Evaluating and Updating a Design Space for Augmented Reality Television

Pejman Saeghe
Department of Computer Science
The University of Manchester
Manchester, United Kingdom
pejman.saeghe@manchester.ac.uk

Mark McGill
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
mark.mcgill@glasgow.ac.uk

Bruce Weir
Research and Development
The British Broadcasting Corporation
Salford, United Kingdom
bruce.weir@bbc.co.uk

Sarah Clinch
Department of Computer Science
The University of Manchester
Manchester, United Kingdom
sarah.clinch@manchester.ac.uk

Robert Stevens
Department of Computer Science
The University of Manchester
Manchester, United Kingdom
robert.stevens@manchester.ac.uk

ABSTRACT

As Augmented Reality Television (ARTV) transitions out of the feasibility phase, it is crucial to understand the impact of design decisions on the viewers' ARTV experiences. In a previous study, six ARTV design dimensions were identified by relying on insights from existing prototypes. However, the set of possible dimensions is likely to be broader. Building on top of previous work, we create an ARTV design space and present it in a textual cheat sheet. We subsequently evaluate the cheat sheet in a between-subject study ($n = 10$), with participants with wide-ranging expertise. We identified six new dimensions (genre, broadcast mode, audience demographics, cartoonish vs. photoreal representation, modality, and privacy), and a new aspect (360°) for the display dimension. In light of our observations, we provide an updated ARTV design space and observe that asking participants to write ARTV scenarios can be an effective method for harvesting novel design dimensions.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

augmented reality, mixed reality, television, design space, qualitative analysis

ACM Reference Format:
Pejman Saeghe, Mark McGill, Bruce Weir, Sarah Clinch, and Robert Stevens. 2022. Evaluating and Updating a Design Space for 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, 15 pages. https://doi.org/10.1145/3505284.3529965

1 INTRODUCTION

The emergence of consumer-grade Augmented Reality (AR) devices has presented researchers and content creators with stimulating questions, such as *how, if at all, AR can be combined with TV, in the context of content creation, distribution, and consumption?* The works of Saeghe et al. [23] and Vatavu et al. [27] highlighted various ways in which researchers and content creators have attempted to combine AR with TV—a hybrid medium referred to as *Augmented Reality Television (ARTV)*. These include, but are not limited to, using AR to extend the real estate of the TV screen (e.g., [12, 14]); delivering story-related holographic artefacts for entertainment (e.g., [24]) and education (e.g., [22, 29]); delivering synchronised sign language interpretation [28]; delivering additional virtual TV screens and providing advanced remote-control functionality for TV content (e.g., [3, 26]); repurposing and delivering archived TV content based on viewers' context (e.g., [5, 25]); bridging intergenerational gaps through playful interaction (e.g., [19]); and even replacing the 2-D TV screen completely by delivering content in an unframed way (e.g., [31]).

A potential problem, however, with the current state of research is a tendency to focus on the feasibility aspect of a concept—typically conjured up by the researchers themselves—without taking into account audiences' or content creators' expectations and preferences. Exceptions to this trend include works such as Geerts et al. [7] who used a co-creation approach to develop two future TV scenarios, and Popovici and Vatavu [20] and Popovici et al. [21] who used 20 sentence-length predefined ARTV scenarios to elicit viewers' preferences regarding ARTV.

In this paper, we use six ARTV design dimensions previously identified by Saeghe et al. [23] as a starting point and expand the concepts to capture a wider spectrum of the theoretical design space underlying ARTV experiences; we present the design ideas as questions in a cheat sheet (Appendix F), and investigate the following research questions:

RQ 1: To what extent can a design space presented as questions in a textual cheat sheet format be easily understood?
RQ 2: To what extent can the aforementioned cheat sheet be used/applied to conceptualise novel ARTV experiences?

In a between-subject study, we asked $n = 10$ participants with wide-ranging expertise (e.g., researchers, engineers, and producers in TV and AR) to write two ARTV scenarios each, and subsequently interviewed them. While all participants received a *handout* outlining basic operational definitions, half the participants—*study*

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

group—also received *the ARTV design space cheat sheet* before writing the second scenario.

Our findings suggest that the ARTV design space cheat sheet can be useful in conceptualising novel ARTV scenarios, and by "providing a *checklist effect*"[1] it could ensure that no pertinent aspect of an ARTV experience is unintentionally left out.

Furthermore, we were able to expand the design space by identifying six new dimensions, repositioned a previously identified dimension—editorial control—as an aspect of the interaction dimension, and added 360° as an aspect to the display dimension.

Our findings contribute to the emerging field of ARTV; an ARTV design space described in an accessible way can provoke new ideas and concepts, and help novices to better grasp the possibilities afforded by this hybrid medium.

Our contributions in this work are:

1) A refined ARTV cheat sheet, demonstrated to promote conceptualisation of novel ARTV concepts.
2) An expanded ARTV design space, incorporating novel ARTV experiences as conceptualised by participants with wide-ranging expertise using the cheat sheet.

2 RELATED WORK

We overview prior ARTV research, and provide a brief description of Saeghe et al.'s [23] design dimensions, which was used as a basis for our work.

2.1 Augmented Reality Television

Saeghe et al. [23] systematically reviewed 42 papers that used AR in the context of TV broadcasting. They identified six themes: 1) enhancing a conventional TV viewing experience, 2) production of TV content, 3) alternative TV experiences, 4) connecting remote viewers, 5) live-video augmentation, and 6) photogrammetry.

The most widely addressed theme was *enhancing a conventional TV viewing experience*. Use-cases in this theme included using AR to deliver holographic content for a TV show (e.g., [13]), using AR to deliver virtual TV screens around a TV set (e.g., [3]), using AR to provide context for a TV set akin to a *focus + context* metaphor [4] (e.g., [14]), or using AR to replace the TV set and deliver content that appeared to be present in the living room (e.g., [31]).

Vatavu et al. [27] conceptualised *ARTV experiences for the living room* by expanding Milgram and Kishino's [17] *reality-virtuality continuum* to two dimensions; where TV and the world were each positioned on a separate reality-virtuality continua (see [27], Fig 4). This resulted in nine variations of the ARTV concept for the living room: 1) Physical world/physical TV (i.e., a conventional TV viewing experience), 2) Physical world/physical TV with on-TV augmentation (e.g., [18]), 3) Physical world/physical TV with off-TV augmentation (e.g., [13]), 4) Physical world/virtual TV (e.g., [26]), 5) Augmented world/physical TV (e.g., [12]), 6) Augmented world/physical TV with augmentation, 7) Augmented world/virtual TV, 8) Virtual world/physical TV, 9) Virtual world/virtual TV. A lack of examples from the literature regarding the last few items in the above list highlights untapped areas in the ARTV research field.

[1] https://en.wikipedia.org/wiki/The_Checklist_Manifesto

2.2 ARTV scenarios

Geerts et al. [7] conducted two co-design workshops with nine families to come up with scenarios for a future TV viewing experience. One of the scenarios developed in their paper addressed *'immersion'*, which consisted of an ARTV-like experience, where in addition to a travel-documentary video being displayed on a TV screen, realistic local animals appeared to "walk through the living room" [7].

To elicit potential viewers' expectation of ARTV in a large scale, Popovici and Vatavu [20] conducted a survey asking $n = 172$ European participants to rate the perceived value of twenty sentence-length pre-written ARTV scenarios. The top three high-ranking scenarios were:

1) I would like to be able to control and interact with AR content displayed around or in front of the TV set.
2) Additional content, such as character names or details about the transmission, displayed next to the TV set.
3) A very large field of view using video projections in the entire room.

Later, Popovici et al. [21] conducted a similar survey, this time with $n = 147$ Chinese participants to investigate cultural differences in viewers' expectations of ARTV. They asked participants to rate the same twenty scenarios. The top three high-ranking scenarios were:

1) I would like to be able to control and interact with AR content displayed around or in front of the TV set.
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.
3) Different perspectives of the TV broadcast, such as a movie or show filmed from different angles, displayed next to the TV set.

The findings of these two large scale cross-cultural surveys indicate that while there are cultural differences, the perceived value of ARTV appears to come from content delivery with AR, novel control over and interaction with content afforded by AR, and the ability to place content near relevant objects in the viewing environment. This perceived value in content is a direct contrast to the lower rated items that focus on menus, channels, and subtitles (see [21], Table 7).

2.3 ARTV design space

In this paper, we used the concepts presented by Saeghe et al. [23] as a basis for creating a design space cheat sheet; henceforth referred to as *the ARTV design space cheat sheet*.

Saeghe et al. [23] identified six dimensions of the ARTV design space:

1) Abstraction: describing the semantic relationship between AR and TV content, where either AR and TV are independently *complete* experiences (e.g., [8]), or both AR and TV content are required for an ARTV to make sense (e.g., [22]), or where either AR or TV play an additional role to an otherwise already *complete* experience (e.g., [3]).
2) Interaction: describing audiences' interactions with content, for instance to change a programme, to resize or re-position

content or change the viewing angle (e.g., [26]); or to interact with content in a game-like manner (e.g., [22])
3) Time: describing the relationship between the timelines of AR content and TV content, where AR and TV are intentionally not synchronised (e.g., [8]), where AR and TV are synchronised and presented together (e.g., [13]), or AR and TV are synchronised but presented intermittently (e.g., [22])
4) Display: describing where the visual elements of AR and TV content are presented to the viewers, for instance, on the same display device (e.g., [6]), or on separate devices (e.g., [13])
5) Context: describing the ways in which a viewer's experience is affected due to the presence of other people, and the features of space and objects.
6) Editorial control: describing a viewer's ability to influence the way in which they consume content, for instance, by changing the camera angle (e.g., [9]).

3 METHOD

We conducted a between-subject study with $n = 10$ participants, with wide-ranging expertise, to investigate the usefulness of a set of design space concepts—presented in textual cheat sheet format—in facilitating the conceptualisation of ARTV scenarios. The study was approved by the departmental ethics committee at The University of Manchester (Reference: 2020-10054-16247).

3.1 ARTV design space cheat sheet

We further developed the dimensions identified by Saeghe et al. [23] by separating *content* (i.e., AR content, TV content, and together ARTV content) from *non-content* (i.e., people, space, and objects) into two classes, and subsequently considered the relationships between the components of each class, within the class and across two classes.

We ran a pilot with one participant asking them to write two ARTV scenarios. Before writing the first scenario, we provided them with a link to a YouTube video and a PDF handout. In the video a researcher described the basic definitions used in this study (i.e., AR content, TV content, and ARTV) and asked the participant to write an ARTV scenario. The handout presented the same material in text.

Once we received their first scenario, we provided them with a link to a second YouTube video and the ARTV design space cheat sheet. In the second video a researcher described the ARTV design space under investigation, and asked the participant to write a second scenario. The cheat sheet presented the design space dimensions by asking questions regarding each dimension.

Based on the feedback we received from the pilot, the handout and the ARTV design space cheat sheet were updated by including an operational definition for an ARTV scenario, and a note consisting of two points; one regarding participants' role in the ARTV scenario and another regarding the usage of current vs. hypothetical future display technologies (see Appendices B and F).

Table 1: Participants' role in the media and AR industries. Selection of more than one option was allowed.

Role	Media	AR
Consumer	7	6
Producer	3	4
Writer	0	1
Researcher	4	5
Engineer	3	1
Technologist	2	2
Enthusiast	4	7
None	0	1
Not disclosed	1	0
Self-described	Designer	Designer

Table 2: Devices typically used by participants to watch TV content. Selection of more than one option was allowed.

Device	Total
TV set	8
Laptop	8
Mobile phone	5
Tablet	4
Desktop computer	2
Projection	1
HMD	1

3.2 Participants and experimental conditions

Adult participants were recruited using social media (LinkedIn and Twitter) and electronic mailing lists. Recruitment was incentivised with a £10 Amazon voucher.

Ten individuals (8 male, 2 female) opted to participate (M=34.5, SD=6.8). Four participants were in the $25-29$ age group, three participants were in the $35-39$ age group, and another three were in the $40-44$ age group.

3.2.1 Experiences with TV and Media. Seven participants selected more than one media role. The most frequently selected role was *consumer* (seven participants) followed by *researcher* and *enthusiast* (four participants each). One participant preferred not to disclose their role, and one participant self-described their role as designer (Table 1).

All participants reported watching some TV every day, with nine watching more than one hour per day and three at least two hours. The most popular devices were *TV set* and *laptop*, with eight participants selecting them. Third most popular device was *mobile phone*, with five participants selecting it. Eight participants reported using more than one device for TV consumption (Table 2).

3.2.2 Familiarity with, and use of, AR. Eight participants selected more than one role. The most popular role was enthusiast with seven votes. The second most popular was consumer with six votes. One participant selected *none*, and one participant self-described as designer (Table 1).

Three options were provided in the demographic questionnaire: 1) I don't know what AR is, 2) I could probably describe AR, and 3) I

could describe AR confidently. All participants selected the third option.

All participants had used AR before. Six reported recent regular use/development experience. One reported former regular use/development experience. Three reported infrequent use and a lack of development experience. Four reported usage of under one hour. One participant used AR at least one hour per week, while another participant used AR at least two hours per week.

The most popularly used device for consuming AR was *mobile phone*, with seven participants selecting it. The second most popular device was *HMD (inc. smart glasses)* with six votes. *Tablet* was selected three times, and *projection* was selected once.

3.2.3 Experimental Conditions. The study had two groups: a study group, and a control group. Participants were assigned to a group randomly. Random assignment had no obvious drawbacks since care was taken to ensure participants were equivalent in terms of expertise and experience. Each participant wrote two ARTV scenarios. The control group was included to account for a potential practice effect, when the same participant writes two scenarios. This resulted in four sets of scenarios:

1) Study 1: five scenarios written by the study group participants after exposure to a set of basic definitions (Appendices A and B).
2) Study 2: five scenarios written by the study group participants after exposure to the ARTV design space cheat sheet (Appendices E and F).
3) Control 1: five scenarios written by the control group participants after exposure to a set of basic definitions (Appendices A and B).
4) Control 2: five scenarios written by the control group participants after exposure to the same set of basic definitions as before (Appendices C and D).

3.3 Procedure

Fig 1 present a flow diagram of the experimental procedure.

3.3.1 Scenario Writing Tasks. There were two scenario writing tasks. The first task was identical for all participants; it consisted of watching a YouTube video, and writing an ARTV scenario. Participants were sent a link to a YouTube video, and two PDF documents: 1) the script for the YouTube video, and 2) a handout. In the video, a researcher presented the basic definitions used in the study (i.e., AR content, TV content, ARTV, and ARTV scenario), and asked participants to write an ARTV scenario (see Appendix A for the script of the video). The handout presented the same material in text (Appendix B).

The second task consisted of either writing a new ARTV scenario or embellishing the first ARTV scenario. All participants received a link to a YouTube video and two PDF documents: 1) the script for the YouTube video, and 2) a handout. For the second task, the material provided to the participants was different depending on their assigned group. Participants in the control group did not receive any new information. The handout given to them was identical to the handout they had received for the previous task. In the YouTube video, a researcher asked the participants in the control group to use the same material for the task (Appendix C). Participants in the study group received a link to a different YouTube video, where a researcher described the dimensions of the ARTV design space under investigation. The handout was the ARTV design space cheat sheet (Appendix F). Participants were given 48 hours to complete each of their tasks.

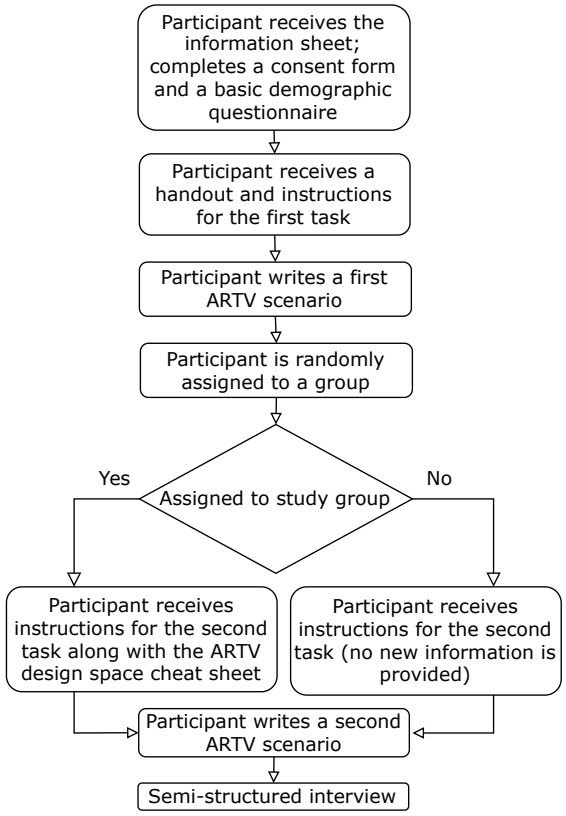

Figure 1: Flow diagram of the experimental procedure.

3.3.2 Semi-structured interview. The third and final task was a semi-structured interview. Participants joined a researcher on a one-to-one Zoom call. We asked each participant about the clarity and usefulness of ARTV concepts, as presented in the instructional videos and associated handouts, with a focus on the ARTV design space cheat sheet. After the interview, participants were debriefed and given the opportunity to ask questions.

4 SCENARIOS

Overall twenty scenarios were written by $n = 10$ participants. While all participants in the control group wrote two unique scenarios, 60% (3 out of 5) of participants in the study group embellished their first scenario (instead of writing a new scenario) for the second task.

4.1 Qualitative analysis

We used a *deductive* content analysis technique [15] to analyse the scenarios. Two researchers independently coded the scenarios and

Table 3: The top ten most used codes and the number of scenarios in which they appeared.

Code	Study 1	Study 2	Control 1	Control 2	All
Content - TV	5	4	3	5	17
AR-TV dependency - additional	5	4	3	3	15
People - single primary viewer	4	2	4	4	14
ARTV genre	2	3	4	4	13
AR and TV content	4	1	4	3	12
Time - sync - continuous	2	4	4	2	12
Visual display - unframed	3	3	3	2	11
People - multiple - concurrent	2	4	2	2	10
Space - indoors	2	3	2	3	10
Content - AR	3	2	2	2	9
People - multiple primary viewers	2	4	2	1	9
Space - private	2	3	2	2	9

resolved disagreements through discussion. The initial code-book consisted of 42 categories extracted from the ARTV design space cheat sheet. The coding process consisted of three iterations; during the first two iterations, a further ten categories were identified and added to the code-book.

First iteration: The first iteration consisted of coding two scenarios from a participant in the study group. The raw inter-rater agreement was 0.72, Cohen's kappa was 0.38. Differences were resolved through discussion. Four new codes were added to the code-book: 1) ARTV genre 2) broadcast mode 3) editorial control 4) audience demographic

Second iteration: The second iteration consisted of coding two scenarios from a participant in the control group. The raw inter-rater agreement was 0.80, Cohen's kappa was 0.44. Differences were once again resolved through discussion. Six new codes were added to the code-book: 1) representation - photoreal 2) representation - cartoonish 3) display - 360° 4) modality - auditory 5) modality - haptic 6) modality - olfactory

After the second iteration, the code-book was fixed with 52 codes. No new codes were added to the code-book after the second iteration.

Third iteration: The third (and final) iteration consisted of coding all the scenarios. The raw agreement was 0.82, Cohen's kappa was 0.55. Once again, differences were resolved through discussion.

Inter-rater agreement: The final level of agreement achieved (Cohen's kappa of 0.55) is typically considered to be *weak* [16]. However, given the shared understanding of the concepts that was developed and articulated during the coding process, we expected the inter-rater agreement to be higher. We hypothesised that the sheer number of codes ($n = 52$) may have caused many instances to be missed by either of the researchers rather than being misunderstood or mislabeled.

To this end, the two coders reviewed their coded passages and compared notes. We found that there were 188 items where the two coders' opinions differed; 129 of these items where codes that were missed by one of the coders (there was no disagreement about the code once a missed instance was pointed out). 59 items were genuine disagreements where the coders had to resolve disagreement by discussion. If the missed codes were to be excluded, the inter-rater agreement of the entire scenario corpus would be a raw agreement of 0.94 and a Cohen's kappa of 0.87, which is considered *strong* [16].

Table 4: Twenty-one high-usage codes appear in all four scenario groups. Codes marked with a *check-mark* were present in Saeghe et al.'s [2020] design space and our original cheat sheet. Others were identified during coding.

Code	Study 1	Study 2	Control 1	Control 2	All
Content - TV (✓)	5	4	3	5	17
AR-TV dependency - additional (✓)	5	4	3	3	15
People - single primary viewer (✓)	4	2	4	4	14
ARTV genre	2	3	4	4	13
AR and TV content (✓)	4	1	4	3	12
Time - sync - continuous (✓)	2	4	4	2	12
Display - unframed (✓)	3	3	3	2	11
People - multiple - concurrent (✓)	2	4	2	2	10
Space - indoors (✓)	2	3	2	3	10
Content - AR (✓)	3	2	2	2	9
People - multiple primary viewers (✓)	2	4	2	1	9
Space - private (✓)	2	3	2	2	9
Broadcast mode	2	2	2	2	8
Social group - between-people interaction (✓)	1	3	1	2	7
People - multiple - at-a-distance (✓)	1	3	1	2	7
People_space - familiar with space (✓)	2	1	1	3	7
Modality - auditory	1	2	2	2	7
Representation - photoreal	1	2	1	2	6
People_objects - familiarity with objects (✓)	2	1	1	1	5
Display - extended frame (✓)	1	2	1	1	5
Social group - friends and family (✓)	1	1	1	1	4

4.2 Codes' usage

Table 3 presents the top ten most widely used codes in the participants' scenarios. The most widely used code was *Content - TV*, which was present in seventeen scenarios from ten unique participants.

From the 52 codes, 21 were used in all experimental conditions, 26 were used in at least one experimental conditions, and five were not used at all; we label these *high*-, *medium*-, and *low*-usage codes, respectively.

4.2.1 High-usage codes: Twenty-one (40%) of the codes appeared in all four experimental conditions (Table 4). Seventeen (81%) originated from the ARTV design space cheat sheet—indicated by a (✓) in Table 4. The remaining four codes emerged during scenarios analysis (Section 4.1).

4.2.2 Medium-usage codes: Twenty-six (50%) of the codes appeared in at least one of the four experimental conditions (Table 5). Twenty of these codes (77%) originated from the ARTV design space cheat sheet—indicated by a (✓) in Table 5. The remaining six codes emerged during scenario analysis (Section 4.1).

4.2.3 Low-usage codes: Five codes did not appear in any of the four experimental conditions (Table 6). Low-usage codes comprise 10% of the total codes.

4.3 Implications

Table 7 presents our initial classification scheme for ARTV design space dimensions, which was used to create the ARTV design space cheat sheet. Furthermore, the structure and wording of Table 7 influenced our code-book during the qualitative analysis of the scenarios.

Based on our analysis of participant scenarios, we propose a revision of Saeghe et al.'s [23] design space that incorporates new dimensions identified during coding, alters the grouping of existing

Table 5: Twenty-six medium-usage codes appear in at least one of the four scenario groups. Codes marked with a *checkmark* were present in Saeghe et al.'s [2020] design space and our original cheat sheet. Others were identified during coding.

Code	Study 1	Study 2	Control 1	Control 2	All
Interaction - display (✓)	3	4	1	0	8
Interaction - game (✓)	0	1	1	4	6
People - multiple - co-located (✓)	2	2	1	0	5
Space - purposeful augmentation (✓)	1	1	0	3	5
People_space - movement (✓)	0	1	0	3	4
Interaction - editorial control	1	2	0	1	4
Display - multiple frames (✓)	1	1	2	0	4
Space - content modification/generation (✓)	0	2	0	2	4
Objects - purposeful augmentation (✓)	2	1	1	0	4
Objects - content modification/generation (✓)	1	2	0	1	4
AR delivering TV content (✓)	0	1	1	1	3
Demographics	1	0	1	1	3
Social group - strangers (✓)	0	3	0	0	3
People - bystanders (✓)	0	2	1	0	3
AR-TV dependency - independent (✓)	1	1	1	0	3
AR-TV dependency - dependent (✓)	0	0	1	2	3
Time - asynchronous (✓)	2	1	0	0	3
Time - sync - intermittent (✓)	1	0	1	1	3
Objects - physical integration (✓)	1	2	0	0	3
People_objects - influencing objects (✓)	1	1	0	0	2
Display - single frame (✓)	0	2	0	0	2
Representation - cartoonish	0	1	0	1	2
Modality - haptic	1	1	0	0	2
Display - 360°	0	0	1	0	1
Modality - olfactory	0	0	0	1	1
Space - outdoors (✓)	0	0	0	1	1

Table 6: Five low-usage codes appear in none of the four scenario groups.

Code	Study 1	Study 2	Control 1	Control 2	All
Space - public	0	0	0	0	0
People - passers-by	0	0	0	0	0
People - multiple - non-concurrent	0	0	0	0	0
People_space - influencing space	0	0	0	0	0
Interaction - story	0	0	0	0	0

Table 7: The classification of ARTV design space dimensions in the cheat sheet.

Content vs. non-content	Content	AR; TV; ARTV
	Non-content	People; space; objects
Relationships	Within content	AR-TV dependency; time; display
	Between content & non-content	Content_people; content_space; content_objects
	Within non-content	People_people; people_space; people_objects

dimensions, and adjusts the terminology used to describe some existing dimensions. The updated scheme consists of four overarching categories: content, people, space, and objects.

The majority of the rewording will be intuitive to the reader. For instance, *People_objects - influencing objects* was shortened to *Influencing objects*. The only relatively major rewording is the changing of *Display* to *Visual display*. This update aims to avoid confusion between various sensory displays (e.g., olfactory and auditory displays), in light of the addition of the *Modality* dimension to the list—consisting of *auditory*, *haptic*, and *olfactory* aspects.

4.3.1 Content: Twenty-two (42%) of the codes describe various aspects of content in the context of an ARTV experience (see Table 8). 50% of content-related codes were used in all four conditions (high-usage), while the other 50% were used in at least one of the four experimental conditions (medium-usage). There were no low-usage dimensions in this category. The top seven most frequently used codes in this category form 58% of the top ten most used codes in the study.

Table 8: Twenty-two content-related codes. The codes above the horizontal dashed line were high-usage and the codes below it were medium usage.

Code	Study 1	Study 2	Control 1	Control 2	All
Content - TV	5	4	3	5	17
Dependency - additional	5	4	3	3	15
ARTV genre	2	3	4	4	13
AR and TV content	4	1	4	3	12
Time - sync - continuous	2	4	4	2	12
Visual display - unframed	3	3	3	2	11
Content - AR	3	2	2	2	9
Broadcast mode	2	2	2	2	8
Modality - auditory	1	2	2	2	7
Representation - photoreal	1	2	1	2	6
Visual display - extended frame	1	2	1	1	5
Visual display - multiple frames	1	1	2	0	4
AR delivering TV content	0	1	1	1	3
Dependency - independent	1	1	1	0	3
Dependency - dependent	0	0	1	2	3
Time - asynchronous	2	1	0	0	3
Time - sync - intermittent	1	0	1	1	3
Representation - cartoonish	0	1	0	1	2
Modality - haptic	1	1	0	0	2
Visual display - single frame	0	2	0	0	2
Modality - olfactory	0	0	0	1	1
Visual display - 360°	0	0	1	0	1

Next we discuss the usage of various content-related dimensions, and bring in examples from the scenarios written by participants.

Genre: Thirteen (65%) of the scenarios were written with a specific genre in mind. Seven genres emerged from our analysis. We provide excerpts from the scenarios for each genre:

1) *Children's show:* For instance, a TV show called "Kid Detective" that centres around a dedicated mobile/tablet app allowing a synchronized play-along experience with a child's favourite cartoon character. "It is the child's job to take up the titular role by watching along, noting down clues to try and solve a problem." [P10S2]
2) *Educational:* For instance, "an interactive teaching programme realising the potential of AR to bring subject topics to life for the students. The teacher could be presented on the TV screen, while the AR content could be spread around the room. " [P7S1]
3) *Game show:* For instance, "[Extending] the famous Japanese game show Takeshi's Castle with Augmented Reality …a wall-sized TV and body recognition system …enables the viewers at home to participate in the games and challenges of the show." [P1S2]
4) *Escape room:* For instance, "A fast-paced time-critical game/ movie/escape room, [where] people have 45 minutes to escape and find out how to kill the serial killer." [P2S2]
5) *Sports:* For instance, an interactive AR application for televised Premiere League matches, where the TV is used to display football content, while "AR …can detect player, game …App allows also to grab the match in the screen and place it somewhere else to watch it with others from different angles." [P4S2]
6) *Dance show:* For example, an ARTV edition of Strictly Come Dancing where "Alice and John could open their ARTV app

on their iPad and point it to the screen. The app will recognise the current professional dancer and celebrity and provide some sort of visual indication to highlight that there's an interactive content about them. They could then tap on this and get information ... Alternatively, Alice and John can interact live with other viewers, [voting] and posting short comments." [P6S2]

7) *Documentary:* For instance, "a documentary about quantum mechanics, say the Wigner's friends thought experiment, ... [where] the orbits and positions of the particles (in a simplest format a bottle, let them be x) observed and the possible state they could be in (0-broken or 1-not broken) presented in 3-D ... All the possible observed states of the objects (x, A+x) would be presented in 3-D in an evolving, chronologically ordered manner." [P3S1]

TV and AR content: TV content was typically referenced using terms associated with current TV-like media, and 2-D presentation. For instance: "short clips" [P1S1], "a film" [P8S1&2], "[an] episode" or a "show" [both P9S1].

AR content was typically presented in 3-D. For instance: "interactive 3-D cross-sectional models of the human body, or of the solar system, or of the different atoms and bonds forming a molecule" [P7S1], or "dancers via AR in 3-D" [P3S2].

Visual display: AR content was typically presented in an *unframed* way (used in 55% of the scenarios). *Extended*, *multiple*, and *single frame* were used less frequently (used in 25%, 20%, and 10% of the scenarios, respectively). Examples of unframed AR content include: "3-D icons ... next to the TV" [P8S1], "objects on [the] table or floor" [P9S1], and "a cat in AR walking around [the] room." [P10S2]

Examples of extended frame include: "the point of view of the player ... extended beyond the screen" [P7S2] and "an explosion is carried out from the screen ... shrapnel flying towards me in AR." [P8S1&2]

Examples of multiple frames include: placement of multiple "AR windows around his living room ... to keep an eye on what's going on in other matches" [P10S1] and responding to an interaction by presenting "more information ... on [the viewer's] phone." [P6S1]

Dependency: Participants typically used either TV content or AR content to present *additional* material (used in 75% of the scenarios). The *independent* and *dependent* aspects were used less frequently (used in 15% of the scenarios, each). Additional AR content was typically used to supplement TV content. For instance: different camera angles displayed as "AR additional content ... around the viewer." [P2S1]

Additional TV-like content was used to supplement an experience that centred around AR content. For instance, "by streaming [a dance seminar in] real-time." [P3S2]

Time: Presenting AR content and TV content in *parallel timelines*—*synchronised continuous*—was used in twelve (60%) scenarios. In contrast, presenting AR content and TV content intermittently or in a non-synchronised way—*synchronous intermittent* and *asynchronous*, respectively—each were used in three (15%) scenarios. Examples of synchronised continuous include: "seamless integration [of TV content] with synchronised 3-D out-of-frame content" [P8S2] and the appearance of "a relevant 3-D object" while "watching a segment on a particular artefact." [P9S2]

An example for synchronised intermittent is: being prompted by "the character on the TV screen ... [when the viewer] needs to switch focus [from] the TV to the tablet device." [P10S2]

An example for asynchronous consists of the user having to manually pause TV content to interact with an AR component of the experience, for instance: "to get more information on some content displayed on TV, [the viewer] pauses the show and then opens [their] ARTV app." [P6S1]

Broadcast mode: Broadcast mode was used in eight (40%) scenarios. It captures the way in which ARTV content is broadcast to the audiences. The two main considerations mentioned in the scenarios were *live* vs. *on-demand*. For instance: "pre-recorded videos [that are] available on-demand" [P1S1] and "a live sport broadcast" [P10S1]

Modality: The *auditory* aspect was used in seven (35%) scenarios, while the *haptic* and *olfactory* aspects were used in two and one, respectively (10% and 5%, respectively).

Examples of the auditory modality include: "3-D audio" for a Formula 1 AR extension [P2S1], spatial audio to enable "full immersion" [P8S1], and enhancing "[a viewer's] sense of being there ... [using] binaural audio." [P10S1]

An example of the haptic modality consisits of: "haptic feedback in [the viewer's] wrist [providing] feedback [when] the screen powers up" and "haptic feedback on [the viewer's] face, as the POV shot goes through vegetation." [both P8S1]

An example of olfactory modality include: "smells designed to [immerse] the players" [P2S2] in an ARTV escape room scenario.

Representation: The two aspects used in the scenarios were *photo-realistic* and *cartoonish*, with 30% and 10% of the scenarios using them, respectively (six and two, respectively). Examples of the photo-realistic aspect typically consist of *volumetrically captured* people and objects. For instance: "Volumetric avatars of the racers, track, and car" [P2S1] in a Formula 1 scenario.

The cartoonish aspect was used in a scenario to overcome a *privacy* concern, when P1 (study group) was prompted—via the ARTV design space cheat sheet—to think about including *strangers* into an ARTV experience: "the players could produce cartoon avatars of themselves and only the player movements are streamed into other households (not the living room or by-sitters)." [P1S1]

4.3.2 People: Twenty-one (40%) codes capture various aspects relating to people (Table 9).

38% of *people*-related codes were used in all four conditions (high-usage), 43% were used in at least one condition (medium-usage), and 19% were used in none of the four experimental conditions (low-usage). The top three most frequently used codes in this category form 25% of the overall top ten most used codes.

Single vs. multiple viewers: A *single* viewer featured in 70% of the scenarios, while *multiple* viewers were present in 45% of the scenarios. More scenarios considered *multiple at-a-distance* viewers than *multiple co-located* viewers, with seven and five, respectively

Table 9: Twenty-one people-related codes. The two horizontal dashed lines separate (top to bottom) high-, medium-, and low-usage codes.

Dimension	Study 1	Study 2	Control 1	Control 2	All
Single primary viewer	4	2	4	4	14
Multiple concurrent viewers	2	4	2	2	10
Multiple primary viewers	2	4	2	1	9
Social group - between people interaction	1	3	1	2	7
Multiple at-a-distance viewers	1	3	1	2	7
Familiarity with space	2	1	1	3	7
Familiarity with objects	2	1	1	1	5
Social group - friends and family	1	1	1	1	4
Interaction - display	3	4	1	0	8
Interaction - game	0	1	1	4	6
Multiple co-located viewers	2	2	1	0	5
Interaction - editorial control	1	2	0	1	4
People moving	0	1	0	3	4
Demographics	1	0	1	1	3
Social group - strangers	0	3	0	0	3
Bystanders	0	2	1	0	3
Influencing objects	1	1	0	0	2
Passers-by	0	0	0	0	0
Multiple non-concurrent viewers	0	0	0	0	0
Interaction - story	0	0	0	0	0
Influencing space	0	0	0	0	0

(35% and 25% of the total scenarios, respectively). No scenario considered *non-concurrent* viewing—an aspect that could arise in scenarios where viewers are at-a-distance or when content is available on-demand.

Interaction: The most widely used type of interaction was *display*-level interaction, with eight (40%) scenarios using it. This type of interaction is typically intended to enable viewers to perform tasks such as: changing the programme (akin to changing the channel in the context of a conventional TV set), accessing extra information, grabbing virtual objects to change their position and enable viewing from other angles, etc. Examples of display-level interaction include: "[AR] app [enabling the viewers] to grab [TV content to] ... place it somewhere else to watch with others from different angles" [P4S2] and the ability to "tap on [an interactive element to] get information about [the TV content]" [P6S2].

Game-level interaction was used in six (30%) scenarios. This type of interaction typically transforms a *passive viewer* into an *active participant* of a show or an ARTV experience. Examples of game-level interaction include: participation of at-a-distance viewers "in the games and challenges of [a game] show"[P1S2] and transforming a class about maths and shapes for primary school students into a game, where "[they] would have to look around and interact with [shapes]." [P7S1]

No *story*-level interaction was used in the scenarios. This type of interaction is intended to capture scenarios in which, through interaction, a viewer can change the narrative of an ARTV experience.

Familiarity of viewers with space and objects: Familiarity with space was used in seven (35%) scenarios. These were typically explicit mentions of a familiar viewing environment for the viewers. For instance: "viewers at home [can interact with AR] on their living room tables." [P1S1]

Familiarity with objects was used in five (25%) scenarios. These were explicit mentions of a familiar object (physical or virtual), for instance: "[the viewer being] used to digital objects [presented] alongside [their] real physical possessions. [P8S1]

Table 10: Five space-related codes. The two horizontal dashed lines separate (top to bottom) high-, medium-, and low-usage codes.

Dimension	Study 1	Study 2	Control 1	Control 2	All
Indoors	2	3	2	3	10
Private	2	3	2	2	9
Purposeful augmentation	1	1	0	3	5
Content modification/generation	0	2	0	2	4
Outdoors	0	0	0	1	1
Public	0	0	0	0	0

Table 11: Three object-related codes appeared in at least one of the four scenario groups.

Dimension	Study 1	Study 2	Control 1	Control 2	All
Purposeful augmentation	2	1	1	0	4
Content modification/generation	1	2	0	1	4
Physical integration	1	2	0	0	3

Bystanders and passers-by: Bystanders were mentioned in three (15%) scenarios. Bystanders are people that are not directly involved in an ARTV experience. For instance, "By-sitters of players, not involved in the game" [P1S2] and family members of a viewer who cannot view AR content "because [the viewer] is the only one wearing the AR headset" [P10S1]

Passers-by were never mentioned. This is not surprising, since viewers are likely to come across passers-by in a *public space* and no scenario considered public space as a viewing environment for their ARTV experience.

4.3.3 Space: Six (12%) codes describe various aspects regarding *space* (Table 10). In the context of an ARTV experience, space typically refers to the *viewing environment* or the site where content is being broadcast from, e.g., a game show set.

33% of space-related codes were used in all four conditions (high-usage), 50% were used in at least one experimental condition (medium-usage), and 17% were never used (low-usage). The top two most frequently used codes in this category form 17% of the top ten most used codes overall in the study.

Indoors and private vs. outdoors and public: Indoors and *private* were both high-usage codes, with ten and nine (50% and 45%) scenarios mentioning them, respectively. In contrast, *outdoors* and *public* were used once and not at all (medium- and low-usage), respectively. This contrast suggests that participants likely conceptualised ARTV as an extension of their typical TV viewing experience, i.e., indoors and in a private space. The avoidance of the public dimension may have also been influenced by the fact that the study was running during a pandemic, where people were actively avoiding unnecessary encounters in public space. The influence of the pandemic on the participants' scenario writing is captured here: "As lockdown has forced schools to close and parents to assume some teaching responsibilities for their children, this ARTV experience would be an interactive teaching programme realising the potential of AR to bring subject topics to life for the students." [P7S1]

4.3.4 Objects: Three (6%) codes describe various aspects regarding *objects* (Table 11). In the context of an ARTV experience, objects consist of both *physical (real)* and *holographic (or otherwise virtual)*

objects that are present in the viewing environment. Objects are, at least initially, not part of the ARTV experience, but may become integrated into it.

All three object-related codes were used in at least one experimental condition (medium-usage). While none of them where used in all four experimental conditions (high-usage). None of these codes appeared in the top ten most frequently used codes.

5 INTERVIEWS

Our analysis in this section focuses on five interviews conducted with the study group participants.

5.1 Qualitative analysis

We used an *inductive* content analysis technique [15] to analyse the interview transcripts. A category was constructed whenever material was found that fitted one of the following descriptions: 1) utility of the concepts, 2) intelligibility of the concepts, and 3) wording, structure, and overall presentation of the material.

The interview transcripts were worked through line by line. An initial reading of the material resulted in 47 labels. Through an iterative approach, to increase clarity and remove overlap between the labels, this number was reduced to nineteen labels pertaining to five overarching categories: 1) cheat sheet and dimensions' utility, 2) definitions, 3) intelligibility, 4) presentation, and 5) new dimensions.

5.2 ARTV design space cheat sheet and dimensions' utility

For consistency purposes, in this section we summarise and present the results in similar fashion as was presented in Section 4.3.

ARTV design space cheat sheet usefulness: Participants found the ARTV design space cheat sheet helpful and "stimulating" [P8]. For instance, P1 reported that "the cheat sheet helped to structure the whole scenario better and to produce a more detailed scenario."

For those participants who chose to write a new scenario for their second task (two out of five study group participants), having the cheat sheet was perceived to be especially useful: "I wanted to do something different and I wanted to force myself into a more complex thing . . . the structure [of the cheat sheet] was well that I could [ask myself] Okay, what have you forgot? Okay, you have those, you have to consider that. It was easier to structure in a new domain." [P1]

The constraints and the final notes in the cheat sheet were also found to be helpful: "'can be no more than a page or a paragraph.' I think that's quite liberating. . . . The final notes are great." [P8]

People: Participants found prompts regarding the number (single vs. multiple) and the role of people (primary viewer, bystander, or passer-by) in the context of an ARTV experience, helpful. For instance, P4 reported that "[having] multiple viewers or just one viewer . . . plays a vital part in creating the content . . . In my second scenario, I added this after reading it in the cheat sheet, it's quite vital, this was quite helpful." P6 reported: "Initially, I was thinking again, just from the perspective of one person watching TV, but then when you mentioned in the cheat sheet, is there anyone else watching, is there anyone passing-by? I said, Okay, so let's think about . . . a couple." P9 reported "being prompted to think about the viewers and other people . . . was really useful."

When considering multiple viewers, the relationship between people and their interactions with each other during an ARTV experience, was reported to be pertinent. For instance: "it's very, very essential to have in mind how people are related to each other" [P1] and "One of the topics you mentioned: . . . interaction with other viewers . . . [prompted me to think] maybe I can address this, it will be a good experience, not only limited for the person to get extra information about something on TV, but also to interact with others . . . something that I hadn't thought about; something that helped me write these scenarios." [P6]

Similarly, a prompt regarding viewers' capabilities to interact with content was found useful. For instance, P6 reported: "[In] the second paragraph of my scenario, . . . people can vote for the contestant . . . When I read that in the cheat sheet I started to think about having some sort of voting or rating mechanism."

Indoors/outdoors and private/public: Although only one scenario considered an outdoor setting, with no scenario considering a public environment (Section 4.3.3), participants reported being prompted about these dimensions helpful. For instance: "Space was helpful, prior to reading the cheat sheet, I was always thinking about an indoors private space" [P6] and "'Is space public or private, indoors or outdoors'; all super relevant. Outdoors would have a bearing upon any kind of image brightness and expected behavioural norms in a space with people wearing headsets or cameras." [P8]

Purposeful augmentations of space and objects, and integration of objects: Participants reported dimensions regarding purposeful augmentation of space and objects useful. For instance: "I think this is a very, very important one and it's something you wouldn't think immediately because the first thing that comes to your mind is Pokemon Go. So you have a Pokemon somewhere floating in the room. But this is where you really have to think of how you interact with objects and room" [P1] and "One reflection was the idea that . . . rather than having, or as well as having, the TV [content] . . . flying out at me, . . . the outline of the TV could be slightly augmented to make it look like the TV itself, and maybe furniture in the house was kind of wobbling and moving around." [P6]

Time: The considerations regarding the AR and TV time-lines was regarded as pertinent, especially in light of the various broadcast models available, and specifically transforming a live broadcast into on demand. For instance: "It's very essential and very complicated to deal with ... What happens if ... [during] a live show ... people [interact with] AR, and you just put it on demand afterwards? ... How do you deal with all the AR interactions? This is ... the main problem of broadcasters at the moment" [P1] and "The time synchronisation is quite significant, especially for live events." [P4]

ARTV Dependency: The main concern here was designing an ARTV experience such that it would be inclusive of viewers who do not have an AR device. For instance: "What happens if someone does not have the additional? What would happen if a by-sitter would not have AR?" [P1] and "If you had a film that was enhanced for AR glasses, with haptics and elements flying out, if you made it completely dependent on the AR, then it limits it." [P8]

Framed vs. unframed: In the context of an ARTV experience in the living room, the unframed concept was associated with enhancing a conventional TV viewing experience. For instance: "Framed and unframed . . . immediately makes me think about stuff coming out of TV, and then it makes me think about the TV itself being [a] dedicated . . . image generation [device] as the focus . . . the AR can then subtly enhance immersion. And, also enhance the path to content. So the framed and unframed basically kind of suggested that, which I felt like a quite an interesting concept." [P8]

5.3 Definitions

ARTV:. Our operational definition of ARTV provided two examples: to use AR for delivering TV content, and to mix AR content with TV content (Appendices B and F). P6 said "before reading the task, my initial idea of ARTV was the mixture of AR and TV content, so when I was reading the cheat sheet, I realised that I can do the other option as well. I hadn't thought about that."

P4 suggested adding "a third [option], for what we had in terms of second screens . . . to add content to the TV experience: visual, audio, or text content."

AR:. A *content-oriented* definition of AR was presented to participants (Appendices B and F). The typical requirements associated with AR technology, namely, 3-D, registered in real space, and interactive [1, 2], were presented as possibilities rather than requirements. The only comment regarding the definition of AR was a suggestion from P9 to replace the term "images" with the term "imagery".

TV:. Similar to AR, our definition of TV was *content-oriented* (Appendices B and F). The only comment in this section was a suggestion from P9 to add a "social element" to our definition.

5.4 Intelligibility

Two concerns were raised regarding intelligibility: 1) a confusion regarding the definitions of the *content class*, and 2) a confusion regarding the meaning of the terminology used regarding *framed* and *unframed* visual content.

1) Content class (missing information): At the top of the cheat sheet we introduced three types of content, namely, AR, TV, and ARTV (Appendix F). Further down, we introduced a dichotomy of content and non-content classes. Given that we had already introduced the elements that constitute content, we only expanded the non-content class. Three participants found this structure to be confusing. For instance: "there is nothing below [the content class] to give you more information." [P6]

2) Framed vs. unframed: Two participants, at least initially, struggled with this terminology. For instance: "I wasn't familiar with the term 'framed'." [P1] P4 suggested introducing the term *object* to further clarify the term *unframed*: "you can place objects in space, which is then unframed." [P4]

5.5 Presentation

The only suggestion made regarding the presentation was offered by P9: "having the final note[s] . . . right at the top would be super useful."

5.6 New dimensions from interviews

Two concepts were discussed during our interviews that were not presented in *the ARTV design space cheat sheet*: 1) Broadcast mode, and 2) Privacy.

Broadcast mode: This dimension emerged during the qualitative analysis of the scenarios (Section 4.3.1) and during the interviews: "Is it on demand? Is it live?" [P1]

Privacy: In Sections 4.3.1 we introduced the cartoonish representation as a new content-related dimension which was used in a scenario as a way to protect the privacy of a remote viewer. The issue of privacy was raised again in the interviews. Since cartoonish vs. photo-realistic representation can apply to all aspects of content (including people's avatars), we consider viewers' privacy as a separate dimension. We suggest that in the context of an ARTV experience, viewers' privacy should be addressed in early stages of design: "if you have collaborative concepts, how do you deal with the privacy of people?. . . people are after work in their living rooms, in a very, very intimate situation and they do not want someone to watch them in their sweatpants . . . with bare feet on the couch." [P1]

6 DISCUSSION

We provide four discussion points regarding typical ARTV conceptualisations, expansion and modification of the ARTV design space, potential trade-offs between a textual checklist style and ideation cards to present ARTV design space concepts, and a hypothesis with regards to study group's tendency to revisit their first scenario.

6.1 Typical ARTV conceptualisations

Section 4.2 provided details of codes' usage in participant scenarios. By reflecting on these findings, we present concise descriptions of typical ARTV conceptualisations. for high-, medium-, and low-usage codes.

High-usage. Popular ARTV conceptualisations in the study typically consisted of scenarios where AR was used to deliver additional photorealistic content in an unframed way, where AR and TV content were delivered simultaneously and AR was used to augment the viewing experience of a single viewer. The viewer was typically in a private indoors space and did not tend to interact with content; rather, the viewer tended to use AR's capabilities to interact with remote viewers (typically friends and family).

Medium-usage. Less popular ARTV conceptualisations typically consisted of scenarios where multiple co-located viewers (who did not necessarily know each other) interacted with content; AR was used to deliver content in multiple frames or to replace the TV display; AR content was not necessarily in sync with, or presented simultaneously as TV content.

Low-usage. There were no ARTV conceptualisations consisting of an ARTV experience in a public space. This may suggest that the participants did not find ARTV to be appropriate in a public space with presence of potential passers-by, which may in part be a result of prior TV viewing experiences (solo/small group, typically indoor). A lack of usage regarding multiple non-concurrent viewers, viewers' ability to influence the viewing environment, and their

ability to influence the story path may be caused by potential complexities associated with conceiving scenarios—and subsequently, technologies—that implement them.

6.2 Expansion/modification of the design space

In light of our findings, we updated the ARTV design space and associated cheat sheet (Appendix G). Furthermore, we observe that an exercise of having practitioners writing scenarios, and the subsequent analysis of such scenarios, can be an effective way to identify further design dimensions. This approach can be used to identify existing overlaps in design between ARTV and other closely related media. For instance, 360° display—an aspect that was identified in participants' scenarios, while missing from ARTV literature—has an active community of researchers and practitioners (e.g., see [11]). Using the expertise and the findings from such closely related fields can enrich ARTV experiences.

Five new dimensions (and their aspects) that were identified in our study expand the ARTV design space in novel ways. These are: broadcast mode (live, on-demand), genre, audience demographic, representation (cartoonish, photoreal), and privacy. Broadcasters' awareness and considerations regarding these dimensions are important and necessary for ARTV to become a successful medium. For instance, the questions surrounding how to create ARTV content for a live broadcast, and how to convert this for on-demand viewers poses challenges in the production and distribution phases. Considerations regarding genre, audience demographics, and the ways in which AR content is represented (photoreal vs. cartoonish) will likely play a key role in creating the right content for the right audiences. For instance, remote participants of a game show may want to be represented in the show using cartoonish avatars, to protect their privacy. Furthermore, children's ARTV preferences in the context of a children's show may be quite different to those of adults.

The identification of three novel types of stimuli (auditory, haptic, olfactory) suggests that the current focus of the ARTV community—which is primarily on the visual stimuli—may need to be expanded. While it is likely for the visual aspect to dominate, considerations regarding other modalities are likely to gain relevance as and when consumer grade devices that target those modalities become available. For instance, Facebook Reality Labs' work on wrist-based haptics can open the door to a new type ARTV experience.

Saeghe et al. [23] considered editorial control to be separate from the interaction dimension. However, in light of its usage in the scenarios, we found that editorial control can be better described and understood as an aspect of the interaction dimension. Two quotes from participants' scenarios exemplify a common conceptualisation of editorial control as a type of interaction that bestows editorial control upon ARTV viewers. For instance, in the context of a remote dance seminar P4 writes: "There would be the option to choose either the lady for copying the steps and moves right next to her, or the leader for dancing with him after I have practised the lady steps." In the context of using AR to augment a conventional TV viewing experience, P18 writes: "The audience will also be able to control the (now) non-linear flow of the programme, skipping over objects they find less interesting or interacting through a visual/interactive timeline or map that could be displayed on a table [or a] wall."

6.3 Cheat sheet as a checklist

Except for the two points mentioned in Section 5.4, overall the contents of the cheat sheet were reported to be easily understood (RQ1). Similarly, as described in 5.2, participants found the cheat sheet to be useful for conceptualising novel ARTV scenarios (RQ2). Particularly, in cases where participants wrote a scenario from scratch, having access to relevant design dimensions in a structured way was reported to be useful. This way of presenting the design space can help content creators to: 1) ensure no pertinent aspect of an ARTV experience is left out inadvertently, and 2) foster creativity, stimulate thought, and generate novel ideas by enabling content creators to consider a potentially wider range of possibilities.

There are other ways for presenting design space concepts. For instance, a set of ARTV ideation cards (e.g., akin to Wetzel et al.'s [30] mixed reality game cards) could be developed to gamify the design process for ARTV experiences. However, while ideation cards are great brainstorming tools and conversation starters [10], especially when more than one designer is involved, going through all possible options in a check list fashion can be a useful way of ensuring no pertinent aspect is left out. A comparative study can highlight the trade-offs between these two methods for presenting ARTV design space concepts.

6.4 Study group's tendency to revisit their first scenario

While all control group participants wrote two unique scenarios, 3 out of 5 study group participants revisited their first scenario and embellished it for their second task. While the relatively small sample size of the study prohibits us from making any concrete conclusions as to why this was the case, one hypothesis is that once study group participants were exposed to the design space concepts—after their first task via the cheat sheet—they used their first scenario to compare their prior conceptualisation of ARTV with that which was presented by the cheat sheet. However, since the control group participants did not see the cheat sheet, their conceptualisation of ARTV was not challenged during the study and they instead focused on writing a new scenario.

7 CONCLUSION

Building on previous work, we refined an ARTV design space. We used a cheat sheet to present design dimensions as questions, and subsequently evaluated the utility of the cheat sheet and its usefulness in conceptualising novel ARTV scenarios. Our findings indicate that the cheat sheet is useful and can support practitioners in conceptualising novel ARTV experiences. Furthermore, we expanded the design space by identifying six new dimensions, repositioned a previously identified dimension—editorial control—as an aspect of the interaction dimension, and added 360° as an aspect to the display dimension. In light of our observations and the new findings, we provided an updated ARTV design space cheat sheet. We further observed that asking practitioners to write scenarios, and the subsequent analysis of these scenarios can be a useful method for harvesting novel design dimensions. Through an iterative approach, this method can supplement the design dimensions extracted from the literature to update and further map the underlying ARTV design space.

ACKNOWLEDGMENTS

Research reported in this paper was funded by the UK EPSRC under grant number EP/R512394/1.

REFERENCES

[1] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. 2001. Recent advances in augmented reality. *IEEE computer graphics and applications* 21, 6 (2001), 34–47.

[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* (Gothenburg, Sweden) *(VRST '17)*. ACM, New York, NY, USA, Article 67, 2 pages. https://doi.org/10.1145/3139131.3141196

[4] Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. 2009. A Review of Overview+detail, Zooming, and Focus+context Interfaces. *ACM Comput. Surv.* 41, 1, Article 2 (jan 2009), 31 pages. https://doi.org/10.1145/1456650.1456652

[5] Federico Debandi, Roberto Iacoviello, Alberto Messina, Maurizio Montagnuolo, Federico Manuri, Andrea Sanna, and Davide Zappia. 2018. Enhancing cultural tourism by a mixed reality application for outdoor navigation and information browsing using immersive devices. In *IOP Conference Series: Materials Science and Engineering*, Vol. 364. IOP Publishing, 012048.

[6] Alberto Denia, Jose Ribelles, Angeles Lopez, and Óscar Belmonte. 2011. Low Cost Virtual Animation Effects for Sports Broadcasting: Mosaics, Flags and Big-Sized Flags. In *2011 International Conference on Computational Science and Its Applications*. IEEE, 10–16.

[7] 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* (Salford (Manchester), United Kingdom) *(TVX '19)*. Association for Computing Machinery, New York, NY, USA, 59–69. https://doi.org/10.1145/3317697.3323356

[8] Sabiha Ghellal and Irma Lindt. 2008. Interactive Movie Elements in a Pervasive Game. *Personal Ubiquitous Comput.* 12, 4 (Feb. 2008), 307–315. https://doi.org/10.1007/s00779-007-0186-8

[9] Jungong Han, Dirk Farin, and Peter de With. 2010. A mixed-reality system for broadcasting sports video to mobile devices. *IEEE MultiMedia* 18, 2 (2010), 72–84.

[10] Eva Hornecker. 2010. Creative idea exploration within the structure of a guiding framework: the card brainstorming game. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. 101–108.

[11] Chris J Hughes and Mario Montagud. 2021. Accessibility in 360° video players. *Multimedia tools and applications* 80, 20 (2021), 30993–31020.

[12] 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* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 869–878. https://doi.org/10.1145/2470654.2466112

[13] Hiroyuki Kawakita and Toshio Nakagawa. 2014. Augmented TV: An augmented reality system for TV programs beyond the TV screen. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 955–960.

[14] 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* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 427, 10 pages. https://doi.org/10.1145/3173574.3174001

[15] Philipp Mayring. 2014. Qualitative content analysis: theoretical foundation, basic procedures and software solution. (2014).

[16] Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica* 22, 3 (2012), 276–282.

[17] 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

[18] Lyndon JB Nixon, Lotte Belice Baltussen, and Johan Oomen. 2015. LinkedCulture: Browsing Related Europeana Objects while Watching a Cultural Heritage TV Programme.. In *PATCH@ IUI*.

[19] Veronica Pialorsi and Pejman Saeghe. 2020. Designing Intergenerational Media Experiences: A Participatory Design Approach. In *Toys and the TV: Serious Play. Workshop at the ACM International Conference on Interactive Media Experiences*.

[20] Irina Popovici and Radu-Daniel Vatavu. 2019. Understanding users' preferences for augmented reality television. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 269–278.

[21] 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* (Virtual Event, USA) *(IMX '21)*. Association for Computing Machinery, New York, NY, USA, 50–60. https://doi.org/10.1145/3452918.3458801

[22] Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. 2015. Electric Agents: Combining Collaborative Mobile Augmented Reality and Web-Based Video to Reinvent Interactive Television. *Comput. Entertain.* 12, 3, Article 1 (Feb. 2015), 21 pages. https://doi.org/10.1145/2702109.2633413

[23] Pejman Saeghe, Gavin Abercrombie, Bruce Weir, Sarah Clinch, Stephen Pettifer, and Robert Stevens. 2020. Augmented Reality and Television: Dimensions and Themes. In *ACM International Conference on Interactive Media Experiences* (Cornella, Barcelona, Spain) *(IMX '20)*. Association for Computing Machinery, New York, NY, USA, 13–23. https://doi.org/10.1145/3391614.3393649

[24] 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* (Salford (Manchester), United Kingdom) *(TVX '19)*. ACM, New York, NY, USA, 255–261. https://doi.org/10.1145/3317697.3325129

[25] Pejman Saeghe, Sarah Clinch, Bruce Weir, Steve Pettifer, and Robert Stevens. 2019. Augmented Reality Broadcasting: Challenges and Opportunities. In *Challenges Using Head-Mounted Displays in Shared and Social Spaces. Workshop at The ACM CHI Conference on Human Factors in Computing Systems (CHI 2019)*.

[26] Radu-Daniel Vatavu. 2012. Point & click mediated interactions for large home entertainment displays. *Multimedia Tools and Applications* 59, 1 (2012), 113–128.

[27] Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. 2020. Conceptualizing Augmented Reality Television for the Living Room. In *ACM International Conference on Interactive Media Experiences* (Cornella, Barcelona, Spain) *(IMX '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3391614.3393660

[28] 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*. 1–12.

[29] Simon von der Au, Leon Giering, Christina Schneegass, and Markus Ludwig. 2020. The SpaceStation App: Design and Evaluation of an AR Application for Educational Television. In *ACM International Conference on Interactive Media Experiences*. 24–33.

[30] Richard Wetzel, Tom Rodden, and Steve Benford. 2017. Developing ideation cards for mixed reality game design. *Transactions of the Digital Games Research Association* 3, 2 (2017).

[31] Christian Zimmer, Nanette Ratz, Michael Bertram, and Christian Geiger. 2018. War Children: Using AR in a Documentary Context. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 390–394.

A FIRST TASK VIDEO SCRIPT

Hi,

Thanks for agreeing to take part in this study, where we aim to explore the subject of content creation for Augmented Reality Television, otherwise known as ARTV. In this video I will first try to describe what we mean by terms such as television, augmented reality, and ARTV. And then describe the first task of this study.

Both AR and TV are somewhat challenging concepts to define precisely. For our purposes, however, we can consider them as two ways of visual storytelling. TV is typically considered a predominantly passive medium that delivers content on a 2D screen. AR is typically considered to be an interactive medium that presents 3D visual content that appear to be aligned to real objects in the viewers' field of view. We can think of ARTV in at least two ways: 1) using AR to deliver TV content, and 2) mixing AR with TV. When AR is used to deliver TV content, the AR specific qualities—that are interaction and alignment of 3D content to real objects—can be applied to TV content directly. For instance, to bring action into the viewers environment and eliminate the conventional TV screen completely. However, AR can also be mixed with TV in such a way that content is distributed over both AR and TV. For instance, a TV character can come out of the TV screen, or a few relevant artefacts from a TV show can be made available for further exploration using AR.

The possibilities are in no way limited to these and part of our motivation for this study is to explore these and other options for visual storytelling when both TV and AR are available as storytelling tools. Scenario writing can be a useful way to explore the possibilities. As it allows a description of possible events, or a story to be conveyed using writing.

In this task we ask you to write a scenario for an ARTV experience describing the content and its use of AR and TV, together with how it should be presented to its viewers. You can think of yourself or other people as viewers of the content. The choice of topic and genre is up to you. There is some extra information on the handout that I sent to you which may also be useful. Please write the scenario and email it to me within the next 48 hours. Anything between a paragraphs up to a page would be great. Please use PDF or Word Doc for the final submission.
Thank you.

B FIRST TASK HANDOUT

Types Of Visual Content:
AR: Dynamic or static images, real or computer generated, 2D or 3D. Can be aligned to the physical space, the objects and people within the space, or aligned to TV content or other AR content. Can be interactive.
TV: Typically dynamic images (film) created from camera capture, drawings or computer models, viewed primarily in a passive mode.
ARTV:

- Using AR to present TV content.
- A mixture of AR content and TV content.

ARTV scenario: A description of a possible event(s), or a description of the story told using ARTV.

Note: When writing your scenario you can imagine the role of the viewer, the content creator, or any other role that helps with your creative process.
Don't limit yourself to existing devices and technologies that are used today to display content.
Task: Watch the first video and write an ARTV scenario.
Length: between a paragraph and a full page.
Format: PDF or Word Document
Time: 48 hours from when the time you receive the email (containing this document and the link to the video)

C SECOND TASK VIDEO SCRIPT - CONTROL GROUP

Hi,

Thanks for sending me your first scenario. The second task is to write another ARTV scenario. I have sent you a PDF file summarising the concepts that you need to consider when writing the second scenario. This contains the same background information and instructions as the first task. Please write the scenario and email it to me within the next 48 hours. Anything between a paragraphs up to a page would be great. Please use PDF or Word Doc for the final submission.
Thanks

D SECOND TASK HANDOUT - CONTROL GROUP

For their second task, control group participants received the same handout as they did for their first task (see Appendix B).

E SECOND TASK VIDEO SCRIPT - STUDY GROUP

Hi,

Thanks for sending me your first scenario. In this video I will first attempt to describe a few concepts of an ARTV design space, and then describe the second task of this study. The following concepts are meant to demonstrate some of the design parameters that can be considered when conceiving content for ARTV. Please use the cheat sheet provided for further information.

When we think of an ARTV viewing experience, we can think of two classes of relevant concepts: 1) the content class; that consists of various elements of visual content, and 2) the non-content class; which consists of people, space and objects. We can ask a series of questions to demonstrate various parameters of this design space. For instance, you might ask yourself whether AR is replacing TV completely or is it being used in addition to a conventional TV display?
Is content "framed", as it is with conventional TV displays, or is it presented "unframed" as is typically the case with AR? Or is it a mixture of both?
Are you considering a public, private, an indoors or an outdoors space?
Are you considering one or multiple viewers?
If the setting is outdoors and public, it may be useful to consider whether there may be bystanders or passers-by, in addition to the primary viewers.

If the content can be divided into two or more distinct visual components, for instance a TV component and an AR component, then we can consider the following questions: How are you distributing the story elements over AR and TV?
Are some of the components only adding something "extra" to the main content, or is every component essential for the whole to be meaningful?
Is every component presented simultaneously, or for instance is the TV paused while the viewer interacts with AR artefact?
Does the augmentations have any content-wise relations to space?
Does the features of space have any effect on generation or modification of content?
Are the objects in the environment being augmented?
Does the features of these objects, or their mere presence, have any effect on generation or modification of content?
Can they be integrated into the experience?
If people can interact with content, then what is the effect of their interaction?
Is it simple manipulation of content, is it changing the story path, or is it scoring points for instance similar to a gamified experience?
If there is more than one viewer, are they co-located?
Are they viewing concurrently?
Do they know each other?
Are they supposed to interact with each other?
Can people move around in space?
Are they familiar with the objects in the environment, can they manipulate these objects?

I have sent you a PDF file which summarises these concepts and provides additional information. The second task is to write another ARTV scenario using these concepts. Please write the scenario and email it to me within the next 48 hours. Anything between a paragraphs up to a page would be great. Please use PDF or Word Doc for the final submission.
Thanks

F SECOND TASK HANDOUT - STUDY GROUP (ARTV DESIGN SPACE CHEAT SHEET)

Types Of Visual Content:
AR: Dynamic or static images, real or computer generated, 2D or 3D. Can be aligned to the physical space, the objects and people within the space, or aligned to TV content or other AR content. Can be interactive.
TV: Typically dynamic images (film) created from camera capture, drawings or computer models, viewed primarily in a passive mode.
ARTV:

- Using AR to present TV content.
- A mixture of AR content and TV content.

Framed versus unframed: Traditionally TV content is presented in a framed format (typically but not necessarily rectangular) as a 'window' to the story-world. AR content can be present in a similar way. Additionally, AR technology enables presentation of content such that it appears to be unframed and "present" in the viewers environment. A mixture of both framed and unframed content presentation is possible.
Two Classes:

1) Content class
2) Non-content class
   a) People
      i) Is the experience designed with one viewer in mind, or will there be multiple viewers?
      ii) If there are multiple potential viewers, are all viewers considered to be primary viewers or will there be bystanders accompanying the primary viewers? Will there be any passers-by?
   b) Space
      i) Is it public or a private space?
      ii) Is it indoors or outdoors?
   c) Objects

Types of Relationships:

1) Within content:
   (between AR content and TV content)
   a) AR – TV dependency: What is the effect of removing either the AR or the TV component? Will the remainder still work as a meaningful piece of content?
   b) Time: What is the relationship between AR and TV contents regarding their timelines?
      i) Are they synchronized?
      ii) If they are synchronized, are both AR and TV elements presented simultaneously, or is one element paused while the other element is being presented?
   c) Display: How is content presented?
      i) In a single or multiple frames.
      ii) An extended frame, for instance a TV screen visually extended with content on the surrounding walls.
      iii) Unframed.
      iv) Mixture of framed and unframed.
2) Between ARTV content and non-content:
   a) Between content and space:
      i) Does content augment the space in a meaningful way?
      ii) Is content being modified or generated based on the features of space?
   b) Between content and people: Can viewers interact with content? What is the effect of their interaction?
      i) Selection and manipulation of content.
      ii) Change of story path.
      iii) Accomplishment of tasks in a game-like fashion.
   c) Between content and objects:
      i) Does content augment objects?
      ii) Is content modified or generated based on features of objects in the environment?
      iii) Can objects in the environment be integrated into the ARTV experience?
3) Within non-content:
   a) Between people and people:
      If there are multiple viewers:
      i) Are the viewers co-located?
      ii) Are they viewing ARTV concurrently?
      iii) Do the viewers know each other prior to the viewing experience?
      iv) Are they meant to talk or interact with each other during the viewing?
   b) Between people and space:

i) How familiar are viewers with the space?
ii) Can they manipulate space?
iii) Can they move around in the space?

c) Between people and objects:
i) How familiar are the viewers with the objects in the environment?
ii) Can they influence or manipulate these objects?

Final notes:

- When writing your scenario you can imagine the role of the viewer, the content creator, or any other role that helps with your creative process.
- Don't limit yourself to existing devices and technologies that are used today to display content.

G UPDATED ARTV DESIGN SPACE CHEAT SHEET

Notes:

- When writing your scenario you can imagine yourself in the role of the viewer, the content creator, or any other role that helps with your creative process.
- Don't limit yourself to existing devices and technologies that are used today to display content.

Basic definitions:
AR: Dynamic or static imagery, real or computer generated, 2D or 3D. Can be aligned to the physical space, the objects and people within the space, or aligned to TV content or other AR content. Can be interactive.
TV: Typically dynamic imagery (film) created from camera capture, drawings or computer models, viewed primarily in a passive mode. TV content can be viewed in a social setting.
ARTV:

- Using AR to present TV content.
- Using AR to deliver additional content.
- A mixture of AR content and TV content.

ARTV scenario: A description of a possible event(s), or a description of the story told using ARTV.
Framed versus unframed: Traditionally TV content is presented in a framed format (typically but not necessarily rectangular) as a 'window' to the story-world. AR content can be present in a similar way. Additionally, AR technology enables presentation of content such that it appears to be unframed and "present" in the viewers environment. A mixture of both framed and unframed content presentation is possible.

The rest of the cheat sheet contains considerations regarding four overarching categories: content, people, space, and objects.
Content:
What is the genre of the ARTV experience?
What modalities will the experience target? For instance, auditory, haptic, olfactory.
How will the content be broadcast? For instance, live or on-demand.
Is there any TV content?
Is there any AR content?
If both AR content and TV content are present, what is the effect of removing either the AR or the TV component? Will the remainder still work as a meaningful piece of content?
Is AR content synchronized with TV content? If so, are both AR and TV components presented simultaneously, or is one paused while the other is being presented?
How will the content be visually presented?

- In a single or multiple frames.
- An extended frame, for instance a TV screen visually extended with content on the surrounding walls.
- Unframed, for instance to put holographic artefacts in the viewing environment.
- 360 degree video.

How will the objects and people be represented? For instance, photo-realistically or as a cartoon.
People:
What is the demographic of the target audience?
How many viewers are watching (or participating in) the ARTV experience?
In case of more than one viewer:

- Are they all primary viewers, or are there any bystanders and passers-by?
- Are they able to interacting with each other during the ARTV experience?
- Do they know each other? For instance, friends and family members.
- Are there any strangers involved in the experience?

In case of more than one primary viewer, are they watching (or participating in) the ARTV experience concurrently? Are the co-located or at-a-distance?
Can viewers interact with content? What is the effect of their interaction?

- Selection and manipulation of content.
- Change of story path.
- Accomplishment of tasks in a game-like fashion.
- Control of what they watch and how they watch it, almost as if they had control over how content was pieced together.

Are they watching content in a familiar space? For instance, their living room.
Are they able to move while watching content?
Are they familiar with objects present in the viewing environment? For instance, the furniture in the living room.
Are they able to influence the viewing environment and the objects? For instance, picking up a physical object present in the viewing environment and moving it to another location.
Have you considered the privacy of the viewers? For instance, when connecting viewers at-a- distance that are strangers.
Space:
Is the viewing environment indoors, outdoors, private, or public?
Does content augment the space in a meaningful way?
Is content being modified or generated based on the features of space?
Objects:
Does content augment objects?
Is content modified or generated based on features of objects in the environment?
Can objects in the environment be integrated into the ARTV experience?