should include

other user segments, such as older adults [12,56], people from different cultural backgrounds [64] and with various abilities [62] to extend our findings about the UX of ARTV. Also, ARTV journeys are interesting to examine in longitudinal studies to better understand user needs for ARTV and to clearly specify organization-level objectives, e.g., business goals, at the *strategy* plane.

5.1.2 Scope. Functional specifications and content requirements for the product are located at the *scope* level. Our focus was on journeys in the ARTV continuum, which translates into the functional specification that the ARTV system should allow transitions from augmentation “A” to augmentation “B,” where “A” and “B” are identifiable points in the ARTV continuum. However, the corresponding content requirements may be different for “A” and “B,” e.g., how much of the TV content to augment (on the vertical axis of the ARTV continuum) and how much of the physical world is augmented (on the horizontal axis); see Figure 2. Also, the expected level of the augmentation, from a simple visual notification (e.g., a popup message rendered in smartglasses about breaking news on channel “6”) to stylized content to realistic looking scenery and characters to photometric lighting and reflections integrating the physical space (e.g., photorealistic content that adopts the lighting sources from the living room), is an important aspect when designing appropriate vehicles to deliver augmented TV content. This also includes choices regarding AR/VR display technology, from simple on-TV augmentations [53,74] to second screens and mobile devices [4,9,18] to smartglasses and HMDs [4,80,81] to video projections in the living room [35,42,73]. Important future work is to outline functional specifications for various regions of the ARTV continuum together with requirements for the content as well as the AR/VR rendering technology to deliver that content.

5.1.3 Structure. At this level, user interactions with the system and the organizational scheme for information that enables users to move through content are specified. In our experiment, we observed that some ARTV scenarios, such as *virtual-screen* and *off-screen-augmentation*, favor “lean back” viewing, while others, such as *augmented-living-room*, are principally “lean forward” and invite users to stand up, walk around, and explore the story surrounding them and unfolding in the room; see Figure 4. Expectations and needs for interacting with ARTV content are thus different for different types of ARTV as are the needs for the structural design of the information space to facilitate access to the augmented TV content. Conjoint explorations of interaction design and information design, e.g., interactions for lean back and lean forward TV watching in the context of various levels of augmentations as information design, are interesting to examine for ARTV transitions.

5.1.4 Skeleton. The *skeleton* plane defines the form in which *structure* becomes concrete and provides users with the ability to do things (i.e., interface design), to access various parts of the interface and content (i.e., navigation design), and involves communicating ideas to the user (i.e., information design); see Garrett [21, p. 109]. In our implementation, navigation design was kept simple (walking in the room to explore the augmented content and using the Bluetooth keyboard as an input modality to implement transitions between ARTV scenarios) so that participants could focus on the content. Input modalities specific to AR/VR environments [11,45], such as

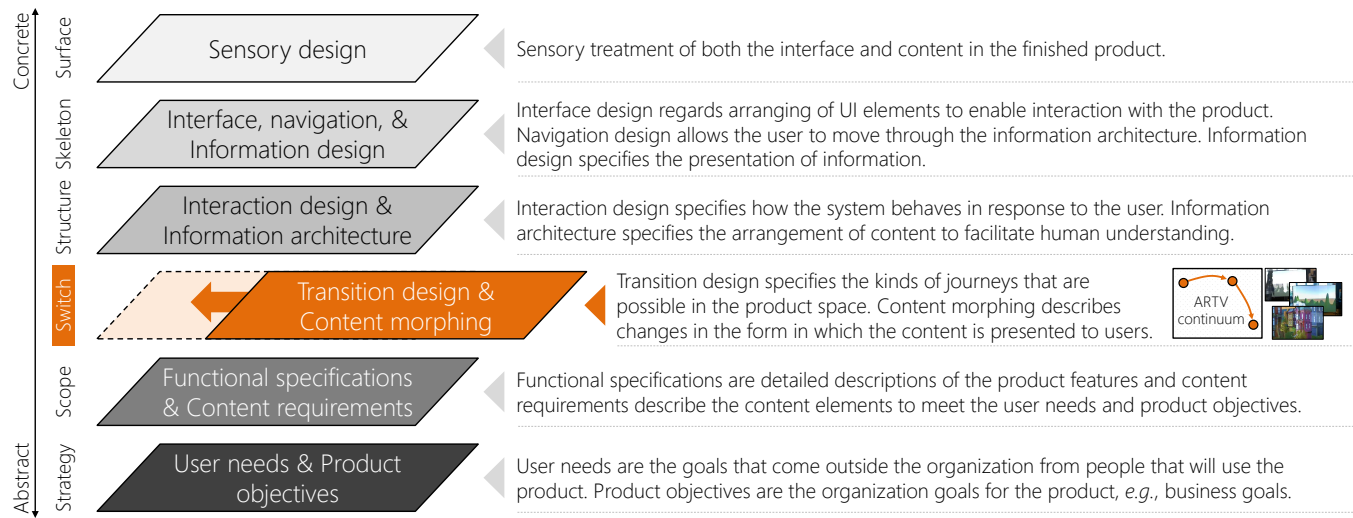


Figure 9: The new UX plane, *switch*, which we add to Garrett’s [21] five-plane model, conveniently separates at conceptual level ARTV journeys from generic functional specifications and specific aspects of the interaction and arrangement of content.

voice and gesture, should be further examined for more nuanced interactions during ARTV journeys. For example, interactions designed to control various parameters of ARTV, e.g., perspective and level of augmentation, or the development of the televised show towards exploring alternative plots, such as interactive movietelling [26] by recombination of story segments.

5.1.5 Surface. The *surface* plane addresses the sensory experience of the user. At this level, characteristics of the virtual content rendered in the augmented world, such as the level of photorealism, affect users’ emotional engagement [84]. Our implementation addressed visual output delivered by HoloLens, but other output channels are interesting to explore in future work towards multi-sensory experiences [78] applied to TV and storytelling [1,37,79]. Opportunities for multimodal experiences are already present in the ARTV continuum [77], for which the axes are open in terms of the sensory channels targeted by the augmentation; see Milgram and Kishino [48] for auditory, haptic, and vestibular AR.

5.2 A Sixth Plane for the UX of ARTV Journeys

The previous subsection showed how the ARTV UX can be described using Garrett’s [21] five-plane model, and included directions for future work at the level of each UX plane. However, the key feature of ARTV journeys implemented by our HoloLens application has not been covered by our discussion so far. Such journeys could be addressed at the *structure* plane during interaction design and information architecture design, while the specifics implemented at the *skeleton* and *surface* planes. However, we argue that the need of a distinct UX plane emerges for ARTV journeys because of the nature of different forms of ARTV involved in the journeys. These forms actually represent standalone ARTV products that deliver distinct UX to their users, as our findings from Figure 5 suggest. From this perspective, journeys in the ARTV continuum are more than an interaction feature to be specified at the *structure*

plane, but actually a bridge between different ARTV products that exist in the ARTV continuum, each with their own UX.

To address this aspect, we conceptualize journeys in augmented worlds at the level of a new UX plane, which we call “*switch*” to adhere to the letter “S” terminology employed by Garrett [21], and which we position between *structure* and *scope*; see Figure 9. At the *switch* level, a journey describes the capacity of the ARTV system to present augmented TV content differently and the possibility for users to control transitions between different augmentations. We differentiate between “product as functionality” and “product as information,” following Garrett’s [21] approach. Regarding the former, we are interested in the types of transitions that are possible in the ARTV continuum, i.e., “journey design.” Regarding the latter, the characteristics of the content presented to the user come into focus to enable morphing from one ARTV form to another, e.g., from *virtual-screen* to *off-screen-augmentation* to *augmented-living-room*, constituting “content morphing design.” Although precedents exist for switching between realities in the VR literature [39,44,48,66,82], such as Milgram and Colquhoun’s [47] description of journeys along the Reality-Virtuality axis, they have not been formalized at the UX level. Moreover, switching is new in terms of understanding the UX of ARTV, where only specific ARTV scenarios have been examined; see our discussion from Subsection 2.3.

The *switch* plane of the ARTV UX is convenient to separate at a conceptual level the transitional aspect from the more general functional specifications and content requirements found at the *scope* level and the more specific outlining of interactions and arrangement of content from the *structure* level. This conceptual separation is especially useful when the number of possible journeys is large. For example, our HoloLens application implemented four types of augmentations representing 12 possible transitions (Figure 3), which represent just a few of the possibilities offered by the ARTV continuum [77]; see Popovici *et al.* [63,64] for twenty

scenarios of ARTV, which amount to a number of $20 \times 19 = 380$ transitions. Such specifications can be addressed at the *switch* plane, without agglomerating or despecializing *scope* or *structure*. Moreover, we estimate that switching between various ARTV scenarios may have an impact on user attention, where such aspects could also be addressed at the *switch* level in correspondence to Bakker and Niemantsverdriet's [5] interaction-attention continuum. In this continuum, interfaces facilitate interaction at varied attentional levels—focused, peripheral, and implicit—as well as switching between them in AR [69]. To exemplify the *switch* plane, we resume the discussion of our two examples from Section 1.

5.2.1 One user, multiple augmentations. In the first example, a user switches between four levels of augmentation of a soccer game: from the conventional TV screen to information about the teams displayed next to the TV to a live sports commentator in the room to replays of key moments from the game as immersive holograms in the living room. By specifying the twelve transitions that are possible with these four scenarios at the *switch* level, design possibilities for the journey and content morphing during the journey, respectively, become conceptually separated from the adjacent levels. For example, the linear zapping behavior for ARTV, largely observed in our experiment, could represent the default option for the ARTV system to progressively increase the level of augmentation during the live soccer transmission. At various moments during the game, the system could automatically and adaptively increase or decrease the augmentation, from displaying information and statistics about the two teams to bringing into view the live sports commentator. This system behavior can be specified in terms of transitions in the ARTV continuum at the *switch* plane, without affecting the functional and content requirements of *scope* or the interaction design and information architecture of the *structure* plane.

5.2.2 Multiple users, multiple augmentations. In the second example, a friend joins the living room from a remote location through holographic teleportation to discuss the mystery TV show in the context of social TV watching. For the first user, this means transitioning to an increased augmentation of the physical world, specified by a transition to the right on the horizontal dimension of the ARTV continuum, which is independent of the current level of TV augmentation. Such a transition is more convenient conceptually to position at the *switch* plane, instead of agglomerating *scope* and *structure*. Moreover, the two users could watch the same show at different levels of augmentation: while the story is the same, its presentation and corresponding level of the augmentation are different for each user. Transitioning from the augmentation of one user to the other's or synchronizing the two augmentations are design aspects that are characteristic to the *switch* level. As the number of users increases, specifying all the possible transitions at the *switch* plane is handled independently of *scope* and *structure*.

6 CONCLUSION

We addressed in this work journeys in the ARTV continuum, which we characterized with a diversified set of UX measures and discussed in the context of Garrett's [21] five-plane model of UX. Our empirical findings revealed an UX of ARTV characterized by high perceived usability, captivation, comprehension, and involvement

with a low to medium workload and moderate feeling of dissociation from the physical world. Our explorations led to the proposal of a sixth UX plane to address transitions in the ARTV continuum with conceptual convenience. We hope that our results will inspire future work where the immersive technology of AR/VR intersects TV to accommodate different users' preferences, needs, and abilities to consume augmented TV content.

ACKNOWLEDGMENTS

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project number PN-III-P4-ID-PCE-2020-0434 (PCE29/2021), within PNCDI III. The authors acknowledge support from OSF Global Services, the Mobile Division, Suceava that kindly provided the HoloLens HMD that was used for the experiment reported in this work.

REFERENCES

- [1] Damien Ablart, Carlos Velasco, and Marianna Obrist. 2017. Integrating Mid-Air Haptics into Movie Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17)*. ACM, New York, NY, USA, 77–84. <https://doi.org/10.1145/3077548.3077551>
- [2] Amazon. 2019. *Prime Video VR*. Amazon. Retrieved January 2022 from <https://www.amazon.com/adlp/primevideovr>
- [3] 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>
- [4] 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>
- [5] Saskia Bakker and Karin Niemantsverdriet. 2016. The Interaction-Attention Continuum: Considering Various Levels of Human Attention in Interaction Design. *International Journal of Design* 10, 2 (2016), 1–14. <http://www.ijdesign.org/index.php/IJDesign/article/view/2341/737>
- [6] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies* 4, 3 (May 2009), 114–123. <https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/>
- [7] BBC Research & Development. 2013. *Unconventional Screens. Exploring the potential of future display technologies*. BBC Research & Development. Retrieved January 2022 from <https://www.bbc.co.uk/rd/projects/unconventional-screens>
- [8] Joey Benedek and Trish Miner. 2002. Measuring Desirability: New Methods for Evaluating Desirability in a Usability Lab Setting. In *Proceedings of Usability Professionals' Association Conference*. Microsoft, Redmond, Washington, USA, 5 pages. <https://www.microsoft.com/usability/uepostings/desirabilitytoolkit.doc>
- [9] Toni Bibiloni, Miquel Mascaró, Pere Palmer, and Antoni Oliver. 2015. A Second-Screen Meets Hypervideo, Delivering Content Through HbbTV. In *Proceedings 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>
- [10] BigScreen. 2016. *Your Ultimate Virtual Reality Hangout*. BigScreen, Inc. Retrieved January 2022 from <https://www.bigscreenvr.com>
- [11] Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A Survey of Augmented Reality. *Foundations and Trends in Human-Computer Interaction* 8, 2–3 (March 2015), 73–272. <http://dx.doi.org/10.1561/11000000049>
- [12] Jan Bobeth, Susanne Schmehl, Ernst Kruijff, Stephanie Deutsch, and Manfred Tscheligi. 2012. Evaluating Performance and Acceptance of Older Adults Using Freehand Gestures for TV Menu Control. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. ACM, New York, NY, USA, 35–44. <https://doi.org/10.1145/2325616.2325625>
- [13] John Brooke. 1996. SUS: A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*, Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester (Eds.). CRC Press, London, UK, 189–194. <https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke>
- [14] Xiaotong Chen and Misha Sra. 2021. IntoTheVideos: Exploration of Dynamic 3D Space Reconstruction From Single Sports Videos. In *The Adjunct Publication of the 34th Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 14–16. <https://doi.org/10.1145/3474349.3480215>
- [15] CINEVR. 2017. *CINEVR: Movie Theater on Demand in Virtual Reality*. CINEVR. Retrieved January 2022 from <https://cinevr.io/en>

- [16] Cédric Courtois and Evelien D'heer. 2012. Second Screen Applications and Tablet Users: Constellation, Awareness, Experience, and Interest. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroTV '12)*. ACM, New York, NY, USA, 153–156. <https://doi.org/10.1145/2325616.2325646>
- [17] Disney. 2016. *Disney Movies VR*. Disney. Retrieved January 2022 from <https://www.disneymoviesvr.com>
- [18] eyecandylab. 2019. *The Power of Video-Triggered AR*. EyeCandy Lab. Retrieved January 2022 from https://augmen.tv/docs/eyecandylab_ThePowerofVideoTriggeredAR_September2019.pdf
- [19] Facebook Technologies. 2019. *Oculus TV on Oculus Quest 2*. Facebook. Retrieved January 2022 from <https://www.oculus.com/experiences/quest/1931356740318898>
- [20] Matthieu Fradet, Caroline Baillard, Anthony Laurent, Tao Luo, Philippe Robert, Vincent Alleaume, Pierrick Jouet, and Fabien Servant. 2017. MR TV Mozaik: A New Mixed Reality Interactive TV Experience. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, Washington, DC, USA, 155–159. <https://doi.org/10.1109/ISMAR-Adjunct.2017.53>
- [21] Jesse James Garrett. 2010. *The Elements of User Experience: User-Centered Design for the Web and Beyond* (2nd ed.). New Riders Publishing, USA. <https://www.oreilly.com/library/view/the-elements-of/9780321688651>
- [22] Jesse James Garrett. 2021. *I Helped Pioneer UX Design. What I See Today Disturbs Me. Where Did We Go Wrong?* Fast Company Inc. <https://www.fastcompany.com/90642462/i-helped-pioneer-ux-design-what-i-see-today-horrifies-me>
- [23] David Geerts, Evert van Beek, and Fernanda Chocron Miranda. 2019. Viewers' Visions of the Future. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 59–69. <https://doi.org/10.1145/3317697.3323356>
- [24] Jeff Gothelf and Josh Seiden. 2013. *Lean UX: Applying Lean Principles to Improve User Experience*. O'Reilly Media, Inc., Sebastopol, CA, USA. <https://www.oreilly.com/library/view/lean-ux/9781449366834/>
- [25] Rebecca A. Grier. 2016. How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. *Proceedings of the 59th Annual Meeting of the Human Factors and Ergonomics Society* 59, 1 (2016), 1727–1731. <https://doi.org/10.1177/2F1541931215591373>
- [26] Fabrizio Guerrini, Nicola Adami, Sergio Benini, Alberto Piacenza, Julie Porteous, Marc Cavazza, and Riccardo Leonardi. 2017. Interactive Film Recombination. *ACM Trans. Multimedia Comput. Commun. Appl.* 13, 4, Article 52 (2017), 22 pages. <https://doi.org/10.1145/3103241>
- [27] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52 (1988), 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [28] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In *Proceedings of the CHI Conf. on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300577>
- [29] Robert P. Hawkins, Suzanne Pingree, Jacqueline Hitchon, Barry Radler, Bradley W. Gorham, Leeann Kahlor, Eileen Gilligan, Ronald C. Serlin, Toni Schmidt, Prathana Kannaovakun, and Gudbjorg Hildur Kolbeins. 2006. What Produces Television Attention and Attention Style?: Genre, Situation, and Individual Differences as Predictors. *Human Communication Research* 31, 1 (2006), 162–187. <https://doi.org/10.1111/j.1468-2958.2005.tb00868.x>
- [30] Scott Hayden. 2015. *Cinemas Social Home Theater Space for Oculus Rift Launches from Creators of Convrge*. Convrge. Retrieved January 2022 from <https://www.roadtovr.com/convrge-launches-social-home-theater-space-convrge-cinemas>
- [31] Jiwoo Hong and Woohun Lee. 2021. Augmenting Television Narratives Using Interactive Home Appliances. *Digital Creativity* n.a., n.a. (2021), 1–14. <https://doi.org/10.1080/14626268.2021.2013898>
- [32] InAIR. 2015. *InAIR TV*. InAIR. Retrieved January 2022 from https://www.youtube.com/watch?v=z3xIEC_m9Fg
- [33] International Organization for Standardization. 2019. *Ergonomics of Human-System Interaction-Part 210: Human-Centred Design for Interactive Systems*. ISO. Retrieved January 2022 from <https://www.iso.org/standard/77520.html>
- [34] 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 *Proceedings 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>
- [35] Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proceedings 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>
- [36] Hiroyuki Kawakita and Toshio Nakagawa. 2014. Augmented TV: An Augmented Reality System for TV Programs Beyond the TV Screen. In *Proceedings of the International Conference on Multimedia Computing and Systems (ICMCS '14)*. IEEE, Washington, DC, USA, 955–960. <https://doi.org/10.1109/ICMCS.2014.6911158>
- [37] Rohit Ashok Khot and Jung-Ying (Lois) Yi. 2020. GustaCine: Towards Designing a Gustatory Cinematic Experience. In *Proceedings of the 14th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '20)*. ACM, New York, NY, USA, 757–770. <https://doi.org/10.1145/3374920.3375010>
- [38] Naoki Kimura and Jun Rekimoto. 2018. ExtVision: Augmentation of Visual Experiences with Generation of Context Images for a Peripheral Vision Using Deep Neural Network. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–10. <https://doi.org/10.1145/3173574.3174001>
- [39] Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. 2000. Seamless Design for 3D Object Creation. *IEEE MultiMedia* 7, 1 (jan 2000), 22–33. <https://doi.org/10.1109/93.839308>
- [40] Jake Knapp, John Zeratsky, and Braden Kowitz. 2016. *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*. Simon & Schuster, New York, NY, USA. <https://www.thesprintbook.com/book>
- [41] Jungjin Lee, Sangwoo Lee, Younghui Kim, and Junyong Noh. 2017. ScreenX: Public Immersive Theatres with Uniform Movie Viewing Experiences. *IEEE Transactions on Visualization and Computer Graphics* 23, 2 (2017), 1124–1138. <https://doi.org/10.1109/TVCG.2016.2532327>
- [42] Asterios Leonidis, Maria Korzi, Vasilios Kouroumalis, Emmanouil Adamakis, Dimitrios Milathianakis, and Constantine Stephanidis. 2021. Going Beyond Second Screens: Applications for the Multi-Display Intelligent Living Room. In *Proceedings of the ACM Int. Conference on Interactive Media Experiences (IMX '21)*. ACM, New York, NY, USA, 187–193. <https://doi.org/10.1145/3452918.3465486>
- [43] Valentin Lohmüller and Christian Wolff. 2019. Towards a Comprehensive Definition of Second Screen. In *Proceedings of Mensch Und Computer 2019 (MuC'19)*. ACM, New York, NY, USA, 167–177. <https://doi.org/10.1145/3340764.3340781>
- [44] Celina Ma, Haohong Wang, Hao Sun, Elliot van Huijgevoort, Mea Wang, and Zhihai He. 2021. Powering TV Experiences with Anytime Environmental Exploration. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 363, 6 pages. <https://doi.org/10.1145/3411763.3451600>
- [45] Daniel Martin, Sandra Malpica, Diego Gutierrez, Belen Masia, and Ana Serrano. 2021. Multimodality in VR: A Survey. *ACM Comput. Surv.* n.a., n.a. (2021), 34 pages. <https://doi.org/10.1145/3508361>
- [46] Mark McGill, John H. Williamson, and Stephen Brewster. 2016. Examining The Role of Smart TVs and VR HMDs in Synchronous At-a-Distance Media Consumption. *ACM Trans. Comput.-Hum. Interact.* 23, 5, Article 33 (nov 2016), 57 pages. <https://doi.org/10.1145/2983530>
- [47] 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
- [48] 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
- [49] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers 2351, Telemanipulator and Telepresence Technologies*, Vol. 2351. SPIE, USA, 282–292. <https://doi.org/10.1117/1.197321>
- [50] Mindprobe Labs. 2016. *CINEVEO - VR Cinema*. Cineveo. Retrieved January 2022 from <https://www.facebook.com/cineveovr>
- [51] Mukesh Nathan, Chris Harrison, Svetlana Yarosh, Loren Terveen, Larry Stead, and Brian Amento. 2008. CollaboraTV: Making Television Viewing Social Again. In *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video (UXTV '08)*. ACM, New York, NY, USA, 85–94. <https://doi.org/10.1145/1453805.1453824>
- [52] Netflix. 2018. *Netflix VR*. Netflix. Retrieved January 2022 from https://play.google.com/store/apps/details?id=com.netflix.android_vr
- [53] Lyndon Nixon, Lotte Belice Baltussen, and Johan Oomen. 2015. LinkedCulture: Browsing Related Europeana Objects While Watching a Cultural Heritage TV Programme. In *Proceedings of the 8th International Conference on Personalized Access to Cultural Heritage - Volume 1352 (PATCH '15)*. CEUR-WS.org, Atlanta, USA, 37–40. <http://ceur-ws.org/Vol-1352/paper7.pdf>
- [54] Don Norman and Jakob Nielsen. n.d.. *The Definition of User Experience (UX)*. Nielsen Norman Group. Retrieved January 2022 from <https://www.nngroup.com/articles/definition-user-experience>
- [55] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. 2016. Holoporation: Virtual 3D Teleportation in Real-Time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 741–754. <https://doi.org/10.1145/2984511.2984517>
- [56] Filomena Papa, Bartolomeo Sapio, and M. Felicia Pelagalli. 2011. User Experience of Elderly People with Digital Television: A Qualitative Investigation. In *Proceedings of the 9th European Conference on Interactive TV and Video (EuroTV '11)*.

- ACM, New York, NY, USA, 223–226. <https://doi.org/10.1145/2000119.2000165>
- [57] Hyerim Park, Maryam Shakeri, Ikbeom Jeon, Jangyoon Kim, Abolghasem Sadeghi-Niaraki, and Woontack Woo. 2022. Spatial Transition Management for Improving Outdoor Cinematic Augmented Reality Experience of the TV Show. *Virtual Reality* n.a., n.a. (2022), 19 pages. <https://doi.org/10.1007/s10055-021-00617-z>
- [58] Hyerim Park and Woontack Woo. 2015. Metadata Design for Location-Based Film Experience in Augmented Places. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. IEEE, Washington, DC, USA, 40–45. <https://doi.org/10.1109/ISMAR-MASHD.2015.12>
- [59] Pixel Edge Games. 2016. *CMOAR VR Cinema*. Valve Corporation. Retrieved January 2022 from https://store.steampowered.com/app/527160/Cmoar_VR_Cinema
- [60] Plex. 2018. *Virtual Reality: Interactive Environments, Watching Together, Voice Chat, and More Bring Your Media to Life*. Plex. Retrieved January 2022 from <https://www.plex.tv/your-media/virtual-reality>
- [61] Irina Popovici and Radu-Daniel Vatavu. 2019. Consolidating the Research Agenda of Augmented Reality Television with Insights from Potential End-Users. In *Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Washington, D.C., USA, 73–74. <https://doi.org/10.1109/ISMAR-Adjunct.2019.00033>
- [62] Irina Popovici and Radu-Daniel Vatavu. 2019. Towards Visual Augmentation of the Television Watching Experience: Manifesto and Agenda. In *Proceedings of the 2019 ACM International 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>
- [63] Irina Popovici and Radu-Daniel Vatavu. 2019. Understanding Users' Preferences for Augmented Reality Television. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Washington, DC, USA, 269–278. <https://doi.org/10.1109/ISMAR.2019.00024>
- [64] Irina Popovici, Radu-Daniel Vatavu, Pu Feng, and Wenjun Wu. 2021. AR-TV and AR-Diànshi: Cultural Differences in Users' Preferences for Augmented Reality Television. In *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '21)*. ACM, New York, NY, USA, 50–60. <https://doi.org/10.1145/3452918.3458801>
- [65] Jacob M. Rigby, Duncan P Brumby, Sandy J. J. Gould, and Anna L Cox. 2019. Development of a Questionnaire to Measure Immersion in Video Media: The Film IEQ. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 35–46. <https://doi.org/10.1145/3317697.3323361>
- [66] Joan Sol Roo and Martin Hachet. 2017. One Reality: Augmenting How the Physical World is Experienced by Combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 787–795. <https://doi.org/10.1145/3126594.3126638>
- [67] Pejman Saeghe, Gavin Abercrombie, Bruce Weir, Sarah Clinch, Stephen Pettifer, and Robert Stevens. 2020. Augmented Reality and Television: Dimensions and Themes. In *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '20)*. ACM, New York, NY, USA, 13–23. <https://doi.org/10.1145/3391614.3393649>
- [68] 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 *Proceedings 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>
- [69] Ovidiu-Andrei Schipor, Radu-Daniel Vatavu, and Wenjun Wu. 2019. Integrating Peripheral Interaction Into Augmented Reality Applications. In *Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Washington, D.C., USA, 358–359. <https://doi.org/10.1109/ISMAR-Adjunct.2019.00-12>
- [70] Richard Skarbez, Missie Smith, and Mary C. Whitton. 2021. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality* 2 (2021), 27. <https://doi.org/10.3389/frv.2021.647997>
- [71] TIME. 2017. *Facebook Reveals Augmented Reality, Virtual Reality and More at Annual F8 Conference*. Facebook. Retrieved January 2022 from <https://www.youtube.com/watch?v=YtUye84PuFY>
- [72] Radu-Daniel Vatavu. 2012. Point & Click Mediated Interactions for Large Home Entertainment Displays. *Multimedia Tools Appl.* 59, 1 (July 2012), 113–128. <https://doi.org/10.1007/s11042-010-0698-5>
- [73] Radu-Daniel Vatavu. 2013. There's a World Outside Your TV: Exploring Interactions beyond the Physical TV Screen. In *Proceedings 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>
- [74] Radu-Daniel Vatavu. 2015. Audience Silhouettes: Peripheral Awareness of Synchronous Audience Kinesics for Social Television. In *Proceedings 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>
- [75] Radu-Daniel Vatavu and Matei Mancas. 2014. Visual Attention Measures for Multi-Screen TV. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14)*. ACM, New York, NY, USA, 111–118. <https://doi.org/10.1145/2602299.2602305>
- [76] Radu-Daniel Vatavu and Stefan-Gheorghe Pentiu. 2008. Interactive Coffee Tables: Interfacing TV within an Intuitive, Fun and Shared Experience. In *Proceedings of the European Conference on Interactive Television (EuroITV '08)*. Springer, Berlin, Heidelberg, 183–187. https://doi.org/10.1007/978-3-540-69478-6_24
- [77] Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. 2020. Conceptualizing Augmented Reality Television for the Living Room. In *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '20)*. ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3391614.3393660>
- [78] Carlos Velasco and Marianna Obrist. 2020. *Multisensory Experiences: Where the Senses Meet Technology*. Oxford University Press, Oxford, UK. <https://doi.org/10.1093/oso/9780198849629.001.0001>
- [79] Carlos Velasco, Yunwen Tu, and Marianna Obrist. 2018. Towards Multisensory Storytelling with Taste and Flavor. In *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction (MHFI'18)*. ACM, New York, NY, USA, Article 2, 7 pages. <https://doi.org/10.1145/3279954.3279956>
- [80] Vinoba Vinayagamoorthy, Maxine Glancy, Paul Debenham, Alastair Bruce, Christoph Ziegler, and Richard Schäffer. 2018. Personalising the TV Experience with Augmented Reality Technology: Synchronised Sign Language Interpretation. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 179–184. <https://doi.org/10.1145/3210825.3213562>
- [81] 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 *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300762>
- [82] Chiu-Hsuan Wang, Chia-En Tsai, Seraphina Yong, and Liwei Chan. 2020. Slice of Light: Transparent and Integrative Transition Among Realities in a Multi-HMD-User Environment. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 805–817. <https://doi.org/10.1145/3379337.3415868>
- [83] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (06 1998), 225–240. <https://doi.org/10.1162/105474698565686>
- [84] Katja Zibrek, Sean Martin, and Rachel McDonnell. 2019. Is Photorealism Important for Perception of Expressive Virtual Humans in Virtual Reality? *ACM Trans. Appl. Percept.* 16, 3, Article 14 (sep 2019), 19 pages. <https://doi.org/10.1145/3349609>
- [85] Christian Zimmer, Nanette Ratz, Michael Bertram, and Christian Geiger. 2018. War Children: Using AR in a Documentary Context. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct*. IEEE, Washington, DC, USA, 390–394. <https://doi.org/10.1109/ISMAR-Adjunct.2018.00112>