CONCEPTUALISATION AND REALISATION OF A DESIGN SPACE FOR AUGMENTED REALITY TELEVISION

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF SCIENCE AND ENGINEERING

2022

Pejman Saeghe

Department of Computer Science

Contents

Word Count: 36,797

List of Tables

List of Figures

10

Abstract

Over the years, a set of story-telling conventions has been established for Television (TV), between content creators and their audiences. These conventions typically involve the creation, and subsequent consumption, of 2-dimensional (2-D) imagery accompanied by sound, requiring limited, if any, interaction between the viewers and content. As consumer Augmented Reality (AR) devices become a commodity, researchers and broadcasters are investigating the ways in which AR can be used to enhance/transform a conventional TV viewing experience—a concept referred to as Augmented Reality Television (ARTV). However, the addition of AR to TV has the potential to disrupt the conventional TV viewing assumptions, since AR's key characteristics are arguably its ability to deliver interactive 3-dimensional (3-D) imagery that is registered to the physical space; a type of content that TV viewers are likely not used to consuming. This raises the question of *how to create ARTV content?*

This thesis conceptualises ARTV as an emerging mass medium, where content creators use both AR and TV to tell their stories to their audiences. We start by providing a few preliminary working definitions (e.g., TV, AR, and ARTV), then systematically review prior research and identify six major themes and six prominent design dimensions (namely, abstraction, interaction, time, display, context, and editorial control). These design dimensions are then used as a basis to create an ARTV design space, from which we derive a cheat sheet to support practitioners. We subsequently evaluate the intelligibility of this cheat sheet, and its impact on conceptualising novel ARTV experiences, in a user study with $n = 10$ participants. Our findings indicate that the cheat sheet can be useful in conceptualising novel ARTV scenarios, and prompt a series of refinements to the design space and associated cheat sheet. Given the dominance of ARTV to enhance conventional TV viewing experiences that take place in the living room, we further develop this aspect with its own conceptual framework and classification. To demonstrate the applicability of our design space and conceptual framework, we explore two specific types of living room based ARTV experiences in

two user studies. In the first, we add holograms to an existing nature documentary and invite $n = 12$ participants to watch a six-minute ARTV content—where TV content is presented on a physical TV display, and AR content is displayed on a Microsoft Hololens head-mounted display—in a user testing facility that is designed to look and feel like a conventional living room. Our findings suggest that the viewers are likely to find such an experience engaging. However, some may feel a degree of anxiety and a fear of missing parts of the content and underlying narrative. In the second study ($n = 10$), we use points on the display dimension of our design space to vary presentation of a lifelike programme-related hologram as part of a nature documentary show. We explore six points on the display dimension in the context of an ARTV experience where both the TV and the living room are virtual. We investigate the impact of the hologram's movement behaviour on the viewers. Our findings highlight the role of personal preferences and the perceived role of the hologram in the underlying narrative and TV content.

Overall, the thesis highlights a need to carefully consider the impact of additional augmentations on the viewers. Furthermore, viewers should be given the ability to customise their ARTV experiences, based on personal preferences.

Our contributions are the following:

1) a systematic literature review, highlighting six common themes of research in ARTV and six commonly used design dimensions;

2) creation and evaluation of a design space for ARTV and its subsequent refinement based on the evidence gathered from a user study;

3) a conceptual framework for ARTV experiences in the living room, identifying nine types of ARTV and highlighting areas for future research;

4) two user-studies investigating two types of ARTV, highlighting a need to accommodate personal preferences for ARTV content.

Declaration

No portion of the work referred to in this thesis has been
submitted in support of an application for another degree or
qualification of this or any other university or other institute
of learning.

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see `http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487`), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see `http://www.manchester.ac.uk/library/aboutus/regulations`) and in The University's policy on presentation of Theses

Acknowledgements

I want to thank my supervisors Dr Sarah Clinch, Prof Robert Stevens, and Dr Bruce Weir for their guidance and support throughout the journey.

Chapter 1

Introduction

Traditionally a television (TV) set created a focal point in a household, around which family members gathered to share their time by watching TV content. Augmented Reality (AR) has the capability to deliver visual content that appears to be registered in 3-dimensional (3-D) space. Combining this capability of AR with a traditional TV viewing experience can result in scenarios where, for instance, a TV character appears to come out, or jump out, of a TV screen and move around the viewing environment (e.g., the living room) [Vatavu et al., 2020].

Over the years, custom and practice has established a set of rules and conventions that enable the creation of well-crafted TV content. However, there is a lack of know-how as regards creating TV content that takes advantage of AR's capabilities in broadcasting. As consumer AR devices become readily commercially available, researchers and broadcasters are investigating the ways in which AR can be used to enhance/transform a traditional TV viewing experience with AR [Saeghe et al., 2020].

AR's affordances can be applied to production, distribution, and consumption phases of TV broadcasting [Saeghe et al., 2020]. However, while the challenges regarding production and distribution of content remain largely technical, the consumption of TV content, in the presence of AR technology, introduces challenges that go beyond the technical and relate to content viewing behaviours, habits, and expectations [Saeghe et al., 2019a].

In recent years, with the advent of streaming services, and the ubiquity of hand-held devices, novel ways of consuming TV content have emerged [Bentley et al., 2019]. However, these devices and services have, for the most part, adhered to the original conventions of TV content, namely, presenting audio-visual content, the visual aspect of which is presented via a 2-D display. Using AR in the context of a TV viewing

experience presents new challenges for content creators and broadcasters, since not all the existing TV conventions will directly transfer to a TV + AR experience.

1.1 Preliminaries

This section provides preliminary insights into three concepts: TV, AR, and Augmented Reality Television (ARTV). While all three concepts are revisited in other chapters (specifically, Chapters 2 and 4), ARTV in particular—as the focus of this thesis—is covered later in the thesis in more detail via a systematic literature review (Chapter 2) and classification of prior work according to various existing frameworks (Chapter 4).

1.1.1 Television

It is difficult to pinpoint a definition for Television that satisfies everyone and without needing to update it frequently; this can be attributed to two factors: 1) the ever-evolving nature of the underlying technology used to create, distribute, and consume TV content, and 2) the ever-evolving consumer habits, partly affected by the advances in TV-related technology and partly shaped by the broader spectrum of changes in the ways we live.

Broadly speaking though, TV can refer to an institution, a platform, the underlying technology, a piece of content, or a device that is used to receive and display content. At its heart, TV remains a mass medium that broadcasts content from a wide array of genres (e.g., news, sports, drama, etc.) to a wide range of audiences. Currently, this content consists of audio-visual components, but other modalities (e.g., haptics and olfactory) can be integrated into TV content, when problems related to their input, output, and distribution are solved.

When considering TV content, one question that may be raised is *how is TV different from film and cinema?* Here, analogies may be useful to clarify; TV is to film what radio is to a song, and TV is to cinema what radio is to a concert hall. However, since the emergence of streaming services—both for music (e.g., Spotify) and for film (e.g., Netflix)—the distinctions have been even more blurred; both TV dramas and cinematic dramas can be streamed over the Internet and viewed on a laptop, a mobile device, or indeed on a TV set.

As a successor to radio, TV started out by adding video to audio. The next stage in the evolution of TV was the addition of interactive services (e.g., teletext, electronic programme guides, and the BBC's Red Button) which aimed to increase audience

engagement [Vatavu et al., 2020]. Recent trends in TV-related research aim to per-
sonalise content, for instance, using companion apps (e.g., Geerts et al. [2014]), or to
display content outside a conventional TV screen, using emerging technologies such as
Augmented Reality (e.g., Saeghe et al. [2019a]).

As novel technologies, such as AR, become more widely available, researchers and
broadcasters are investigating the ways in which they can incorporate the affordances of
AR into various aspects of TV. In particular, they are interested in taking advantage of
AR's interactive and personalising capabilities to reach a wider audience and to engage
with TV audiences more effectively.

1.1.2 Augmented Reality

One of the oldest and most widely cited definitions of AR originates from Milgram and
Kishino [1994]'s "reality-virtuality continuum" (Figure 1.1). In this definition AR is
considered a specific case of *mixed reality*, where computer generated graphics are used
to augment a view of the real world.

As shown in Figure 1.1, on one end of the "reality-virtuality continuum" resides
reality (i.e., a real/physical environment), while on the other end of it resides *virtual
reality* (i.e., a fully virtual environment); everything in between is considered mixed
reality. In turn, mixed reality is split into "augmented reality"—if reality dominates,
and virtual components are used to augment it—and "augmented virtuality"—if virtual
reality dominates and real objects are used to augment the virtual world.

Figure 1.1: Reality-Virtuality continuum based on Milgram and Kishino [1994].
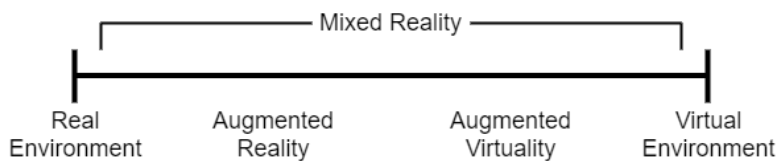
Another widely cited definition of AR is based on Azuma's [1997] survey; proposing
the following three key characteristics for AR systems: 1) combining real and virtual,
2) interactive in real-time, 3) registered in 3-D space.

However, in a series of interviews with experts, Speicher et al. [2019] found that
there was no unique working definition of AR "in the wild"; while some experts
considered simple overlays of virtual artefacts on a view of the real world to constitute
AR, other experts considered 3-D registration and interaction as requirements.

One of the limitations of Milgram and Kishino's [1994] and Azuma's [1997] definitions is their focus on the visual domain; as researchers from fields other than computer graphics start using the term "augmented reality" to refer to concepts and prototypes that augment reality via non-visual modalities (e.g., sounds, haptics, olfactory), a more comprehensive definition for AR is needed.

Another limitation of the definitions provided by Milgram and Kishino [1994] and Azuma [1997] is that they are device-oriented. This limitation originates from the fact that both Milgram and Kishino [1994] and Azuma [1997] derived their definitions based on existing AR systems, and at the time of their publication, the focus of the AR community (and the related fields) was still on the feasibility of AR display technology.

Chapter 4 provides a detailed overview of past definitions of AR (see Chapter 4 Figure 3), highlighting its evolution with emphasis shifting from initial device-oriented definitions (Sutherland [1968], Milgram and Kishino [1994], Azuma [1997]) towards experiential and content-oriented definitions (Azuma [2016, 2017, 2019]).

In this thesis, we adopt an experiential and content-oriented view of AR. However, the core of our investigations remains confined within the visual modality. In light of our observations regarding prototypes and proof-of-concept systems in prior work, in Chapter 3, we generated a practical definition for AR content as:

> Dynamic or static imagery, real or computer generated, 2-D or 3-D. 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.

1.1.3 Applications of Augmented Reality in Television

Researchers and broadcasters have been experimenting with various ways in which AR can be used in the context of TV broadcasting. They have created prototypes and proof-of-concept systems that use AR in production (e.g., Denia et al. [2011]) and distribution (e.g., Sotelo et al. [2017]) of TV content. These systems typically intend to reduce the cost and increase the efficiency of the production and distribution pipelines. Chapter 2 expands these findings. We provide a few preliminary insights in this section.

Beyond the production and distribution of TV content, one of the main anticipated applications of AR in the context of TV broadcasting, originates from AR's ability to display content anywhere in the viewing environment; content that can be 3-D and registered in 3-D and interactive. These affordances can impact a viewer's experience to the point of transforming a conventional TV viewing experience. In this context,

investigations into AR's application in TV consumption can be categorised in two themes: 1) enhancing/transforming a conventional TV viewing experience in the living room, and 2) novel content viewing experiences outside the living room.

ARTV in the living room

In a systematic literature review Saeghe et al. [2020] found that the most commonly targeted application area for AR in the context of TV viewing, has been to enhance or transform a conventional TV viewing experience. This included systems and prototypes that use AR to deliver additional content (e.g., Kawakita and Nakagawa [2014], or additional screens and control (e.g., Baillard et al. [2017], Vatavu [2012]); to extended the real estate of a TV display (e.g., Kawamura and Otsuki [2018]); to replace a TV screen completely (e.g., Zimmer et al. [2018]); and to connect at-a-distance viewers (e.g., Vatavu [2015]).

ARTV outside the living room

The investigations in delivering TV content outside the living room, typically focus on systems that deliver location-based TV content, for instance to enhance cultural tourism (e.g., Debandi et al. [2018]), or to foster film tourism (e.g., Park and Woo [2015])—a phenomenon that typically consists of visiting film locations.

1.2   Aims and research questions

In this thesis, we explore three research questions:

RQ1 What is the underlying design space of ARTV?

RQ2 How can we support practitioners to navigate this design space?

RQ3 What are viewers' perceptions of ARTV?

To answer RQ1, we first identify six design dimensions commonly used in ARTV work, via a systematic literature review (Chapter 2). We then conduct a user study ($n = 10$) asking practitioners to write ARTV scenarios and interview them. Based on the findings from scenario analysis and subsequent interviews, we then refine the design space (Chapter 3).

Regarding RQ2, we create a "cheat sheet" that presents design dimensions as questions and subsequently evaluate the utility of the cheat sheet and its usefulness

in conceptualising novel ARTV scenarios, in a user study with $n = 10$ practitioners (Chapter 3).

Regarding RQ3, since a conventional TV viewing experience is traditionally associated with the living room, we first create a conceptual framework for ARTV experiences in the living room (Chapter 4). We then examine viewers' perceptions in the context of ARTV experiences in the living room, with a focus on nature documentaries in two user studies. In the first ($n = 12$), we add holograms—displayed via a Microsoft HoloLens device—to an existing nature documentary, displayed on a physical TV set. In the second ($n = 10$), we focus on investigating the extent to which AR holograms can be used to bring living creatures out of the TV and into the viewing environment, and the impact this has on the viewers. We investigate this in the context of an ARTV experience where both the TV and the living room are virtual, and explore six points on the display dimension of the hologram.

Future work can expand our approach and explore viewers' perceptions regarding other genres and other design aspects of ARTV.

1.3 Contributions

The overarching contribution of this thesis is to provide content creators and broadcasters, the target audience of this thesis, with a set of methods and practical tools to create content that takes advantage of the affordances of AR in the context of audio-visual content viewing. This thesis contributes to the field of Human-Computer Interaction (HCI) and the emerging field of Augmented Reality Television. Our contributions are:

1. A refined ARTV design space—presented as a set of questions—demonstrated to help practitioners in conceptualising novel ARTV experiences.

2. A conceptual framework for ARTV experiences in the living room, defining nine specific types of ARTV and highlighting areas open for further research.

3. Novel insights into viewers' perceptions of ARTV in the context of nature documentaries augmented with lifelike holograms.

1.4 Thesis structure

This thesis is submitted in a journal/alternative format with permission from the supervisory team from the Faculty of Science and Engineering. This means that the thesis'

core chapters (i.e., Chapters 2 to 6) are published papers (details below). To maintain a clear narrative throughout the thesis, a section named "Thesis context" is added at the beginning of each of these chapters to link them to the previous chapters and highlight their contributions to the overall work. We chose to present this thesis in a journal format because each of its main chapters depicts its own picture that can be read and comprehended individually. The content of each chapter, along with its publication, are described below:

- **Chapter 2** presents a systematic literature review of prior work in the field of ARTV. We identify six main themes in which AR has been used in the context of TV broadcasting; with the most widely used theme being *enhancing a conventional TV viewing experience*. Furthermore, we identify six commonly used ARTV design dimensions and propose six design guidelines.

  The content of this chapter is adapted from: Pejman Saeghe, Gavin Abercrombie, Bruce Weir, Sarah Clinch, Stephen Pettifer, and Robert Stevens. Augmented Reality and Television: Dimensions and Themes. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 13–23, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393649. URL https://doi.org/10.1145/3391614.3393649.

- **Chapter 3** builds on the six design dimensions identified in Chapter 2 to generate an ARTV design space. To support implementation of the design space, we present its dimensions as a set of questions compiled into a cheat sheet document. We therefore evaluate the intelligibility and usefulness of this cheat sheet in a user study ($n = 10$). Our findings suggest that the cheat sheet is useful in conceptualising novel ARTV scenarios. In light of our findings, we update the previous ARTV design space and cheat sheet.

  The content of this chapter is adapted from: Pejman Saeghe, Mark McGill, Bruce Weir, Sarah Clinch, and Robert Stevens. Evaluating and updating a design space for augmented reality television. In *ACM International Conference on Interactive Media Experiences*, IMX '22, page 79–94, New York, NY, USA, 2022a. Association for Computing Machinery. ISBN 9781450392129. doi: 10.1145/3505284.3529965. URL https://doi.org/10.1145/3505284.3529965.

- **Chapter 4** creates a conceptual framework for ARTV in the context of enhancing/transforming a conventional TV viewing experience in the living room.

The content of this chapter is adapted from: Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. Conceptualizing Augmented Reality Television for the Living Room. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 1–12, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393660. URL https://doi.org/10.1145/3391614. 3393660.

- **Chapter 5** presents a pilot study ($n = 12$), where a conventional TV viewing experience is augmented using programme-related AR content. Our findings suggest that while the experience was engaging, some participants reported a worry that they may have missed parts of the underlying TV content and the narrative.

  The content of this chapter is adapted from: Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, pages 255–261, New York, NY, USA, 2019a. ACM. ISBN 978-1-4503-6017-3. doi: 10.1145/3317697.3325129. URL http://doi.acm.org/10.1145/3317697.3325129.

- **Chapter 6** presents a user-study ($n = 10$), where a conventional TV viewing experience is simulated in VR, and a lifelike programme-related hologram is presented to the viewer on the backdrop of a nature documentary. Our findings highlight the importance of personal preferences in the context of an ARTV experience.

  The content of this chapter is adapted from: Pejman Saeghe, Bruce Weir, Mark McGill, Sarah Clinch, and Robert Stevens. Augmenting a nature documentary with a lifelike hologram in virtual reality. In *ACM International Conference on Interactive Media Experiences*, IMX '22, page 275–280, New York, NY, USA, 2022b. Association for Computing Machinery. ISBN 9781450392129. doi: 10.1145/3505284.3532974. URL https://doi.org/10.1145/3505284. 3532974.

- **Chapter 7** concludes this thesis by presenting a summary and final versions of the design space concepts and reflecting on the results to present ARTV as

a new storytelling vehicle. This is followed by sections describing limitations, recommending potential future work, and revisiting the research questions and outlining the main findings.

Chapter 2

ARTV: Dimensions and Themes

2.0 Chapter overview

2.0.1 Thesis context

The previous chapter provided preliminary definitions and insight into this thesis'
key concepts (i.e., TV, AR, and ARTV). This chapter presents a systematic literature
review of ARTV work and identifies six key research themes and six underlying design
dimensions commonly used in the work. It also provides suggestions to operationalise
the design dimensions and usage guidelines.

The main content of this chapter is adapted from: Pejman Saeghe, Gavin Abercrom-
bie, Bruce Weir, Sarah Clinch, Stephen Pettifer, and Robert Stevens. Augmented Reality
and Television: Dimensions and Themes. In *ACM International Conference on Interac-
tive Media Experiences*, IMX '20, page 13–23, New York, NY, USA, 2020. Association
for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393649.
URL https://doi.org/10.1145/3391614.3393649.

2.0.2 Author's contributions

Pejman Saeghe designed and executed the systematic literature review, analysed the
results, synthesised the emergent design dimensions, wrote and presented the paper.
Dr Gavin Abercrombie helped with the systematic literature review as a second coder.
Both Dr Gavin Abercrombie and Dr Sarah Clinch participated in a group meeting

with Pejman Saeghe to identify emergent dimensions and themes and helped with the classification of work. Prof Robert Stevens, Dr Bruce Weir, and Dr Sarah Clinch provided continuous support, critical feedback, and edits in the writing of the paper.

2.0.3   Published abstract

Commercialisation of augmented reality (AR) devices has led to their growing application in domestic environments and leisure activities. One such domain is that of television, where AR is one of several technologies driving innovation (c.f. Internet broadcasting, second screen devices). We conduct a systematic literature review to quantify research at the intersection of AR and broadcast television. We identify six common themes and a set of cross-cutting design decisions. We distill this information into a design space incorporating six dimensions: *abstraction*, *interaction*, *time*, *display*, *context* and *editorial control*. We provide methods to operationalise the dimensions to enable research and development of novel concepts, and through this generate six design guidelines to shape future activity at the intersection of AR and television.

2.1   Introduction

Advances in technology lead to changes in broadcast media. This introduces challenges and opportunities for content creators. In this paper we seek to provide insights into the design space of a hybrid medium composed of mixing Augmented Reality (AR) with TV.

Various aspects of broadcast media that relate to the production, distribution, and consumption of content, are affected by advances in technology. This in turn affects the viewers' experience. For instance, the introduction of the Internet changed the way the audiences received TV-related data services—from Teletext data embedded in the broadcast signal, to programme specific websites delivered via TCP/IP.

Similarly, recent advances in technologies that enable immersive representations, such as virtual, mixed and augmented reality, have raised the interest of the broadcast and research communities.

There are two ways in which the immersive paradigm can be used: on its own, or combined with another medium. The success of various aspects of the immersive paradigm—as standalone media—may depend on long-term experimentation. These experimentations are often necessary to establish conventions and rules that the content

creators implement and the audiences understand. In the meantime, the new medium can be used to remediate [Bolter et al., 2000] existing content and conventions, or in combination with an already established medium such as television.

We select AR from the range of possibilities within the immersive paradigm[1] and investigate how it can be combined with television to create novel experiences by transforming the conventional TV viewing experience.

Our overall aim is to provide a framework that broadcasters and content creators can use when creating content that mixes conventional TV with AR. We do this by making the following contributions:

1) A systematic literature review (SLR) that captures publications at the intersection of TV broadcasting and AR between 2008 and 2018;

2) Identification of six themes (i.e. common motivations and/or problem domains) that emerge from the SLR that can be used to understand the state-of-the-art in the field;

3) Abstraction of design decisions into a design space composed of six dimensions;

4) An initial exploration of the design space through (a) two methods of generating 'patterns' that operationalise the dimensions to realise points on the design space, and (b) six practical design guidelines to be considered by content creators when designing novel experiences for the hybrid medium.

When creating experiences that mix AR with TV, content creators are often faced with questions such as:

- What is the relationship between the AR and TV components in such experiences?

- How should the viewers interact with the experience?

- How and when should various components be presented to the viewers?

- What is the role of the viewers' environment in their experience?

We set out to find ways that could be useful in providing answers to questions of this nature and provide concrete guidelines where possible.

Our effort is focused on visual augmentations. Auditory and other sensory augmentations are outside the scope of this paper. We excluded VR from our work, since the

---

[1]AR has the potential to become ubiquitous, due to its non-invasive nature.

broader scope needed to include it in the same study would lead to sacrificing depth for breadth. We suggest that similar investigative work needs to address comparable challenges when combining traditional media (such as TV), with other forms of the immersive paradigm (such as virtual reality).

2.2   Definitions

We use existing definitions of the terms and concepts in the immersive paradigm and television to explore the characteristics of the emerging hybrid medium that mixes aspects of the two (AR and TV).

Some terms in the immersive landscape have highly standardized definitions (e.g.,Virtual Reality), while others have been overloaded with more than one meaning (e.g., Mixed Reality). In such cases, the context in which a term is used may help to disambiguate its intended meaning [Speicher et al., 2019]. For clarity and to provide context we provide widely accepted definitions of key concepts related to this field.

Figure 2.1: Reality-Virtuality continuum based on Milgram and Kishino's [1994] taxonomy of mixed reality displays.
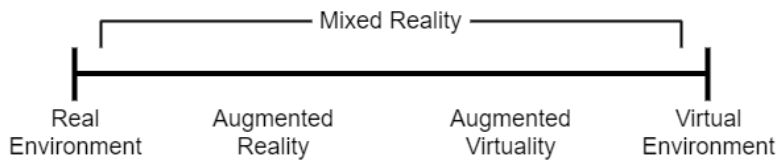
2.2.1   The immersive paradigm

Milgram and Kishino [1994] were the first people to present the idea of a continuum to describe this paradigm. Subsequently, Milgram et al. [1995, 1999] referred to this as the Reality-Virtuality continuum (RV).

The *real environment* and the *virtual environment* are placed at opposite ends of the RV. Between the two extremes, there are three concepts: Augmented Reality (AR), Mixed Reality (MR), and Augmented Virtuality (AV) (see Figure 2.1).

Real and virtual environments   The *real environment* refers to a viewer's immediate physical environment. Their view may be direct (e.g., naked eye, see-through glasses), or indirect (e.g., live video feed, pre-recorded video). Conversely, the *virtual environment* refers to *a computer generated model* of an imaginary (or a real) environment.

Mixed Reality   Milgram and Kishino [1994] referred to any experience that falls between the two ends of RV as *mixed reality*. However, since then, this term has been used inconsistently. People from neither the researcher nor practitioner communities have an agreed upon definition for this term [Speicher et al., 2019]. MR is often used interchangeably with AR [Craig, 2013].

Augmented Virtuality   This refers to cases where the virtual world dominates the viewers' view of the world, and real objects are used to augment this view.

Augmented Reality   AR was first implemented in 1968 by Sutherland to display three-dimensional information to the user [Sutherland, 1968]. It was later defined by Milgram and Kishino to refer "to all cases in which the display of an otherwise real environment is augmented by means of virtual (computer graphic) objects" [Milgram and Kishino, 1994]. Since then, this definition has been revised and refined multiple times.

Azuma [1997] defined AR as "a variation of Virtual Environments" which enables "the user to see the real world, with virtual objects superimposed upon or composited with the real world" on the same display. He later suggested three requirements for any AR system: 1) combining "real and virtual objects in a real environment", 2) running "interactively, and in real time", and 3) aligning "real and virtual objects with each other" [Azuma et al., 2001].

In this paper we adopt Milgram and Kishino's [1994] definition of AR, since it is the most widely accepted definition [Speicher et al., 2019].

2.2.2   Television

Television was originally conceived in the 1920's as a one-to-many audio-visual successor to radio. It has, however, evolved, benefiting from other technologies and media, in order to improve on how it delivers content and engages with its audiences.

The term TV is overloaded with multiple meanings to reflect various aspects of the industry involved in production, distribution and consumption of audio-visual content. The online English Oxford dictionary defines *television* as: "*a device* shaped like a box with a screen that receives electrical signals and changes them into moving images and sound, or *the method* or *business* of sending images and sound by electrical signals".

The Internet, as a novel way of distribution, and portable displays, as novel ways of consumption, have transformed the audiences' perception of television and the ways in

which TV content is consumed. While many people may not use traditional broadcast channels or use a physical TV set, they may use streaming services (e.g., YouTube and Netflix) and their hand-held displays to watch content that was originally made for TV broadcast by broadcasting corporations (e.g., watching the BBC's The Blue Planet series on Netflix).

This suggests that, Television should no longer be considered merely the physical artefact which is used to display images, but instead as a medium through which a viewer consumes, typically passively, audio-visual content. For example, TV content can be viewed on a physical box in the corner of a living room, streamed to a mobile phone, or rendered onto the side of a building using an AR display.

In this study we include any research that uses the term *television* (or *TV*) to describe the work.[2]

2.3   Method

Our research is founded on systematic literature review and qualitative analysis methodology. Focusing on recent work, we followed the PRISMA methodology [Moher et al., 2009] to identify relevant literature published in the 11 years between 2008 and 2018. A representation of the five stages of this process is shown in Figure 2.2.

2.3.1   Paper selection

Our initial search terms can be represented as follows:

```
(``augmented reality'' OR ``mixed reality'')
                    AND
        (television OR TV OR broadcast*)
        [title/abstract/author keywords]
```

We identified three target databases (SCOPUS, ACM Digital Library[3], and IEEE) and generated platform-compatible queries for each. Each database search was conducted using the web interface provided. A total of 417 publications were returned by the search (SCOPUS $n = 193$, ACM $n = 57$, IEEE $n = 167$).

---

[2]Except for when the TV is used as a mere display device.

[3]A new ACM digital library (DL) was launched publicly in December 2019. Our search was conducted prior to this and thus used the old DL. We have since rerun the query on the new DL and no difference in results was seen.

Figure 2.2: Adapted PRISMA flowchart for selecting publications.
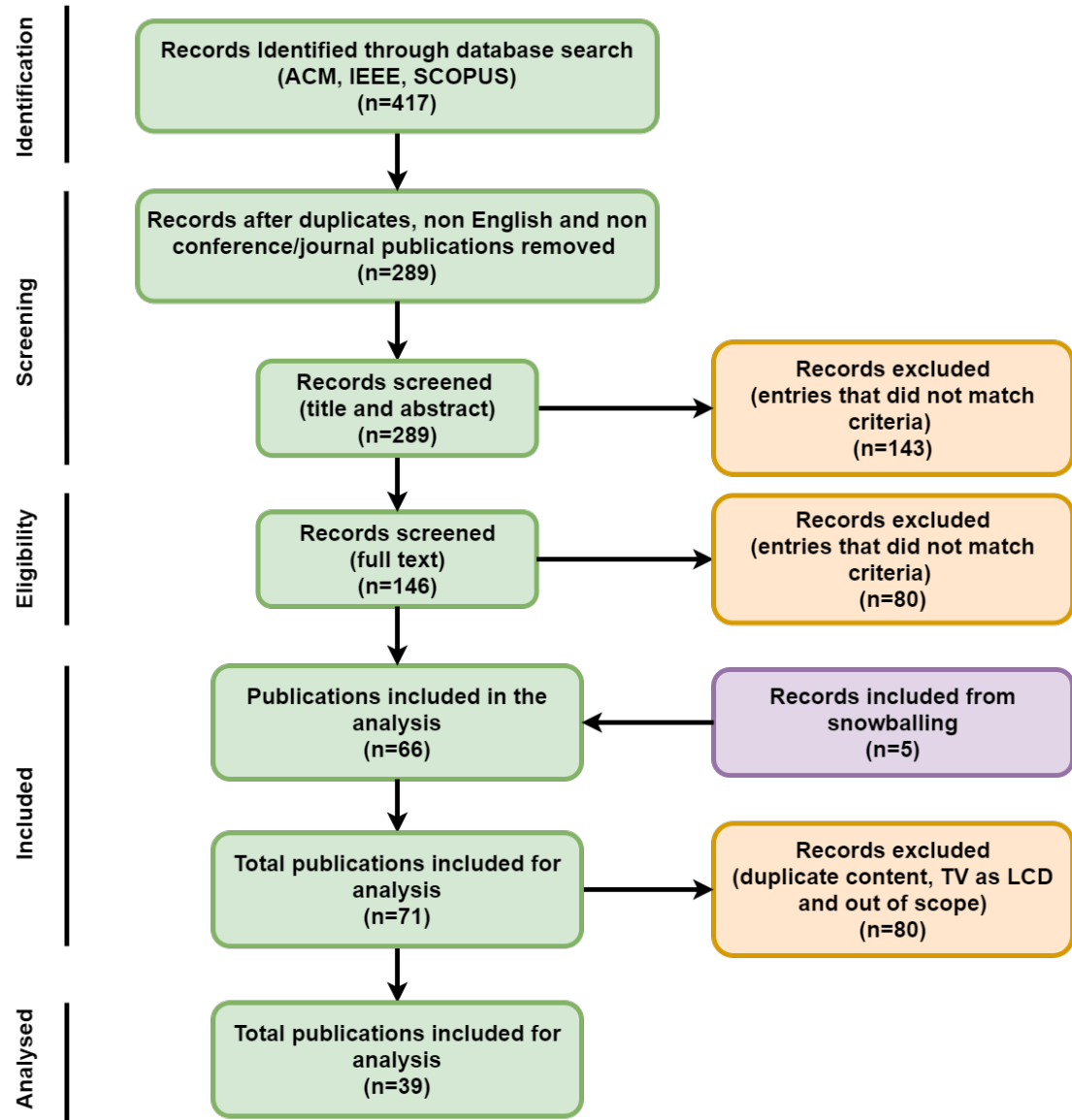
We excluded duplicates ($n = 121$), papers written in languages other than English ($n = 2$), and results that had not been published as peer-reviewed conference or journal papers ($n = 5$).

The remaining 289 publications were screened by the first two authors, who independently read the titles, abstracts and keywords of the papers. Using the inclusion/exclusion criteria (Table 2.1), a further 143 papers were excluded.

Table 2.1: Inclusion/exclusion criteria used to screen the papers returned from the database search.

Inclusion criteria
Returned by one or more target databases
Peer reviewed conference/journal paper
Uses AR/MR in accordance with widely accepted definitions [Milgram and Kishino, 1994].
Incorporates some television broadcasting element.
Describes concrete scenario or implemented prototype.
Exclusion criteria
Non-English language title/abstract/full text.
Duplicates a system or scenario present in another (more recent) paper in the dataset.
No AR (MR) OR no TV present
Posters, talks, courses, panels, demos, white papers etc.

Inter-coder agreement on this task was calculated to be Cohen's *kappa* (k) $= 0.49$, and the authors subsequently resolved all disagreements through discussion.

The same two authors then screened the full texts of the remaining 146 papers, again applying the inclusion/exclusion criteria. This resulted in the removal of a further 80 papers. Agreement at this stage was calculated to be $k = 0.85$, and disagreements were once again resolved through discussion.

These screening stages produced a pool of 66 papers. Forward and backward citations were followed in accordance with Wohlin's guidelines for snowballing [Wohlin, 2014], resulting in the inclusion of five further publications.[4]

A further 14 papers were excluded from this set due to duplicate content, where the same authors described the same concept but published in a different venue. In these cases we kept the most recent publication. A further 18 papers were excluded due to their use of the term *TV* solely as a display device to show non TV content. This reduced the number of included publications to 39.

Relying on our knowledge of this area, we added a further three papers (Vatavu

---

[4]These had not been picked up in the original search as their publishers (Springer Link, Semantic Scholar and the Wiley Online Library) are not indexed by any of the three target databases.

[2012, 2013, 2015]) that were published within the same time-frame but were not returned by the systematic search. These papers were not retrieved due to discrepancies in the terminology used. For instance, Vatavu [2015] used the term *augmented TV*, without mentioning directly the term *augmented reality*, in the title, abstract, and keywords. The final number of papers analysed was 42.

Inclusion of MR   We used MR in our query, to ensure the capture of all AR-related papers, since AR and MR are often used interchangeably [Craig, 2013] (see Section 2.2 for further detail).

2.4   Themes

We present six themes that emerge from the qualitative analysis of the publications. These are: 1) Enhancing TV experiences in living room, 2) Production of TV content, 3) Alternative TV experiences beyond living room, 4) Connecting remote viewers, 5) Live video augmentation, and 6) Photogrammetry.

Table 2.2 presents the classification of the papers by theme.

Table 2.2: Publications per themes.

Theme	Papers
Living room	Baillard et al. [2017], Bibiloni et al. [2015], Chandrasekaran and Kesavan [2017], Gómez et al. [2014], Jones et al. [2013], Kimura and Rekimoto [2018], Kawakita and Nakagawa [2014], Kawamura and Otsuki [2018], Kim et al. [2017], Revelle et al. [2015], Rodrigues et al. [2017], Sotelo et al. [2017], Vatavu [2012, 2013], Vinayagamoorthy et al. [2018], Zimmer et al. [2018]
Production	Cho et al. [2018], De Gaspari et al. [2014], Herder et al. [2018], Hough et al. [2014], Kim [2011], Li et al. [2012], Oyarzun et al. [2010], Sanna et al. [2017], Schiller et al. [2010], Simsch and Herder [2014], Yan and Zhang [2011], Sitnik et al. [2008], Zhang and Zhu [2018]
Beyond living room	Debandi et al. [2018], Ghellal and Lindt [2008], Park and Woo [2015, 2017]
Remote viewers	Lee et al. [2017], Li [2010], Schreer et al. [2016], Vatavu [2015]
Live video	Cushen and Nixon [2012], Denia et al. [2011], Maalej et al. [2015]
Photogrammetry	Han et al. [2010], Saito et al. [2014]

Our aim is to identify the themes that emerge from a representative sample of the literature. A comprehensive overview of each theme is outside the scope of this paper.

2.4.1   Enhancing TV experiences in living room

This theme covers the largest number of publications in our review ($n = 16$). These studies describe prototypes or scenarios where AR is used to enhance or to transform a traditional TV viewing experience in the living room. We identified three sub-themes that best describe the research in this area: *TV + AR*, *focus + context*, and *TV through AR*.

TV + AR   AR can be used to deliver virtual content [Revelle et al., 2015, Kawakita and Nakagawa, 2014], or virtual screens and controls [Baillard et al., 2017, Vatavu, 2012, 2013] outside the physical TV frame.

Focus + context   The implementation of a *focus + context* [Cockburn et al., 2009] metaphor, allows the viewers to see contextual images in the area around the TV frame. AR can be used in this context to increase immersion [Kimura and Rekimoto, 2018, Jones et al., 2013].

TV through AR   AR can replace the physical TV set altogether. By eliminating the conventional TV frame (both physical and virtual) and delivering content that appears to be present in the viewers' environment [Zimmer et al., 2018].

2.4.2   Production of TV content

This theme contains the second largest number of papers in the survey ($n = 13$). Two sub-themes best describe the research in this area: *pre-production* of TV content, and *virtual studio*. The main differentiating factor between the two is that the virtual content used in a *virtual studio* is visible to the viewers at home, whereas the virtual content used in *pre-production* is ultimately replaced by either real artefacts or higher quality renders.

Virtual studio   Virtual studio is a mature technology that adds locked-to-camera-view virtual objects to a TV image in real-time. See Gibbs et al. [1998] for an overview. Although researchers often motivated their work by proposing systems for *reducing costs* and *increasing productivity* [Cho et al., 2018, De Gaspari et al., 2014, Kim, 2011, Sitnik et al., 2008] when creating TV content, a virtual studio can also be used to create a production-set which would be physically prohibitive to create.

Pre-production   In the pre-production phase of content creation, AR can be used to help directors and set designers visualise a production-set. Using AR, they can place and manipulate various virtual artefacts before committing to a final decision [Sanna et al., 2017]. Alternatively, it can be used during filming to provide a low quality, real-time view of the virtual elements to allow correct shot framing before the creation of the final composition.

2.4.3   Alternative TV experiences beyond living room

There are studies that suggest use-cases, at the cross-section of AR and TV that take the TV viewing experience outside the living room. We identified two sub-themes in this category that achieve this in different ways: *Cultural Tourism* and *Film Tourism*.

Cultural tourism   Broadcasters typically own large archives of audio-visual content. This content can often be remediated to create novel experiences using AR. For instance, delivering extra information about a historical landmark using AR [Debandi et al., 2018] or displaying how a landmark looked in certain eras of history.[5]

Film tourism   Film tourism is defined as tourist visits to destinations as a result of those locations being featured on television or film [Hudson and Ritchie, 2006]. AR can be used to enhance film tourism in two ways: by enabling the visitors to watch clips from TV shows that were filmed in important landmarks, when they are vising these landmarks; or by enabling a TV show's fans to follow a specified route and visit multiple filming locations, watching clips from the scenes that were filmed in those locations [Park and Woo, 2015, 2017].

2.4.4   Connecting remote viewers

Television has traditionally been viewed as a medium that brings family and friends together for a shared viewing experience. As the options for viewing content proliferate, there may be fewer opportunities for co-located TV viewing. AR's affordances can be used to simulate a shared viewing experience, for instance, by using virtual avatars [Lee et al., 2017] or overlaying body silhouettes [Vatavu, 2015] on the broadcast video.

---

[5]https://www.bbc.co.uk/rd/blog/2019-02-5g-mobile-augmented-reality-bath

2.4.5   Live video augmentation

Advancements in computer vision and related fields can enable the real-time identification and tracking of people and objects present in a live video feed. This information can subsequently be used to overlay virtual content and related information aligned to the people and objects in the scene. In sports broadcasting this can be used to re-texture players' shirts with localized advertisements or statistics [Cushen and Nixon, 2012], or to simulate animated effects performed by the audience of a sports match in a cost-effective way [Denia et al., 2011].

2.4.6   Photogrammetry

Photogrammetry attempts to obtain "reliable information about the physical objects and the environment through the process of recording, measuring, and interpreting photographic images."[6] In the context of AR and TV, this has been applied to generate 3-D information from 2-D images in order to create a virtual model of the scene [Han et al., 2010]. This typically attempts to provide the viewers with a set of options regarding camera angles and ultimately a *free-viewpoint* viewing experience.

2.5   Dimensions of the design space

The *themes* describe the research efforts from a high-level perspective. There are, however, cross-cutting design decisions within the themes that can be further distilled. We present these as six primary dimensions that emerge from a qualitative analysis of the publications.

Our motivation is to enable further exploration of the design space in order to help the content creators and to identify under-researched areas.

The dimensions are:

1) The level of *abstraction* between AR and TV, 2) The nature and level of viewers' *interaction* with the experience, 3) The relationship between AR and TV in terms of *time*, 4) Their relationship in terms of *display* location, 5) The influence of the viewers' *context* on the experience, and 6) The level of *editorial control* on the experience.

The dimensions are described in terms of their 'aspects'. Table 2.3 presents these dimensions and their aspects.

---

[6]https://en.wikipedia.org/wiki/Photogrammetry

Table 2.3: Six dimensions of the ARTV design space and their aspects.

Dimension	Aspect
Abstraction	Independent; dependent; additional AR; additional TV
Interaction	Display-level; structure-level; content-level
Time	Continuous; intermittent; asynchronous
Display	Same display; separate displays; TV via AR
Context	People; place; objects
Editorial Control	Point of view; sequence

2.5.1 Abstraction

We build on top of Jones et al.'s [2013] *abstraction* dimension, and extend it to describe the semantic relationship between AR and TV. We consider the 'wholeness' and the 'completeness' of the experience to be the key deciding factors. In other words, either one or both components (AR and TV) are independently complete and meaningful, or the completeness of the experience relies on the presence of one or both components.

Independent   This is when AR and TV experiences are independently complete. They may, however, be semantically related, for instance because they both originate from the same idea or are both based on the same script. This approach can be used to create self-contained mixed and augmented reality experiences based on a TV show [Ghellal and Lindt, 2008].[7]

Dependent   In this case the 'wholeness' of the overall experience depends on the existence of both components. For instance, the content created in a *virtual studio* setting often requires the existence of both elements; removing a virtual presenter [Oyarzun et al., 2010] from the TV show would make the entire experience meaningless. This type of *abstraction* can be used to create novel engaging experiences across TV and AR [Revelle et al., 2015].

Additional AR   There are experiences that mix AR and TV in such a way that the TV experience remains meaningful after elimination of AR. In these cases, AR is *dependent* on the TV content while TV is meaningful *independently*. In such cases AR is often used to enhance the TV viewing experience by providing extra features or novel interaction techniques [Baillard et al., 2017, Vatavu, 2012, 2013, 2015].

---

[7]The BBC's Civilisations AR app and Pokémon Go are other such examples.

Additional TV   We found no instances where the experience remains meaningful once the TV component is eliminated.[8]

2.5.2   Interaction

Interactions enable the viewers to dynamically change an aspect of their experience. We found three types of interactions in the literature that reflect the goal of the designer: *display-level*, *structure-level*, and *content-level*.

Display-level   These interactions enable the viewers to perform tasks such as: changing the programme, resizing or repositioning the content, etc. [Baillard et al., 2017, Vatavu, 2012, 2013]. A more advanced *display-level* interaction may enable the viewers to change camera angles [Han et al., 2010].

Structure-level   This level of interaction breaks the linear structure and enables the viewers to only view the elements that interest them in any order [Bibiloni et al., 2015]. This is akin to the concept of hyperlinks in the textual paradigm.

Content-level   This level of interaction is akin to that in the gaming paradigm which enables direct manipulation of the elements within the content, such as characters and objects [Ghellal and Lindt, 2008, Oyarzun et al., 2010, Kawakita and Nakagawa, 2014, Kim et al., 2017, Revelle et al., 2015, Schreer et al., 2016, Sotelo et al., 2017]

2.5.3   Time

Traditional television content is typically delivered on a timeline. Similarly, the presentation of AR content can be visualised on a timeline. When delivering experiences that mix AR with TV, researchers have often attempted to synchronise these distinct timelines [Kawakita and Nakagawa, 2014, Kawamura and Otsuki, 2018]. From the point of view of the viewers AR and TV components can be viewed in three ways: *continuous*, *intermittent*, or *asynchronous*.

Continuous   Here, both components (AR and TV) are presented to the viewers at the same time [Kawakita and Nakagawa, 2014, Kawamura and Otsuki, 2018] (Figure 2.3).

---

[8]Except when the two experiences are independent, or when AR is used to deliver TV content.

Intermittent In this case, the experience consists of presenting content over TV and AR intermittently, for instance to mix an AR game with an educational TV programme [Revelle et al., 2015] (Figure 2.4).

Asynchronous There are cases where synchronisation of AR with TV is not the aim of the designer. Often, these are *independent* AR and TV experiences (see Abstraction) [Ghellal and Lindt, 2008].

Figure 2.3: Visualisation of continuous time.
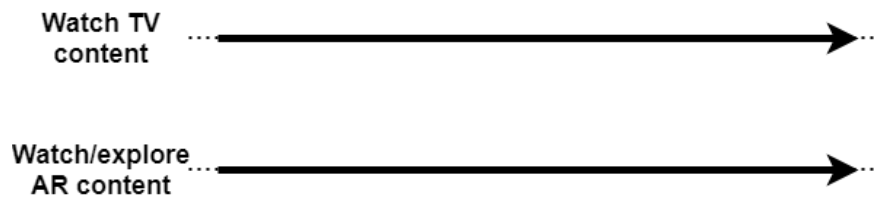
Figure 2.4: Visualisation of intermittent time.
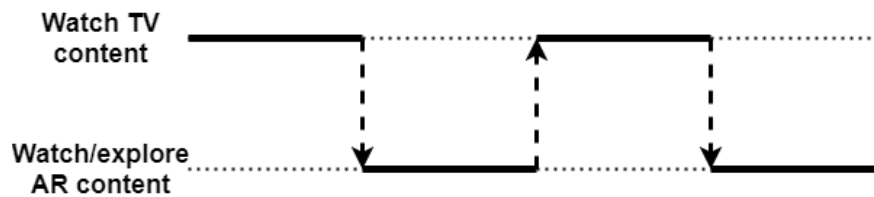
2.5.4 Context

Context reflects the designers' goals regarding the ways in which, if at all, the viewers' surrounding environment affects the experience. We present these in three classes of: *people*, *place*, and *objects*.

People There are instances of mixing AR with TV that have been designed with multiple viewers in mind. For instance, to enable friends and family members [Baillard et al., 2017, Lee et al., 2017, Revelle et al., 2015, Schreer et al., 2016, Vatavu, 2012, 2015], or colleagues [Herder et al., 2018] to share aspects of an experience.

Transforming or adapting such experiences based on the presence of bystanders and passers-by, and the the primary viewers' relationship and interactions with them, may be a necessary step to make such experiences acceptable in public.

Place   In some experiences the content is anchored to specific physical locations and becomes available when the viewer reaches these locations [Debandi et al., 2018, Park and Woo, 2015, 2017].

Objects   In its simplest form, the system is aware of the existence and location of the TV screen. This enables AR content to be displayed in the area surrounding the TV [Vatavu, 2013]. Other implementations use knowledge of the surface colours and geometry, and apply *radiometric compensation* [Bimber and Raskar, 2005, Bimber and Iwai, 2008, Raskar et al., 2001] to adapt the appearance of the content to the viewers' environment [Jones et al., 2013].

2.5.5   Display

When combining TV and AR, content can be displayed in three ways: on the same display, on separate displays, or TV content presented using AR.

AR and TV on the same display   Often the results of mixing AR with TV is displayed on the same display. Examples include the content created in a *virtual studio* setup [Denia et al., 2011, Oyarzun et al., 2010], or experiences that enable viewers to view or interact with additional content using AR [Sotelo et al., 2017, Vatavu, 2015].

AR and TV on separate displays   This is akin to the use of second screens for viewing additional content while watching TV [Geerts et al., 2014, Lohmüller and Wolff, 2019]. In the context of mixing AR with TV, the secondary display can be used to create novel TV viewing and cross-media experiences [Kawakita and Nakagawa, 2014, Kawamura and Otsuki, 2018, Revelle et al., 2015, Vinayagamoorthy et al., 2018].

TV viewing with AR   An AR display can be used to replace the TV screen in two ways: viewing of TV content presented in *virtual screens* [Park and Woo, 2015, 2017, Vatavu, 2013], and display of TV-like content beyond any visible rectangular frame [Zimmer et al., 2018]. While the former can be readily applied to existing TV content, the latter requires creation of content using technologies such as chroma-keying and volumetric capture.

2.5.6 Editorial control

In a conventional TV programme, content creators often have control over two aspects of the viewers' experience: sequencing of media elements, and viewers' point of view. Mixing AR with TV enables new opportunities to delegate aspects of control back to the viewers.

Sequencing of media content TV content is often created by editing shots and sequences in a specific way, dictated by the director's style which is influenced by existing conventions. Viewers of a TV programme typically have no control over the order in which various elements of the programme are presented to them. In this context AR can provide opportunities to allow the viewers a free-exploration style approach to viewing the content [Bibiloni et al., 2015, Gómez et al., 2014, Park and Woo, 2015, 2017, Revelle et al., 2015].

Creating a successful experience may rely on the right combination of the two ends of this spectrum (fully edited vs fully exploratory), based on design intentions, genre, demographics of the viewers, etc.

Point of view Traditional audio-visual media (e.g., film and TV) fix the viewers' POV on the screen. This allows the director to dynamically change the POV—based on their stylistic choices and existing conventions—to guide the viewers' gaze through the story.

AR can be used to enable the viewers to select from a few possible camera angles [Han et al., 2010], or to provide a completely free-view point [Kawamura and Otsuki, 2018, Sotelo et al., 2017].

While the affordances of a screen (e.g., a TV display) make it suitable for delivering content with a fixed POV, AR can be used to deliver content with a fixed POV [Vinayagamoorthy et al., 2018], context-aware POV [Zimmer et al., 2018], viewers' choice of camera angle [Han et al., 2010], or fully free POV [Kawamura and Otsuki, 2018].

2.6 Patterns

When mixing TV and AR, the dimensions can be combined to create unique patterns. A set comprised of the Cartesian product of all aspects of the dimensions results in 648 unique patterns. A closer look at this set, however, reveals that not all patterns are valid; Regarding the *abstraction* dimension (see Dimensions in Section 2.5), when creating *independent* AR and TV experiences there is no need for the two components to be

synchronised. Conversely, *dependent*, *additional AR*, and *additional TV* aspects of this dimension require a form of synchronisation between the TV and AR components. Eliminating the invalid patterns that combine these aspects reduces the number to 378 possible patterns.

While detailed enumeration and discussion of all possible patterns is outside the scope of this paper, we give examples of how to generate patterns and how to subsequently use them for design purposes. We suggest two methods for exploring the design space implied by the dimensions: sampling from the set of all possible patterns, and ideation cards.

2.6.1   Sampling from the set

We present two random samples from the set generated by the Cartesian product of all dimensions (and their aspects) and explore the possibilities by writing two scenarios.

Sample 1

- Abstraction: dependent

- Interaction: content-level

- Time: intermittent

- Display: separate displays

- Context: location-aware

- Editorial Control: exploratory AR

Scenario 1   Imagine a crime series where the story elements are divided and told across both TV and AR. While the established conventions of television are used to create content that the viewers watch on the TV, following the end of the episode, the viewers need to find a clue to help the detective solve a crime. For instance, the living room may be transformed into the crime scene, or the viewers may need to visit a local café to find the *virtual* clue using their smart phone. They can only watch the next episode after 'throwing' the clue back into the physical TV set.

Sample 2

- Abstraction: independent

- Interaction: display-level

- Time: asynchronous

- Display: separate displays

- Context: location-aware

- Editorial Control: edited content

Scenario 2 Imagine in the previous crime scene a spin-off story is created for the AR platform. Hard-core fans of the show can experience these spin-offs of various characters in various locations. To view the content related to the backstory of a character that supposedly lives in London's East End, local fans can go to an old church in the area, find the specified bench and sit there next to the character and listen to them tell their backstory. The content is *independent* in the sense that it contains a stand-alone story and can even be understood by someone who does not watch the TV show. Interaction is display-level enabling the viewer to select content that they want to view.

2.6.2 Ideation cards

Mixed reality card games have been used as tools to enable collaborative design in the creation of mixed reality games in a playful manner [Wetzel et al., 2017]. We suggest a similar approach to foster collaborative content creation for experiences that mix TV and AR.

This can bring a team of interdisciplinary professionals (e.g., producers, engineers, screenwriters, filmmakers, playwrights, animators, computer graphics artists, game designers, etc.) together in the ideation phase and enable them to explore the design space from various perspectives.

We suggest an iterative approach to the designing of the cards. For instance using colours to represent the dimensions (e.g., red for *abstraction*). Various aspects of each dimension can then be represented using various shades of those colours.

In an imaginary ideation meeting the cards can be used in the following way: One person is in charge of the cards. They separate the cards based on colours and go through

each colour discussing the possibilities laid out by each dimension. For instance, regarding *abstraction* they can discuss the suitability of *dependent*, *independent*, *additional AR*, or *additional TV* aspects in creating the experience. Each party can present opportunities and challenges that each aspect presents to the project from their point of view. For instance, the screenwriter and the playwright discuss various potentials regarding character development, action, and location of the story; the engineer, animator and computer graphic artist can provide informed suggestions on the capabilities and limitations of the technology; the producers can have input on budget and editorial aspects of the content; the filmmaker discusses various styles of the audio-visual presentation provided by each aspect; and the game designer can provide input on how various aspects of *abstraction* may influence their choice when later they have to decide on the aspects of the *interaction* dimension.

They go through each dimension, discussing their ideas. It may be necessary to do this iteratively before the design is finalised.

2.7   Discussion and sample guidelines

2.7.1   Production, distribution, consumption, and method

Broadcasters' efforts can typically be categorised in three phases as: *production*, *distribution*, and *consumption*. These three categories are reflected in the contributions of the papers in the survey.

- Production-oriented papers often focus on enhancing the tools and processes involved in capturing and producing TV content.

- Distribution-oriented papers attempt to improve on existing ways or introduce novel ways of distributing the content.

- Consumption-oriented publications focus on enhancing existing experiences or creating novel viewing experiences.

We found papers with contributions regarding a *method* or an algorithm, in addition to (or instead of) directly contributing to the three aforementioned phases. Table 2.4 presents the papers and their contributions regarding these four categories.

The majority of the publications focus on either *production* or *consumption* with fewer papers focusing on *distribution*. The application of AR in *production* and *consumption* of TV content can be understood in the context of reducing production costs,

and creating novel experiences for the viewers, respectively. The limited number of contributions regarding *distribution* may be due to researchers' assumption that the content will be distributed using the Internet, as this has been the general tendency to distribute additional TV services.

There are, however, papers that cover all aspects of the broadcast chain [Han et al., 2010, Li, 2010, Kawakita and Nakagawa, 2014, Park and Woo, 2015, Schreer et al., 2016]. For example, enabling viewers to participate in a pre-recorded quiz show [Schreer et al., 2016] and broadcast the modified content to their friends. This suggests a shift in the viewers' role; from passively consuming content to actively creating new content—or manipulating existing content—and broadcasting it to their friends. Commercial availability of high quality tools to capture, manipulate and view content, may encourage more viewers to take part in this form of activity.

Design guideline 1   When creating experiences for the hybrid medium, consider providing features that enable the viewers to manipulate existing content, or to create original content, that the viewers can subsequently distribute via the Internet.

2.7.2   Viewers' attention

Attention can be seen as a limited resource and is commonly described using a *'spotlight'* metaphor, where "a particular region of the visual scene" is selected "for more detailed processing" [Driver, 2001]. When content is delivered over two separate displays in *continuous time* (see Section 2.5), some viewers may experience confusion and a sense of 'missing out' [Saeghe et al., 2019a]. Thoughtful orchestration of content and using appropriate cueing mechanisms—to orient [Posner, 1980] viewers' attention—can reduce this effect by reducing viewers' mental load.

Design guideline 2   When creating content for the hybrid medium, if the content is delivered in *continuous time* over separate screens, consider using visual and auditory cues to guide the viewers' attention.

It may also be possible to use other approaches to reduce the viewers' mental load. We present them as guidelines 3 and 4.

Design guideline 3   If the content is being delivered over separate displays, consider using *intermittent time* instead of *continuous time*.

Design guideline 4   When delivering AR content using a separate display in *continuous time*, visually limit the space where the AR content can appear (e.g., to the area surrounding the TV).

2.7.3 Interdisciplinary work

A wide range of expertise and knowledge may be required to create a successful experience that mixes AR with TV.

In TV often the space of action is limited to a 2-D screen. In contrast, AR enables the action to take place anywhere in the viewers' environment. From the viewers' point of view AR may be more similar to theatre than TV. In addition, AR enables transforming the viewers' environment into a story-space (e.g., by adding objects from the story). This suggests that the knowledge and expertise of a set-designer may be useful. Further, the viewers can interact with objects and characters within the story-space, which suggests that the knowledge and expertise of a game designer may be useful.

Design guideline 5   When creating content for the hybrid medium, consider using knowledge and expertise from a wide range of fields such as TV, gaming, theatre, set design, etc.

2.7.4 Dimensions' aspects

Abstraction, *time*, and *display* allow exactly one aspect to be selected per pattern. For instance, when combining TV with AR the relationship between the two can be either of *independent*, *dependent*, *additional AR*, or *additional TV*. However, *interaction*, *context*, and *editorial control* allow any number of aspects to be selected from each dimension (including no aspect at all). For instance, an experience—which mixes AR with TV—can allow *display-level* interaction as well as allowing *content-level* interaction. Conversely, it can be designed without any interactions at all. In cases where multiple aspects are allowed, designers can create richer experiences for the viewers by including more than one aspect per dimension.

Design guideline 6   When considering dimensions that allow multiple aspects, consider implementing more than one aspect to create a richer experience for the viewers.

2.7.5 Context-aware experiences

Experiences that mix AR with TV can be generated or adapted to take into account the viewers' context. So far the prototypes and scenarios use this dimension in a limited way and based on the available technology (e.g., anchoring video clips or 3-D models to a physical location). Generating ideas that take full advantage of this dimension may be used to direct the technology. For instance, adapting the content based on the viewers' relationship with bystanders or passers-by may be useful to avoid socially awkward situations.

2.7.6 Meaningless patterns

There are patterns that combine aspects of dimensions in a way that the resulting design pattern is invalid. Some aspects of dimensions cannot fit in the same pattern with certain other aspects of other dimensions.

We eliminated 135 patterns that contained aspects of *abstraction* and *time* that made them unusable or illogical. A systematic exploration of the patterns can reduce the size of design space further.

2.8 Conclusion

The technology that enables the realisation of AR experiences for the public is maturing. However, creating successful AR experiences requires an understanding of opportunities and challenges that this medium provides [Saeghe et al., 2019b]. In the context of broadcasting, individual efforts have been made to use AR's affordances to enhance or transform various aspects of the broadcast chain, although these efforts appear to be disjointed.

In this paper, our overall aim has been to provide an overview of the efforts in this field. This can help to identify opportunities for creating experiences that mix AR with TV. It can also highlight under-researched areas and guide future research in the field.

We have distilled information in the field—gathered from a representative sample of the literature—into an operable framework consisting of six primary dimensions of the design space—*abstraction*, *interaction*, *time*, *display*, *context*, and *editorial control*—and their respective aspects.

Researchers should be able to use the methods provided in this paper to systematically explore the design space as a way of generating guidelines.

Content creators and producers should be able to use these methods directly to make informed design decisions when creating experiences that mix AR with TV.

Table 2.4: Summary of the 42 papers analysed in terms of contribution areas. The majority of papers make contributions in only one broadcasting phase: *production* (pd.), *distribution* (dt.), or *consumption* (cs.). Some additionally provide *methodological* (md.) contributions (Maalej et al. is novel in contributing only through their method rather than to any specific broadcasting activity phase).

Paper	Contribution			
	Pd.	Dt.	Cs.	Md.
Ghellal and Lindt [2008]	−	−	✓	−
Sitnik et al. [2008]	✓	−	−	−
Han et al. [2010]	✓	✓	✓	✓
Li [2010]	✓	✓	✓	−
Oyarzun et al. [2010]	✓	−	✓	−
Schiller et al. [2010]	✓	−	−	✓
Denia et al. [2011]	✓	−	−	✓
Kim [2011]	✓	−	−	✓
Yan and Zhang [2011]	✓	−	−	✓
Cushen and Nixon [2012]	−	−	−	✓
Li et al. [2012]	✓	−	−	✓
Vatavu [2012]	−	−	✓	✓
Jones et al. [2013]	−	−	✓	✓
Vatavu [2013]	−	−	✓	✓
De Gaspari et al. [2014]	✓	−	−	−
Gómez et al. [2014]	−	−	✓	✓
Hough et al. [2014]	✓	−	−	−
Kawakita and Nakagawa [2014]	✓	✓	✓	✓
Saito et al. [2014]	−	−	−	✓
Simsch and Herder [2014]	✓	−	−	✓
Bibiloni et al. [2015]	✓	−	✓	−
Maalej et al. [2015]	−	−	−	✓
Park and Woo [2015]	✓	✓	✓	✓
Revelle et al. [2015]	−	−	✓	✓
Vatavu [2015]	−	−	✓	✓
Schreer et al. [2016]	✓	✓	✓	✓
Baillard et al. [2017]	−	✓	✓	−
Chandrasekaran and Kesavan [2017]	−	✓	✓	−
Kim et al. [2017]	−	✓	✓	✓
Lee et al. [2017]	−	✓	✓	✓
Park and Woo [2017]	−	−	✓	✓
Rodrigues et al. [2017]	−	−	✓	✓
Sanna et al. [2017]	✓	−	−	✓
Sotelo et al. [2017]	−	✓	✓	✓
Cho et al. [2018]	✓	−	−	−
Debandi et al. [2018]	−	−	✓	−
Herder et al. [2018]	✓	−	−	−
Kawamura and Otsuki [2018]	✓	✓	−	−
Kimura and Rekimoto [2018]	−	−	✓	✓
Vinayagamoorthy et al. [2018]	✓	−	✓	−
Zhang and Zhu [2018]	✓	−	−	✓
Zimmer et al. [2018]	✓	−	✓	−

Chapter 3

Evaluating an ARTV Design Space

3.0 Chapter overview

3.0.1 Thesis context

The previous chapter identified six design dimensions commonly used in ARTV work, via a systematic literature review. This chapter expands the findings of the previous chapter; it synthesises an ARTV design space, uses a "cheat sheet" to present design dimensions as questions, and subsequently evaluates the intelligibility and usefulness of the cheat sheet in two scenario writing tasks. The findings suggest the usefulness of the cheat sheet in conceptualising novel ARTV scenarios. In light of the findings, the ARTV design space and the cheat sheet are updated.

The main content of this chapter is adapted from: Pejman Saeghe, Mark McGill, Bruce Weir, Sarah Clinch, and Robert Stevens. Evaluating and updating a design space for augmented reality television. In *ACM International Conference on Interactive Media Experiences*, IMX '22, page 79–94, New York, NY, USA, 2022a. Association for Computing Machinery. ISBN 9781450392129. doi: 10.1145/3505284.3529965. URL `https://doi.org/10.1145/3505284.3529965`.

3.0.2 Author's contributions

Pejman Saeghe designed the study, created the material (e.g., video presentations and handouts), recruited participants, carried out data collection, conducted the interviews, analysed the results, and wrote the paper. Dr Mark McGill coded the scenarios written by the participants, as a second coder, and provided critical input during scenario

analysis. Prof Robert Stevens, Dr Bruce Weir, and Dr Sarah Clinch provided continuous support and critical feedback throughout all stages of the study.

3.0.3 Published 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 from a 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.

3.1 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. [2020] and Vatavu et al. [2020] 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., Kimura and Rekimoto [2018]); delivering story-related holographic artefacts for entertainment (e.g., Saeghe et al. [2019a]) and education (e.g., von der Au et al. [2020]); delivering additional virtual TV screens and providing advanced remote-control functionality for TV content (e.g., Baillard et al. [2017]), and even replacing the 2-D TV screen completely by delivering content in an unframed way (e.g., Zimmer et al. [2018]).

A potential problem, however, with the current state of research in this field 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. [2019] who used a co-creation approach to develop two future TV scenarios, and Popovici and Vatavu [2019a] and Popovici et al. [2021] who used a list of twenty sentence-length ARTV scenarios to elicit viewers' preferences regarding ARTV with the European and Chinese audiences, resectively.

In this chapter, we use six ARTV design dimensions previously identified by Saeghe et al. [2020] 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 A.1.3), and investigate the following research questions (RQ):

RQ 1: To what extent can a design space presented in a textual cheat sheet format be easily understood?

RQ 2: To what extent can a design space cheat sheet (presented in text as questions) 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 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 stimulating thought to conceptualise 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.

---

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

2) An expanded ARTV design space, incorporating novel ARTV experiences as conceptualised by participants with wide-ranging expertise using the cheat sheet.

3.2 Related work

We provide an overview of prior research that has combined AR with TV, and papers that looked at generating scenarios for future TV experiences. We also provide a brief description of Saeghe et al.'s [2020] design space, which was used as a basis for the ARTV design space cheat sheet used in the current work.

3.2.1 Augmented Reality Television

Saeghe et al. [2020] 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., Kawakita and Nakagawa [2014]), using AR to deliver virtual TV screens around a TV set (e.g., Baillard et al. [2017]), using AR to provide context for a TV set akin to a *focus + context* metaphor [Cockburn et al., 2009] (e.g., Kimura and Rekimoto [2018]), or using AR to replace the TV set and deliver content that appeared to be present in the living room (e.g., Zimmer et al. [2018]).

Vatavu et al. [2020] conceptualised *ARTV experiences for the living room* by expanding Milgram and Kishino's [1994] *reality-virtuality continuum* to two dimensions; where TV and the world were each positioned on a separate, and perpendicular, reality-virtuality continua. 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., Nixon et al. [2015]), 3) Physical world/physical TV with off-TV augmentation (e.g., Kawakita and Nakagawa [2014]), 4) Physical world/virtual TV (e.g., Vatavu [2012]), 5) Augmented world/physical TV (e.g., Jones et al. [2013]), 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.

3.2.2   ARTV scenarios

Geerts et al. [2019] 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" [Geerts et al., 2019].

To elicit potential viewers' expectation of ARTV in a large scale, Popovici and Vatavu [2019a] 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. [2021] conducted a similar survey, this time with $n = 147$ Chinese participants to investigated 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 (cf. Popovici et al. [2021], Table 7).

3.2.3 ARTV design space

In this chapter, we used the concepts presented by Saeghe et al. [2020] as a basis for creating a design space cheat sheet; henceforth referred to as *the ARTV design space cheat sheet*.

Saeghe et al. [2020] 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., Ghellal and Lindt [2008]), or both AR and TV content are required for an ARTV to make sense (e.g., Revelle et al. [2015]), or where either AR or TV play an additional role to an otherwise already *complete* experience (e.g., Baillard et al. [2017]).

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., Vatavu [2012]); or to interact with content in a game-like manner (e.g., Revelle et al. [2015])

3) Time: describing the relationship between the timelines of AR content and TV content, where AR and TV are intentionally not synchronised (e.g., Ghellal and Lindt [2008]), where AR and TV are synchronised and presented together (e.g., Kawakita and Nakagawa [2014]), or AR and TV are synchronised but presented intermittently (e.g., Revelle et al. [2015])

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., Denia et al. [2011]), or on separate devices (e.g., Kawakita and Nakagawa [2014])

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 influencing the sequence (e.g., Bibiloni et al. [2015]), or the camera angle (e.g., Han et al. [2010]).

3.3 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).

This section provides the following information:

1) A brief overview of the the ARTV design space cheat sheet (Section 3.3.1).

2) An overview of the procedure (Section 3.3.2)

3) An overview of participants' demographics and experimental conditions (Section 3.3.3)

3.3.1   ARTV design space cheat sheet

We based our ARTV design space cheat sheet on Saeghe et al. [2020]'s ARTV design space dimensions. We further developed the dimensions 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 (see Appendix A.1.1).

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 (Appendix A.1.3).

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 Appendix A). The ARTV design space cheat sheet can be found in Appendix A.1.3.

3.3.2 Procedure

 Participants received the participant's information sheet, and completed a consent form and a basic demographic questionnaire.

Scenario writing tasks

There were two scenario writing tasks. The first task was identical for all participants, regardless of their assigned group; it consisted of watching a YouTube video, and writing an ARTV scenario. For this task, participants were sent a link to a YouTube video, and two PDF documents: 1) the video transcript, 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. The handout presented the same material in text (see Appendix A.1.1). Participants were given 48 hours to complete the first task.

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 video transcript, 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 their second task (Appendix A.1.2). 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 sent at this stage to these participants was the ARTV design space cheat sheet (Appendix A.1.3). All participants were given 48 hours to complete the second task.

Semi-structured interview

The third 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 to them 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.

Participants received a £10 Amazon gift voucher on completion of the study.

Table 3.1: Participants' role in the media industry. Selection of more than one option was permitted.

Role	Consumer	Researcher	Enthusiast	Engineer	Producer	Technologist	Not disclosed	Self-described
Total	7	4	4	3	3	2	1	Designer

Table 3.2: Devices typically used by participants to watch TV content. Selection of more than one device was allowed.

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

3.3.3   Participants and experimental conditions

Adult participants were recruited using social media (LinkedIn and Twitter) and electronic mailing lists. Our inclusion criteria for participants was threefold; a participant had to be a researcher/content creator of immersive media (e.g., AR and VR), they had to be over eighteen years old, and have good written English.

Ten individuals (8 male, 2 female) opted to participate. 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.

Experiences with TV and media

Table 3.1 presents participants' role in the media industry. Selection of more than one role was allowed. Seven participants selected more than one 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*.

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. Second most popular device was *mobile phone*, with five participants selecting it. Eight participants reported using more than one device for TV consumption (see Table 3.2).

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 (see Table 3.3).

Three options were provided in the demographic questionnaire: 1) I don't know

Table 3.3: Participants' role in the AR industry. Selection of more than one option was permitted.

Role	Enthusiast	Consumer	Researcher	Producer	Technologist	Engineer	Writer
Total	7	6	5	4	2	1	1

what AR is, 2) I could probably describe AR, 3) I could describe AR confidently. All participants selected the last option, namely, "I could describe AR confidently".

All participants had used AR before. Six reported recent regular use/develop experience. One participant reported former regular use/development experience. Three reported infrequent use and a lack of development experience. Four of participants 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.

Experimental conditions

The study had two groups: a study group, and a control group. Participants were assigned to a group randomly. Each participant wrote two ARTV scenarios. The control group was added 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 (see Appendix A.1.1).

2) **Study 2:** five scenarios written by the study group participants after exposure to the ARTV design space cheat sheet (see Appendix A.1.3).

3) **Control 1**: five scenarios written by the control group participants after exposure to a set of basic definitions (see Appendix A.1.1).

4) **Control 2:** five scenarios written by the control group participants after exposure to the same set of basic definitions as before (see Appendix A.1.2).

3.4 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.

On average, from scenario 1 to scenario 2, participants in the study group increased their word-count by 20% (from 319.2 to 381.8 words per scenario), while participants in the control group decreased their word-count by 20% (from 298.2 to 237.4 words per scenario).

3.4.1   Qualitative analysis

We used a *deductive* content analysis technique [Mayring, 2014] to analyse the scenarios. Two researchers independently coded the scenarios and resolved disagreements through discussion. The initial code-book consisted of 42 categories extracted from the ARTV design space cheat sheet (Appendix A.1.3). The coding process consisted of three iterations, during the first two iterations, a further ten categories were identified and added to the code-book. The rest of this section details the outcomes for each of the three iterations, and discusses inter-rater agreements.

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.

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

Table 3.4: Descriptive statistics regarding scenarios' word-count. Mean, median, min, max, range, and sum are presented for task 1, task 2, and overall. Sub-columns provide the statistics for the study group (stdy), control group (ctrl), and study and control groups combined (all).

	Task 1			Task 2			Overall		
	Stdy	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**
Mean	319.2	298.2	308.7	381.8	237.4	309.6	350.5	267.8	309.15
Median	299	298	298.5	267	205	260	283	251.5	282.5
Min	94	130	94	229	153	153	94	130	94
Max	612	491	612	850	335	850	850	491	850
Range	518	361	518	621	182	697	756	361	756
Sum	1596	1491	3087	1909	1187	3096	3505	2678	6183

was 0.44. Differences were once again resolved through discussion.

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* [McHugh, 2012]. However, given the shared understanding of the concepts that was developed and articulated during the coding process, we expected the inter-rated agreement to be higher. We hypothesised that the sheer number of codes (52) may have caused many instances to be missed by either of the researchers rather than 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* [McHugh, 2012].

3.4.2 Codes' usage

Table 3.5 presents the top ten most widely used codes in the participants' scenarios. The mostly 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 dimensions were not used at all; we label these *high-*, *medium-*, and *low-*usage codes, respectively.

Table 3.5: The top ten most used codes and the number of scenarios in which they appeared. Numbers in study (stdy) 1 and 2, and control (ctrl) 1 and 2 are out of five. Numbers in tasks 1 and 2, and stdy and ctrl are out of 10. The numbers in the all column is out of 20.

Code	Task 1			Task 2			Overall		
	Stdy	Ctrl	All	Stdy	Ctrl	All	Stdy	Ctrl	All
Content - TV	5	3	8	4	5	9	4	8	17
AR-TV Dependency - additional	5	3	8	4	3	7	4	6	15
People - single primary viewer	4	4	8	2	4	6	2	8	14
ARTV genre	2	4	6	3	4	7	3	8	13
AR and TV content	4	4	8	1	3	4	1	7	12
Time - sync - continuous	2	4	6	4	2	6	4	6	12
Visual display - unframed	3	3	6	3	2	5	3	5	11
People - multiple - concurrent	2	2	4	4	2	6	4	4	10
Space - indoors	2	2	4	3	3	6	3	5	10
Content - AR	3	2	5	2	2	4	2	4	9
People - multiple primary viewers	2	2	4	4	1	5	4	3	9
Space - private	2	2	4	3	2	5	3	4	9

Table 3.6: 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	Task 1			Task 2			Overall		
	Stdy	Ctrl	All	Stdy	Ctrl	All	Stdy	Ctrl	All
Content - TV (✓)	5	3	8	4	5	9	9	8	17
AR-TV dependency - additional (✓)	5	3	8	4	3	7	9	6	15
People - single primary viewer (✓)	4	4	8	2	4	6	6	8	14
ARTV genre	2	4	6	3	4	7	5	8	13
AR and TV content (✓)	4	4	8	1	3	4	5	7	12
Time - sync - continuous (✓)	2	4	6	4	2	6	6	6	12
Display - unframed (✓)	3	3	6	3	2	5	6	5	11
People - multiple - concurrent (✓)	2	2	4	4	2	6	6	4	10
Space - indoors (✓)	2	2	4	3	3	6	5	5	10
Content - AR (✓)	3	2	5	2	2	4	5	4	9
People - multiple primary viewers (✓)	2	2	4	4	1	5	6	3	9
Space - private (✓)	2	2	4	3	2	5	5	4	9
Broadcast mode	2	2	4	2	2	4	4	4	8
Social group - between-people interaction (✓)	1	1	2	3	2	5	4	3	7
People - multiple - at-a-distance (✓)	1	1	2	3	2	5	4	3	7
People_space - familiar with space (✓)	2	1	3	1	3	4	3	4	7
Modality - auditory	1	2	3	2	2	4	3	4	7
Representation - photoreal	1	1	2	2	2	4	3	3	6
People_objects - familiarity with objects (✓)	2	1	3	1	1	2	3	2	5
Display - extended frame (✓)	1	1	2	2	1	3	3	2	5
Social group - friends and family (✓)	1	1	2	1	1	2	2	2	4

Table 3.7: Twenty-six medium-usage codes appear in at least one of the 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	Task 1			Task 2			Overall		
	Stdy	Ctrl	All	Stdy	Ctrl	All	Stdy	Ctrl	All
Interaction - display (✓)	3	1	4	4	0	4	7	1	8
Interaction - game (✓)	0	1	1	1	4	5	1	5	6
People - multiple - co-located (✓)	2	1	3	2	0	2	4	1	5
Space - purposeful augmentation (✓)	1	0	1	1	3	4	2	3	5
People_space - movement (✓)	0	0	0	1	3	4	1	3	4
Interaction - editorial control	1	0	1	2	1	3	3	1	4
Display - multiple frames (✓)	1	2	3	1	0	1	2	2	4
Space - content modification/generation (✓)	0	0	0	2	2	4	2	2	4
Objects - purposeful augmentation (✓)	2	1	3	1	0	1	3	1	4
Objects - content modification/generation (✓)	1	0	1	2	1	3	3	1	4
AR delivering TV content (✓)	0	1	1	1	1	2	1	2	3
Demographics	1	1	2	0	1	1	1	2	3
Social group - strangers (✓)	0	0	0	3	0	3	3	0	3
People - bystanders (✓)	0	1	1	2	0	2	2	1	3
AR-TV dependency - independent (✓)	1	1	2	1	0	1	2	1	3
AR-TV dependency - dependent (✓)	0	1	1	0	2	2	0	3	3
Time - asynchronous (✓)	2	0	2	1	0	1	3	0	3
Time - sync - intermittent (✓)	1	1	2	0	1	1	1	2	3
Objects - physical integration (✓)	1	0	1	2	0	2	3	0	3
People_objects - influencing objects (✓)	1	0	1	1	0	1	2	0	2
Display - single frame (✓)	0	0	0	2	0	2	2	0	2
Representation - cartoonish	0	0	0	1	1	2	1	1	2
Modality - haptic	1	0	1	1	0	1	2	0	2
Display - 360	0	1	1	0	0	0	0	1	1
Modality - olfactory	0	0	0	0	1	1	0	1	1
Space - outdoors (✓)	0	0	0	0	1	1	0	1	1

High-usage codes:

Twenty-one (40%) of the codes appeared in all four experimental conditions (see Table 3.6). Seventeen (81%) originated from the ARTV design space cheat sheet—indicated by a (✓) in Table 3.6. The remaining four codes emerged during scenarios analysis (see Section 3.4.1).

Medium-usage codes:

Twenty-six (50%) of the codes appeared in at least one of the four experimental conditions (see Table 3.7). Twenty of these codes (77%) originated from the ARTV design space cheat sheet—indicated by a (✓) in Table 3.7. The remaining six codes emerged during scenario analysis (see Section 3.4.1).

Table 3.8: Five low-usage codes appear in none of the four scenario groups.

Code	Task 1			Task 2			Overall		
	Stdy	Ctrl	All	Stdy	Ctrl	All	Stdy	Ctrl	All
Space - public	0	0	0	0	0	0	0	0	0
People - passers-by	0	0	0	0	0	0	0	0	0
People - multiple - non-concurrent	0	0	0	0	0	0	0	0	0
People_space - influencing space	0	0	0	0	0	0	0	0	0
Interaction - story	0	0	0	0	0	0	0	0	0

Low-usage codes:

Five codes did not appear in any of the four experimental conditions (see Table 3.8). Low-usage codes comprise 10% of the total codes.

Table 3.9: 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

3.4.3 Implications

Table 3.9 presents our initial classification scheme for ARTV design space dimensions. We used this scheme to create the ARTV design space cheat (Appendix A.1.3). Furthermore, the structure and wording of Table 3.9 influenced our code-book during the qualitative analysis of the scenarios (Section 3.4.1).

Based on our analysis of participant scenarios, we propose a revision of Saeghe et al. [2020]'s design space that incorporates new dimensions identified during coding, alters the grouping of existing 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.

Content:

Twenty-two (42%) of the codes describe various aspects of content in the context of an ARTV experience. Table 3.10 presents *content*-related codes, ranked according to their usage.

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 (Table 3.5) in the study.

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.

Table 3.10: Twenty-two content-related codes. The codes above the horizontal dashed line were high-usage and the codes below it were medium usage.

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

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" [P8S1S2], "[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." [P8S1S2]

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's screen." [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]

And 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 time-lines—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 emerged during our qualitative analysis of the scenarios (see Section 3.4.1); 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: *Modality* emerged during our qualitative analysis of the scenarios (see Section 3.4.1). The *auditory* aspect was used in seven (35%) scenarios, while the *haptic*

and *olfactory* aspects were used 10% and 5%, respectively (two and one, respectively).

Examples of the auditory modality being included into an ARTV experience 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 being included into an ARTV experience include: "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: *Representation* emerged during our qualitative analysis of the scenarios (see Section 3.4.1). 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 raised in one scenario to overcome a *privacy* concern when a participant 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]

People:

Twenty-one (40%) codes capture various aspects relating to people, in the context of an ARTV experience.

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 (Table 3.5).

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 (35% and 25% of the total scenarios, respectively). No scenario considered

Table 3.11: Twenty-one people-related codes. The two horizontal dashed lines separate (top to bottom) high-, medium-, and low-usage codes.

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

non-concurrent viewing; this aspect can arise in scenarios where viewers are at-a-distance or when content is available on-demand.

Interaction: The most widely used type of interaction in the scenarios 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]

Table 3.12: Five space-related codes. The two horizontal dashed lines separate (top to bottom) high-, medium-, and low-usage codes.

	Task 1			Task 2			Overall		
Code	**Stdy**	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**
Indoors	2	2	4	3	3	6	5	5	10
Private	2	2	4	3	2	5	5	4	9
Purposeful augmentation	1	0	1	1	3	4	2	3	5
Content modification/generation	0	0	0	2	2	4	2	2	4
Outdoors	0	0	0	0	1	1	0	1	1
Public	0	0	0	0	0	0	0	0	0

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]

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.

Space:

Six (12%) codes describe various aspects regarding *space*. 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 (Table 3.5) 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]

Objects:

Three (6%) codes describe various aspects regarding *objects*. In the context of an ARTV experience, objects consist of both *physical (real)* and *holographic (or otherwise virtual)* objects that are present in space. Objects are, at least initially, not part of the ARTV experience, but may become integrated into it. Table 3.13 presents object-related codes, ranked according to their usage.

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.

Table 3.13: Three object-related codes appeared in at least one of the four scenario groups.

Code	Task 1			Task 2			Overall		
	Stdy	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**	**Stdy**	**Ctrl**	**All**
Purposeful augmentation	2	1	3	1	0	1	3	1	4
Content modification/generation	1	0	1	2	1	3	3	1	4
Physical integration	1	0	1	2	0	2	3	0	3

3.5 Interviews

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

3.5.1 Qualitative analysis

We used an *inductive* content analysis technique [Mayring, 2014] 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.

3.5.2 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 3.4.3.

ARTV design space cheat sheet usefulness: Participants found the ARTV design space cheat sheet to be 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 was 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 [about] 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 (see Section 3.4.3), 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]

3.5.3 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 (see Appendix A.1.3). 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 (see Appendix A.1.3). The typical requirements associated with AR technology, namely, 3-D, registered in real space, and interactive [Azuma, 1997, Azuma et al., 2001], 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* (see Appendix A.1.3). The only comment in this section was a suggestion from P9 to add a "social element" to our definition.

3.5.4   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 (see Appendix A.1.3). 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 structuring 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 the 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]

3.5.5   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."

3.5.6   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 (see Section 3.4.3) and during the interviews: "Is it on demand? Is it live?" [P1]

Privacy: In Sections 3.4.3 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 viewer's 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]

3.6 Discussion

We provide four discussion points regarding word-count, typical ARTV conceptualisations, expansion and modification of the ARTV design space, and the benefits of using a check-list cheat sheet to present the design space.

3.6.1 Word-count

Although the number of participants ($n = 10$) and the total number of scenarios (twenty) are small, there was a 20% increase in average word-count in the study group, from scenario 1 to scenario 2, which is suggestive of an effect from the ARTV design space cheat sheet; particularly in light of a 20% decrease in average word-count, from scenario 1 to scenario 2, in the control group.

The decrease in word-count from scenario 1 to scenario 2 in the control group may have been caused by a loss of interest; since tasks 1 and 2 were identical, this may have been perceived as repetitive, causing participants in the control group to be less engaged with the second scenario writing task. The increase in word-count from scenario 1 to scenario 2 in the study group may suggest an increase in engagement with the second task, due to exposure to an outline of a design space, which they could then explore.

The same reasoning may, to some extent, explain the changes in minimum and maximum scenario lengths, between two groups; for the study group participants, from scenario 1 to scenario 2, the minimum and maximum scenario lengths increased by 144% (from 94 to 229 words per scenario) and 39% (from 612 to 850 words per scenario), respectively. In comparison, for the control group participants, from scenario 1 to scenario 2, the minimum and maximum scenario lengths increased by 18% (from 130 to 153 words per scenario) and decreased by 32% (from 491 to 335 words per scenario), respectively.

3.6.2 Typical ARTV conceptualisations

Section 3.4.2 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 synced with TV content and it was not presented simultaneously as TV content.

Low-usage   There were no ARTV conceptualisations consisting of an ARTV experience in a public space, in the presence of passers-by. 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.

3.6.3   Expansion and modification of the design space

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 Hughes and Montagud [2021]). Using the expertise and the findings from such closely related fields can result in the creation of richer 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), 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 experiences.

In case of editorial control, Saeghe et al. [2020] considered it as 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."

3.6.4 Cheat sheet as a checklist

Participants found the cheat sheet to be useful for conceptualising novel ARTV scenarios. 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.

We propose for future work to expand the ARTV design space, presented in this paper, by iteratively identifying other potentially relevant design dimensions pertaining to an ARTV experience. Furthermore, we suggest that the updated ARTV design space cheat sheet (Appendix A.1.4) can be used by practitioners and designers in the conceptualising and brainstorming phase of their content creation. For instance, many versions of the same base ARTV concept can be easily developed by changing an aspect of one or more dimensions, or by adding/removing specific dimensions.

Overall, we observe that the ARTV design space cheat sheet has a qualitatively positive effect on conceptualising novel ARTV scenarios.

3.7   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 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. 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.

Chapter 4

Conceptualising ARTV for the Living Room

4.0 Chapter overview

4.0.1 Thesis context

Chapter 2 identified six themes in ARTV work, with enhancing a conventional TV experience as the most popular research theme. The current chapter's focus is on experiences that use AR to enhance or transform a conventional TV viewing experience. Since a conventional TV viewing experience is typically associated with the living room, the chapter creates and promote a conceptual framework for ARTV experience in the living room.

The main content of this chapter is adapted from: Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. Conceptualizing Augmented Reality Television for the Living Room. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 1–12, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393660. URL https://doi.org/10.1145/3391614.3393660.

4.0.2   Author's contributions

Pejman Saeghe conducted the literature survey, classified the documents (cf. Table 4.1), and wrote Section 4.4 and parts of Sections 4.1 and 4.6. Radu-Daniel Vatavu wrote Section 4.3 and parts of Sections 4.1 and 4.6. Teresa Chambel and Vinoba Vinayagamoorthy wrote parts of Section 4.5. All authors provided critical feedback and edits.

4.0.3   Published abstract

We examine the concept and characteristics of "Augmented Reality Television" (ARTV) using a four-step investigation method consisting of 1) an analysis of commonly-accepted perspectives on Augmented and Mixed Reality systems, 2) a literature survey of previous work on ARTV, 3) relevant connections with other areas of scientific investigation from TVX/IMX, such as Ambient Media, Interactive TV, and 3-D TV, and 4) by proposing a conceptual framework for ARTV called the "Augmented Reality Television Continuum." Our work comes at a moment when the excitement and hype about the potential of AR for home entertainment has overlooked rigorous analysis and clear-cut examinations of the concepts involved, which should be the hallmark of any exact science. With this work, our goal is to draw the community's attention toward fundamentals and first principles of ARTV and to tease out its salient qualities on solid foundations.

4.1   Introduction

New forms of computer-generated content, media, and platforms [Paradiso and Landay, 2009, Mann, 1999, Azuma, 2019, Jenkins, 2019], anchored into and aligned with the physical world, continue to change how we perceive and interact with our surroundings. As display and sensing technology evolve, presentation of content that is photorealistic, adaptive, personalized and customizable, real-time responsive and interactive and, consequently, engaging for users becomes attainable. These developments impact how we experience various representations of reality, such as virtual [Jenkins, 2019], mediated [Mann, 1999], multimediated [Mann et al., 2018], amplified [Falk et al., 1999], alternate [Chambel et al., 2016], augmented [Azuma, 1997, 2019], augmediated [Mann et al., 2018], mixed [Milgram and Kishino, 1994, Speicher et al., 2019], blended [Waterworth and Hoshi, 2016], extended [Marr, 2019], and cross-reality [Paradiso and Landay, 2009].

For instance, applied to television, Augmented Reality (AR) can immerse viewers into an interactive storytelling space that enables fantasy worlds to "break out" of the confines of the physical TV frame as well as to "bring in" aspects of the physical world. As part of the televised show, the action can extend to the living room [Vatavu, 2013, Jones et al., 2013, 2014, Vatavu, 2012] to create a sense of "actually being there." Therefore, the combination of AR and television creates the premises for a hybrid medium that opens new horizons for storytelling and engagement with interactive media and digital content. For instance, imagine watching a live soccer game, where the formation of players in each team is symbolized on a miniaturized model of the game field that is rendered right in front of the TV screen, while statistics of the match are always visible right next to the TV set; see TWiT Tech Podcast Network [2018] (minute 02:05). Or, consider extending the field of view of the conventional TV set with synchronized video projections on the wall behind it [Jones et al., 2013, Kimura and Rekimoto, 2018], or putting up on the wall as many virtual TV screens as one wishes, and be able to adjust their location and size to match the architectural design of the room [Vatavu, 2012] with flexible control over which content is rendered where [Vatavu, 2013]. Moreover, imagine a detective TV series, where indispensable parts of the story are told through film with established editing conventions, but before the end of each episode, the viewers' living room is transformed into the crime scene, and viewers can continue to experience the plot of the story at a new level by searching for clues with their AR-enabled smart devices, before the final resolution of the episode.

By putting television and AR together in the form of ARTV (*read*: "Augmented Reality Television"), experiential rich scenarios can become the norm of living room TV-based entertainment [Popovici and Vatavu, 2019c,a, Geerts et al., 2019]. However, while we are seeing more and more developments in AR for television presented and discussed at top-tier venues, such as CHI, UIST, TVX, and ISMAR [Jones et al., 2013, 2014, Popovici and Vatavu, 2019a, Vatavu, 2013, Vinayagamoorthy et al., 2019, Rostami et al., 2018, Saeghe et al., 2019a, Vinayagamoorthy et al., 2018, Popovici and Vatavu, 2019b,c], there is large heterogeneousness in the ARTV landscape as well as in the terminology used by researchers and practitioners, revealing concepts that are not crisply defined and fully understood. A more rigorous description of *what ARTV actually is* represents thus a requirement for our community to sustain growth in this area on solid foundations. A closer look at the larger context of AR research reveals the timely need for this requirement.

4.1.1   The context in AR research for ARTV

A key observation is that *it is not just ARTV and not just the TVX/IMX community that lack the desired level of rigorousness in properly specifying computer-generated augmented realities.* In fact, the terminology used in the scientific literature, industry, and media to refer to Augmented Reality (AR), Mixed Reality (MR), and Extended or Cross-Reality (XR) is split to the extent to which it has become difficult even for experts to define these concepts precisely and confidently [Speicher et al., 2019].

There are several articles and blogs attempting to clarify the differences between AR, MR, VR, XR, and other "R" acronyms for the layman reader [Goode, 2019, Irvine, 2017, Marr, 2019, Scribani, 2019]. However, major industrial players refer to comparable technologies using different terms. For example, Google has adopted the "Augmented Reality" terminology when addressing its community of developers, speaking about "augmented reality experiences" enabled by the ARCore platform[1]; Microsoft promotes the term "Mixed Reality" with the Windows Mixed Reality platform and the HoloLens headset[2]; while other industry players, such as Qualcomm, speak directly about XR and envision "the convergence of the smartphone, mobile VR headset, and AR glasses into a single XR wearable"[3]. Although a few attempts have been made to clarify the terminology[4] and to demystify the VR landscape[5], empirical evidence suggests that the interpretation of relevant terminology by experts in the field remains varied [Speicher et al., 2019].

4.1.2   The responsibility of the TVX/IMX community

As the "R" realities are not crisply defined, the term television itself is starting to become a troublesome one, in need for an updated definition. In this context, ARTV must be properly characterized in a space where television meets computer-generated realities. Exactly because TV and "R" realities are converging, it is our responsibility as a community to provide an informed answer for what is the new emerging concept of ARTV to keep up with the expectations horizon created by a fast developing industry, a creative academia, and an excited public media [TWiT Tech Podcast Network, 2018, Scribani, 2019, TIME, 2017, Somolinos, 2019].

---

[1]https://developers.google.com/ar/develop
[2]https://www.microsoft.com/en-gb/windows/windows-mixed-reality
[3]https://www.qualcomm.com/invention/extended-reality
[4]https://unity3d.com/what-is-xr-glossary
[5]https://www.intel.co.uk/content/www/uk/en/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html

Figure 4.1: Our four-step method for specifying Augmented Reality for Television. Each step addresses a specific question and is covered by a distinct section of this paper, e.g., Section 6 discusses *What is a conceptual framework for ARTV?*

Overview current definitions and perspectives on AR; understand research and technical challenges for the AR field.	Scrutinize previous work implementing AR systems, prototypes, and interaction techniques for home television environments.	Exploit connections with related concepts from immersive media, multimedia alternate realities, 3-D TV, and ambient media.	Identify key characteristics for Augmented Reality for television and outline a conceptual framework for ARTV.
1 **Foundations**	**2** **ARTV**	**3** **Connected areas**	**4** **Framework**
What constitutes the foundation for ARTV?	*What has already been achieved in ARTV?*	*What makes ARTV a specific kind of experience?*	*What is a conceptual framework for ARTV?*
(Section 3)	(Section 4)	(Section 5)	(Section 6)

To understand the hype and context in which our contribution for specifying ARTV is needed, we provide a few examples. In a January 2019 Visual Capitalist article piece, Scribani [2019] noted that "XR brings immersive experiences to the entertainment world, and offers consumers an opportunity to virtually experience live music and sporting events from the comfort of their VR headset"; on their web page dedicated to XR, Qualcomm talks about how "XR could replace all the other screens in your life, like that big TV in your living room"[6]; and, during Facebook's F8 Developer Conference of 2017, Mark Zuckerberg touched on augmented television, among other examples of how AR technology could change users' lives: "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"; see the video recording of the event at TIME [2017] (minute 4:35).

It seems though that excitement and hype about AR, MR, and XR technology for home entertainment, television included, has largely overpassed rigorous examination and understanding of the intrinsic concepts of Augmented Reality and television. However, to advance on scientific grounds, we need rigorous conceptual formalization of what ARTV is. This paper is an attempt toward such a formalization.

4.1.3 Contributions of this work

Our practical contributions are as follows:

1) We conduct the first literature survey on AR for television in order to understand past efforts, critical ideas, and key projects. We examine various perspectives

---

[6]https://www.qualcomm.com/invention/extended-reality

on AR/MR, from which we extract key characteristics for ARTV and identify areas of investigation from TVX/IMX related to ARTV, e.g., Ambient Media, Interactive TV, and 3-D TV.

2) We introduce a conceptual framework for ARTV in the form of the "Augmented Reality Television Continuum," a 2-D representation space for ARTV concepts, devices, systems, and applications inspired by the 1-D Reality-Virtuality Continuum of Milgram and Kishino [1994]. We also differentiate ARTV from AVTV (*read*: "Augmented Virtuality Television"), a complementary concept equally covered by our ARTV conceptual framework.

4.2   The investigation method of this paper

Before we move on, we take a moment to describe the investigation method that we implement in this work to identify key characteristics of ARTV. Our method is composed of four steps, as follows (see Figure 4.2 for a visual illustration):

Step 1: We start from the perspectives and definitions of general-purpose AR and MR systems, examine current classifications, and overview research and technical challenges. This step enables us to provide an answer to *What are the foundations for ARTV?*

Step 2: We perform a targeted literature survey of previous work implementing AR for television. This step enables us to form an understanding of *What has already been achieved in ARTV?*

Step 3: Having established the foundations and understood the state-of-the-art, we proceed to identifying relevant connections between ARTV and other areas of investigation from TVX/IMX, such as Multimedia Alternate Realities, Ambient Media, Interactive TV, and 3-D TV. This step enables us to understand *What makes ARTV a specific kind of television experience?*

Step 4: We identify key components for ARTV, which we build into our new conceptual framework, the "Augmented Reality Television Continuum."

4.3 Foundations: augmented, mixed, and virtual realities

We overview in this section first principles of VR, AR, and MR systems and environments. While VR is commonly understood as fully-immersive environments that substitute real-world sensations with simulated cues [Scribani, 2019, Milgram and Kishino, 1994, Mann et al., 2018], the distinction between AR and MR has been less obvious [Speicher et al., 2019].

The transition from fully-immersive VR toward displays that combine elements from both the physical (real) world and the virtual (computer-generated) world was represented by the first prototypes of AR systems starting with Sutherland's [1968] head-mounted 3-D display; see Azuma's [1997] widely referenced survey of the state-of-the-art in AR in 1997 and Billinghurst et al.'s [2015] comprehensive overview of the field photographed in 2015. However, since the introduction of the concept, AR has received many definitions in the scientific literature corresponding to various perspectives, which are relevant to scrutinize for the purpose of our investigation. This examination of the foundations of computer-generated realities is key to draw implications for ARTV. We start with Milgram and Kishino's [1994] highly-influencing "Reality-Virtuality Continuum" that distinguished MR from AR.

4.3.1 The Reality-Virtuality Continuum

In 1994, Milgram and Kishino [1994] introduced the "Virtuality" continuum—an imaginary line having the real and virtual worlds at its opposite ends. Later, Milgram et al. [1995] and Milgram and Jr. [1999] referred to this line as the "Reality-Virtuality (RV) Continuum"—the name in use today. As one moves along the RV Continuum, the degree of interpolation between the real and the virtual changes, leading to Augmented Reality (AR) and Augmented Virtuality (AV) world mixtures; see Figure 4.2. In fact,

Figure 4.2: Illustration of Milgram and Kishino's [1994] RV Continuum, which we reproduce in this work since it represents the foundation for our ARTV Continuum from Figure 4.4.

Figure 4.3: Evolving perspectives of Augmented Reality (left, shown in chronological order) and implications for ARTV (right).

Definition of AR	Perspective	Impacted by and/or impacts on	Implication for ARTV
2019: Interface and gateway	Integration	Cloud Computing: SaaS, PaaS, IaaS, XaaS & ARCloud	Specific gateway and integration of TV genres
2017: Immersive experience, illusion	Experiential	Human-Computer Interaction	Specific type of immersion
2016: A new form of media	Content consumption	Social & Interactive Media	Specific form of media
1997: Combines the real and the virtual, interactive in real-time, and registered in 3-D	Technology independence	Applied AR	3-D registered and interactive in real time
1997: A specific form of intelligence amplification	Task assistance	Human-Computer Interaction	Specific services
1997: A variation of Virtual Reality	Computer-generated realities	Computer Graphics & Virtual Reality	Anchored in the physical, primary substratum of the living room environment
1994: The left-most part of the RV continuum			
1968: Information surrounding users in 3-D, enabled by see-through head-mounted displays	Virtual Reality	Computer Graphics	Informative and contextual

the primary environment or "substratum" that is augmented determines the distinction between AR and AV. Everything in the RV Continuum, except its ends, was defined as Mixed Reality (MR), a "more encompassing term to supplement the existing definition of Augmented Reality (AR), which leads us to propose definitions of the associated concepts of Augmented Virtuality (AV) and then Mixed Reality (MR)" [Milgram and Jr., 1999].

According to the RV Continuum, MR encompasses AR, but excludes VR: "the most straightforward way to view a Mixed Reality environment, therefore, is one in which real world and virtual world objects are presented together within a single display, that is, anywhere between the extrema of the virtuality continuum" [Milgram and Kishino, 1994]. However, MR and AR have received other connotations in recent years. For example, in a 2019 study, Speicher et al. [2019] reported six definitions for MR from interviews conducted with AR/MR/VR experts from academia and industry as well as from a literature survey. They concluded that "MR can be many things and its understanding is always based on one's context... there is no single definition of MR and it is highly unrealistic to expect one to appear in the future." The authors also recognized that "it is extremely important to be clear and consistent in terminology and communicate one's understanding of MR in order to avoid confusion and ensure constructive discussion" [Speicher et al., 2019]. Milgram and Kishino's [1994] definition of MR based on the RV Continuum was the most frequent perspective found in Speicher et al. [2019]. Thus, we also adopt this perspective in our work.

4.3.2 Augmented Reality *vs.* Augmented Virtuality

Milgram and Kishino's [1994] formalization of the RV Continuum unveiled the concept of AV, where the virtual world is the primary environment that is augmented. By analogy, we differentiate between Augmented Reality TV (ARTV) and Augmented Virtuality TV (AVTV). In the former case, it is the real world (e.g., the living room) that is augmented with virtual objects; in the later, it is a virtual world that is augmented with real objects, e.g., a video feed of a physical TV set. In this work, we are interested in ARTV and, thus, we continue our discussion with an overview of definitions for AR. Section 6 resumes AVTV for the interested reader.

4.3.3 Perspectives on AR and implications for ARTV

Several definitions have been proposed for AR in the scientific literature and it is essential for the purpose of this work to scrutinize them in order to understand the foundations of AR and to extract implications for ARTV. In the following, we present a chronological examination of evolving perspectives on AR, from which we identify key concepts, implications for ARTV, and connected areas of scientific investigation from TVX/IMX; see Figure 4.3 for an overview.

1968: AR is 3-D information surrounding the user, enabled by wearing see-through HMDs

Sutherland's [1968] head-mounted display (HMD) represents the first instance of an AR device and system. Although there is no mention of the term "Augmented Reality" in Sutherland's [1968] paper, one quote characterizes well the ultimate goal in his work: "our objective in this project has been to surround the user with displayed three-dimensional information". In this context, AR is visualization of 3-D spatial information enabled by a see-through HMD, a perspective that will dominate AR research for decades. For instance, in their 1994 paper introducing the concept of Mixed Reality and the RV Continuum, Milgram and Kishino [1994] remarked that the prominent use of the term AR was limited at that time to what they called "Class 3 type displays," which are HMDs with see-through capabilities. In another work, Milgram et al. [1995] presented a definition of AR as "a form of virtual reality where the participant's head-mounted display is transparent, allowing a clear view of the real world", which had originated from a call for papers of a representative journal special issue on AR. Pointing to the limitations of such definitions, Milgram and Kishino [1994] argued that the AR concept

can equally be extended to other types of displays, an observation that Milgram and Jr. [1999] and Milgram et al. [1995] resumed in their follow-up work on real and virtual world display integration.

> **Key AR concepts:** information, HMD, 3-D.
> **Implications for ARTV:** Augmented television renders relevant information in the 3-D space surrounding the viewer and, if physically present, the TV set, enabled by some dedicated visualization device.
> **Connected TVX/IMX area:** 3-D TV.

1994: The real environment is "augmented" by means of virtual (computer graphic) objects

This perspective was adopted by Milgram and Kishino [1994] as an operational definition for AR, because it was encompassing in terms of classes of displays compared to the previous, HMD-based approach. However, the generality of this definition generated terminology problems for Milgram and Kishino [1994] in relation to their Class 5 and Class 6 displays, i.e., completely graphic display environments, that made the authors ponder about the nature of the primary environment that is actually augmented. The result was the distinction between AR and AV and, eventually, the introduction of the term MR to cover all types of mixtures between physical and virtual.

> **Key AR concepts:** virtual objects, computer graphics, real environment.
> **Implications for ARTV:** Augmented television superimposes virtual content onto the real environment, which is the primary substratum that is augmented.
> **Connected TVX/IMX area:** 3-D TV.

1997: A variation of VR

According to this definition and perspective from Azuma [1997], "in contrast [to VR], AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it." Eight years later, in their book on spatial Augmented Reality (sAR), Bimber and Raskar [2005] resumed this definition to specify the characteristics of AR by contrasting it with VR: "The fact is that in contrast to traditional VR, in AR the real environment is not completely suppressed; instead it plays a dominant role.

Rather than immersing a person into a completely synthetic world, AR attempts to embed synthetic supplements into the real environment (or into a live video of the real environment)." In this antinomy, AR is meant to supplement the physical reality instead of completely replacing it [Azuma, 1997], while focusing on embedding synthetic supplements into the physical world [Bimber and Raskar, 2005]. Like Milgram and Kishino [1994], Milgram et al. [1995], and Milgram and Jr. [1999] before them, Bimber and Raskar [2005] recognized the need to go beyond the technology of eye-worn and HMD displays, and defined sAR as "new display paradigms [that] exploit large spatially-aligned optical elements, such as mirror beam combiners, transparent screens, or holograms, as well as video projectors."

> **Key AR concepts:** superimposed, composited, supplemented reality.
> **Implications for ARTV:** Augmented television supplements the content of the primary TV broadcast.
> **Connected TVX/IMX areas:** 3-D TV, Interactive TV (iTV), Ambient Media.

1997: A specific form of "intelligence amplification"

In his 1997 survey of AR systems, applications, and developments, Azuma [1997] provides two other perspectives on AR, besides the antinomy with VR. One of them, more general, capitalizes on the utility of computers to assist users in their tasks, which can be interpreted as access to a higher level of human intelligence. From this perspective, AR implements a specific case of intelligence amplification. The second perspective is more structured and systematic; see next.

> **Key AR concepts:** intelligence amplification, task assistance, the computer as a tool.
> **Implications for ARTV:** Augmented television assists viewers by providing new services.
> **Connected TVX/IMX areas:** Ambient Media, iTV.

1997: A system that combines the real and the virtual, is interactive in real time, and registered in 3-D

Azuma [1997] recognized the large influence of the HMD perspective on AR research and to avoid limiting the field to a specific form of technology, proposed three essential properties of any AR system. These properties constituted into a structured and

systematic definition of what AR is, which has been largely adopted since;[7] see, for example, Billinghurst et al.'s [2015] survey of AR that cites Azuma's [1997] three characteristics right from the first paragraph of their chapter defining AR.

> **Key AR concepts:** real-time interactivity, 3-D.
> **Implications for ARTV:** Augmented television is registered in 3-D and interactive in real time.
> **Connected TVX/IMX area:** iTV, 3-D TV.

2016: A new form of media

This perspective builds on a specific challenge that has been identified for AR systems. According to Azuma [2016], "The ultimate and most important challenge facing AR ... is experiential in nature: How will we establish Augmented Reality as a new form of media, enabling new types of experiences that differ from established media? If AR is to become ubiquitous in consumer usage, then we must ... [develop] new types of experiences that are engaging and compelling in different ways than traditional media such as books, movies, and even Virtual Reality." By adopting the experiential perspective, AR systems can be seen as implementers of a specific kind of (ambient) media with characteristics from both the real and the virtual world.

> **Key AR concepts:** new media, experience.
> **Implications for ARTV:** Augmented television is a new medium for a new digital media experience.
> **Connected TVX/IMX area:** Ambient Media.

2017: An immersive experience and an illusion

As a direct effect of superimposing virtual 3-D objects on top of the users' direct view of the physical world, AR generates the illusion of an immersive mix-world experience [Azuma, 2017].

> **Key AR concepts:** immersion, experience, illusion.
> **Implications for ARTV:** Augmented television is an immersive experience.
> **Connected TVX/IMX areas:** Multimedia Alternate Realities.

---

[7]Azuma [1997] has been cited over 10,000 times, `https://scholar.google.com/scholar?um=1&ie=UTF-8&lr&cites=17196017931627326366`

Table 4.1: Classification of ARTV-related papers according to their contributions, Milgram and Kishino's [1994] display classes and dimensions, such as Extent of World Knowledge and Presence Metaphor. *Note:* papers are listed in chronological order.

Reference	Contributions made					Extent of World Knowledge	Reproduction Fidelity	Extent of Presence Metaphor	Display class	Arguably AR
	New tech	Application	Method	User study	Design					
Stauder and Robert [2002]	✓	✓	✓	–	–	where	video / video	n.a.	1	n.a - XR - Imt
Vatavu [2012]	✓	✓	✓	✓	–	where	HD video / Projected video	Real time imaging	6	n.a. - XR - n.a
Jones et al. [2013]	✓	✓	–	✓	–	where	HD video / Real-time 3D animation	Real time imaging	6	P - XR/AltR - Imt/Img
Vatavu [2013]	✓	✓	✓	✓	✓	where	HD video / Projected video	Real time imaging	6	n.a. - XR - n.a.
Jones et al. [2014]	–	–	–	–	–	world completely modeled	HD video / Real-time 3D animation	Real time imaging	6	P - XR/AltR - Imt/Img
Gómez et al. [2014]	✓	✓	✓	✓	–	where + what	HD video / simple graphics	Monoscopic imaging	1	n.a. - XR - n.a.
Kawakita and Nakagawa [2014]	✓	✓	✓	–	–	where + what	HD video / Real-time 3D animation	Real time imaging	1	P – XR - Img
Revelle et al. [2015]	✓	✓	–	✓	–	where + what	HD video / Real-time 3D animation	Real time imaging	1	P - XR/HR - Img
Vatavu [2015]	✓	✓	✓	✓	✓	where	HD video / Visible surface imaging	Monoscopic imaging	1	n.a. - XR - n.a.
Baillard et al. [2017]	–	✓	–	–	–	where + what	HD video / Real-time 3D animation	Real time imaging	1 & 3	P - XR - Imt
Sotelo et al. [2017]	✓	✓	✓	–	–	n.a.	HD video / Real-time 3D animation	n.a.	1	n.a. - XR - Imt
Kawamura and Otsuki [2018]	✓	✓	✓	–	–	world completely modeled	HD video / Real-time 3D animation	Real time imaging	1	P - XR - Img
Kimura and Rekimoto [2018]	✓	✓	–	✓	–	world unmodeled	HD video / Generated context images	n.a.	5	n.a. - XR - Imt
Zimmer et al. [2018]	–	✓	✓	✓	✓	where	HD video / 3D animation	Monoscopic imaging	1	P - XR - Imt
Popovici and Vatavu [2019a]	–	–	–	✓	–	n.a.	n.a.	n.a.	n.a.	P - XR - Imt
Geerts et al. [2019]	–	–	✓	✓	–	where	n.a.	n.a.	6	P - XR - Imt
Saeghe et al. [2019a]	–	✓	–	✓	✓	where	HD video / Real-time 3D animation	Real time imaging	3	P - XR - Imt
Vinayagamoorthy et al. [2019]	–	✓	–	✓	✓	where	HD video / Holographic video	Monoscopic imaging	3	P - XR - Imt

2019: The interface and gateway to a 1:1 correspondence between the digital and the real world

This recent perspective from Azuma [2019] converges the Internet (as the World Wide Web, cloud computing, and access to remote repositories of content and information) with AR, and contours the vision of an "AR Cloud" representing the gateway access to persistent virtual content attached to real locations.

> **Key AR concepts:** convergence, gateway, interface.
> **Implications for ARTV:** Augmented television is a gateway to cloud-based digital content and corresponding services for television.
> **Connected TVX/IMX area:** on-demand TV, iTV.

All these perspectives on AR highlight distinct concepts, such as *immersion*, *information*, *interface*, *experience*, useful to derive implications for ARTV, as we have been doing in this section. They are also useful to identify areas of scientific investigation from TVX/IMX that connect to ARTV; we relate to these areas in Section 4.5. For now, we continue our examination of ARTV with an overview of the field.

4.4 A literature survey of AR for Television

We overview prior work in AR for television and discuss concepts and/or implementations for AR and the living room.

4.4.1 Method

We conducted a targeted literature survey to locate relevant papers for ARTV. We identified a total number of 17 papers (from 338 candidates) describing ARTV systems

or concepts (see Table 4.1) by running the following platform-compatible query against the ACM Digital Library ($n = 108$ papers), Scopus ($n = 192$ papers), and IEEE Xplore ($n = 38$ papers):

```
(``Augmented Reality'') AND (Television OR TV)
```

We considered papers published in conference proceedings and journals during the last twenty years (2000 to 2019) that explicitly contained our key terms in their title, abstract, or keywords.[8] Based on our experience, we also considered for inclusion other papers that were not picked up by the query, from which we selected Jones et al. [2014] as an extension of Jones et al. [2013].

Next, we review these papers and classify them according to Milgram and Kishino's [1994] three dimensions of the taxonomy for merging real and virtual worlds, as follows:

- *Extent of World Knowledge (EWK)*: refers to how much about the world being modeled is known to the system.

- *Reproduction Fidelity (RF)*: refers to the quality with which the synthesizing display reproduces the actual or intended images of the objects being displayed.

- *Extent of Presence Metaphor (EPM)*: is the extent to which the viewer is intended to feel present within the scene.

We also identify for each paper the classes of the implemented AR displays, according to Milgram and Kishino [1994]:

(Class 1): monitor-based video displays,

(Class 2): HMDs,

(Class 3): HMDs with see-through capability,

(Class 4): video-based see-through HMDs,

(Class 5): completely graphic display environments, and

(Class 6): completely graphic, but partially immersive environments, such as large displays, in which physical objects play a role in the virtual world.

Furthermore, inspired by previous initiatives to specify contributions in HCI [Wobbrock and Kientz, 2016, López Jaquero et al., 2019], we group prior work according to their contributions to ARTV using the following categories: 1) new technology,

---

[8]In cases where the same authors published evolving work in different venues over time, such as Popovici and Vatavu [2019c,b,a], we only kept the latest publication.

2) application, 3) method, 4) user study, and 5) design recommendations. Table 4.1 presents a summary of our literature survey.

4.4.2   Window-on-the-world displays (class 1)

Many of the papers identified in our literature survey implemented the "window-on-the-world" metaphor for augmented television using hand-held devices, such as smartphones and tablets. For example, Gómez et al.'s [2014] system enabled users to browse a tree of resources linked to a physical object using AR on a hand-held device; as a use case, they augmented a map of the world displayed on a TV set, enabling the viewer to access extra information about specific map regions based on the viewer's distance from the TV. Kawakita and Nakagawa [2014] created a system where a 3-D virtual TV character appeared to come out of the TV screen when viewed on a hand-held display. Revelle et al. [2015] proposed a "transmedia" game for children to learn new words. Kawamura and Otsuki [2018] presented an imaginary world on the TV screen which was equally viewable in 3-D free-viewpoint AR on a hand-held display. And Zimmer et al. [2018] immersed viewers into a story by eliminating the TV set and delivering the content using AR exclusively.

Besides hand-held devices, Class 1 displays have been implemented with augmentations on the TV screen itself. For example, Stauder and Robert [2002] proposed a method to ensure photometric realism when virtual artifacts were overlaid on top of the TV content, and Sotelo et al. [2017] described a system enabling the viewer to interact with a 3-D model overlaid on broadcast video. Vatavu's [2015] "Audience Silhouettes" prototype overlaid live 3-D representations of viewers' body movements directly on top of the TV broadcast.

4.4.3   Head-mounted displays (classes 2, 3, and 4)

A few systems have implemented ARTV using HMDs. Baillard et al. [2017] created a multi-user system where AR content was displayed on the periphery of the physical TV set, viewable using both a hand-held and an HMD. Saeghe et al. [2019a] displayed virtual animated content related to the broadcast outside the TV frame. Vinayagamoorthy et al. [2019] described a prototype where a sign language interpreter was presented to the viewer next to the TV set.

4.4.4   Graphic display environments (classes 5 and 6)

In 2012, Vatavu [2012] proposed an interactive home entertainment system where multiple virtual TV screens were video projected on the living room wall and independently controlled. In a follow-up paper, Vatavu [2013] introduced "AroundTV"—a video projection-based system for the area surrounding the physical TV set for interacting with virtual TV screens and graphical user interface widgets. At the same time, Jones et al. [2013] introduced "IllumiRoom" for computer games, a system that projected visualizations of the game in the periphery of the TV screen. The follow-up "RoomAlive" system and installation was meant to transform the entire room into an immersive, augmented space [Jones et al., 2014]. To address peripheral vision, Kimura and Rekimoto [2018] proposed "ExtVision"—a system that generated and displayed context images in the area surrounding the TV set.

4.4.5   User study contributions in ARTV

In our survey, we found that nine papers implemented the window-on-the-world metaphor (Class 1), three papers used see-through HMDs (Class 3), and six implemented graphic environments of Class 5 or 6; see Table 4.1. Some papers did not present actual implementations, since they focused on user studies. For example, Geerts et al. [2019] used a co-design approach to generate a scenario where extra content was displayed in the viewers' environment. Popovici and Vatavu [2019c] proposed an agenda for researchers to consider when designing for ARTV. In two follow-up papers, they elicited user preferences regarding ARTV [Popovici and Vatavu, 2019a] and used the findings to consolidate the original research agenda [Popovici and Vatavu, 2019b].

4.4.6   World knowledge, presence, and fidelity

We used Milgram and Kishino's [1994] dimensions to classify the ARTV literature. We found that all papers used information about the location (*where*) of the modeled world, four papers modeled both the *where* and *what*, while only two papers implemented a complete model of the world (Jones et al. [2014], Kawamura and Otsuki [2018]). As the spatial understanding capability of AR advances, we can expect systems to approach complex models of the displayed world. While six papers described a real-time hi-fidelity 3-D animation, one paper used basic graphics (Gómez et al. [2014]) and another employed holographic video via chroma-keying techniques (Vinayagamoorthy et al. [2019]). Regarding the Extent of Presence Metaphor dimension, four papers (Gómez

et al. [2014], Vinayagamoorthy et al. [2019], Vatavu [2015], Zimmer et al. [2018]) used AR to deliver a monoscopic image, while the rest delivered real-time images with AR artifacts that could be viewed from multiple points of view.

4.5 TVX/IMX areas connected to ARTV

Section 4.3 highlighted key concepts for ARTV, which led to connections to several areas of scientific investigation representative for the TVX/IMX community. We discuss these areas in this section to unveil further aspects of ARTV.

4.5.1 Interactive Television

Broadcasters are trying to engage TV audiences through the addition of data services on top of traditional television to increase participation and feedback [Vinayagamoorthy et al., 2012]. This has resulted in the TV evolving from a purely audiovisual platform to one with in-built interactive services, such as teletext, electronic program guides, or red-button services. Users are increasingly opting to purchase smart TVs with IP connectivity that run applications and integrate with connected devices in the home, such as conversational user interfaces [Westover, 2019, Corpuz, 2019]. Besides the attractive prospect of personalizing the ways in which users might control their connected TVs, there have been ventures to personalize and augment the viewing experience itself [Geerts et al., 2014, Vinayagamoorthy et al., 2012, Lohmüller and Wolff, 2019] through synchronized companion experiences based on audio watermarking, fingerprinting, and HbbTV 2.0 [Vinayagamoorthy et al., 2016]. The potential to personalize television experiences connects interactive TV with ARTV, where virtual objects augment viewers' personal TV watching experience.

On the big screen itself, the move from broadcast to streaming over IP enables more interactive storytelling. The go-to obvious format has been play along quizzes and voting, but the interactive potential can go further [Hanson, 2019b]. For example, Netflix's Bandersnatch [Netflix, 2018] is a non-linear branching narrative that progresses the stories off depending on the choice the viewer makes. Also, as content is increasingly consumed over IP, content creators know more about their audiences through the data they can collect. Unlike branching narratives, object-based media (OBM) allows content to change according to the requirements of each individual audience member [Hanson, 2019c]. This affords a versatile manner in which the story is remixed according to

the audience [Cox et al., 2017, Li et al., 2018, Hanson, 2019c, Bauwens et al., 2019], connecting to the potential of ARTV to render new ways for the virtual story to play out. Examples of OBM experiences include immersive audio [Francombe, 2018, Baume, 2017], branching narratives [Brooks, 2019], personalized documentary [Hanson, 2019a], and personality quizzes [BBC, 2019]. In ARTV, future media offerings will more easily personalize to the audience, which means adding interaction and augmentation to engage the audience in the storytelling process.

4.5.2   3-D Television

3-D photography, cinema, and TV have a long history from the first examples of passive stereoscopic 3-D cinema in the 1990s to manufacturers involved in marketing autostereoscopic 3-D TVs in the late 2000s [Ozaktas and Onural, 2010, Meesters et al., 2004]. The 3-D TV display systems use a combination of a 2-D image and a depth map [Grau and Vinayagamoorthy, 2010] (i.e., depth image-based rendering) to synthesize new virtual views and, consequently, to augment the original image from the 2D-plus-depth data [de Sorbier et al., 2010]. These advances in 3-D TV technology are relevant to support developments in ARTV since, according to Azuma [1997], 3-D registration is one of the three key characteristics of any AR system.

4.5.3   Ambient media

Research in Ambient Media has unveiled a new type of media, conceptually different from television, print, and digital media, that define the media environment of smart spaces. According to Lugmayr et al. [2013], "Ambient media in a larger scale define the media environment and the communication of information in ubiquitous and pervasive environments." Among its characteristics, ambient media is subtle, unmonopolizing, and addressing peripheral awareness [Schmidt et al., 1999], while it can morph and manifest in various ways [Pogorelc et al., 2012, 2013]. These properties make ambient media especially relevant for ARTV where, according to a recent perspective [Azuma, 2016], AR technology itself can be qualified as a specific form of new media.

4.5.4   Immersive media and multimedia alternate realities

Immersion, Interactivity, and Imagination (I^3) [Burdea and Coiffet, 2003] are central concepts associated with the sense of presence in VR, AR, and MR with strong

roots in computer-generated graphics and content. Immersion is influenced by sensory and perceptual modalities associated with "presence", i.e., the feeling of being inside the computer-generated reality due to realistic feedback, participation, and social immersion [Chambel et al., 2013]. Developments in immersive media in the context of interactive television in specialized communities and events, such as TVX/IMX (and, before it, EuroITV), ACM Multimedia, IEEE QoMEX, or general CHI/ISMAR venues, have focused on aspects such as audiovisual immersion, 3-D and panoramic 360° multi-view and holographic video, spatial and stereoscopic audio, perceptual immersion and multi-sensory interaction, interactive and immersive television, film, and cinema [Chambel et al., 2013, 2016].

Chambel et al. [2013, 2016] introduced the concept of "Multimedia Alternate Realities" (MMARs) as "different spaces, times or situations that can be entered thanks to multimedia contents and systems, that coexist with our current reality, and are sometimes so vivid and engaging that we feel we are living in them . . . immersive experiences that may involve the user in a different or augmented world, as an alternate reality" [Chambel et al., 2016]. To properly characterize such realities, a taxonomy with eight dimensions was proposed [Chambel et al., 2016] consisting of: the alternate (e.g., different space, time, context), the virtual-augmented spectrum, the real-fictional spectrum, the level of interactivity, the level of immersion (e.g., presence and belonging, imagination, and engagement), the multisensorial (the media and modalities involved), the personal (adaptation to preferences and contexts), and the social dimension (individualized or shared realities). MMARs go beyond the focus of VR, AR, and MR by addressing new media and immersive experiences. In the context of ARTV, these dimensions become relevant when television is part of the audiovisual content delivered inside the MMAR. Of those, the virtual-augmented and real-fictional spectra connect directly to ARTV.

4.6 A conceptual framework for ARTV

Based on our findings and discussion so far, we introduce a conceptualization for ARTV. To this end, we adopt the following three principles regarding AR technology, the real-world environment of the living room, and end users:

P 1. *Independence of AR rendering technology.* Just like Azuma [1997] for AR, Bimber and Raskar [2005] for spatial AR, and Milgram and Kishino [1994], Milgram et al. [1995], Milgram and Jr. [1999] for MR, we believe

that the technology to render computer-generated graphics in the living room environment should not drive the concept of ARTV. From this perspective, smartphones and tablets [Kawakita and Nakagawa, 2014, Revelle et al., 2015, Kawamura and Otsuki, 2018], smart glasses and HMDs [Vinayag-amoorthy et al., 2018, 2019], video projections [Vatavu, 2012, 2013], and wall- and room-sized projections [Kimura and Rekimoto, 2018, Jones et al., 2014, 2013] are implicitly encompassed by our framework. All that matters is that the TV experience is augmented.

P 2. *Focus on the living room.* We restrict our discussion of ARTV to the context of the living room environment. Thus, we exclude mobile television [Oks-man et al., 2007, McNally and Harrington, 2017], for which a proper investigation will need the context of mobile AR [de Sá and Churchill, 2013]. Our focus on the living room is motivated by traditional TV being predominant in the TV landscape, e.g., 93% of streamers watch traditional TV on a typical day [Nielsen, 2018] as the growth of mobile media levels out [Zenith Media].

P 3. *Focus on the viewers' side of ARTV.* AR has two applications in TV broad-casting: 1) TV production, which aims to increase productivity and/or reduce costs, often referred to as the "virtual studio" [Gibbs et al., 1998], and 2) application on the viewers' side, which aims to create novel experi-ences in the viewers' environment during television watching. In this work, our focus has been on the latter.

4.6.1   The ARTV Continuum

Our exploration from the previous sections revealed many key characteristics of ARTV. We have seen that the RV Continuum [Milgram and Kishino, 1994] (Figure 4.2) rep-resents the most accepted, go-to source for defining AR and to distinguish MR from AR [Speicher et al., 2019]. In the following, we adopt the RV Continuum as the basis for our conceptualization of ARTV. But, while this continuum can be successfully employed to characterize the degree of mixture between real and virtual, its 1-D design is insufficient for our purpose. By considering the TV viewing experience where a *physical TV set* is placed in a *physical living room*, we recognize the fact that each of these two entities, world and TV, can be independently augmented. By adopting this perspective, we propose and introduce a 2-D conceptual framework for ARTV; see

Figure 4.4: The ARTV Continuum. *Note:* the red dotted line loosely marks a delineation between ARTV and AVTV.
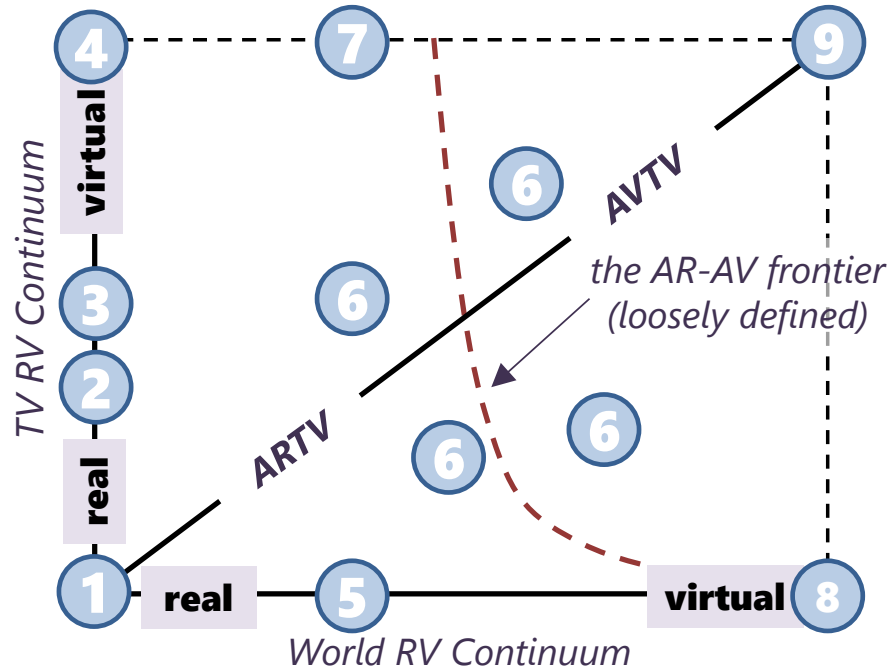
Figure 4.4.

The horizontal axis of Figure 4.4, going from a completely real to a completely virtual environment is Milgram and Kishino's [1994] RV Continuum that applies to the living room environment (the world). The vertical axis, going from a completely physical TV device to fully virtual televised content is our adaptation of Milgram and Kishino's [1994] RV Continuum for television. Together, these two orthogonal axes characterize the various ways in which a television viewing experience can be augmented, e.g., in terms of the world, the televised content, or both. In this conceptual framework, content can flexibly flow between the living room and the TV set, while the living room and the TV set can independently flow across their respective continua from real to virtual. We call this space the "Augmented Reality Television Continuum."

4.6.2 Examples of using the ARTV Continuum

To demonstrate the utility of the ARTV conceptual framework, we enumerate various types of ARTV generated from the mixture of the two axes of Figure 4.4. For each category, we note examples from the literature where available and, where not, we highlight open areas for future work.

1. *Conventional world/conventional TV.* This category represents the conventional TV viewing experience where neither the TV nor the room are augmented. The corresponding position in the ARTV Continuum is the bottom-left of Figure 4.4. Similar to how the RV Continuum captures the real world [Milgram and Kishino, 1994], the conventional TV viewing experience is equally captured by our framework.

2. *Conventional world/conventional TV with on-TV augmentation.* In this category, contextual augmentations are shown on the TV screen. Examples include LinkedCulture [Nixon et al., 2015], CollaboraTV [Nathan et al., 2008], and Audience Silhouettes [Vatavu, 2015]. Teletext and Picture-in-Picture can also be included.

3. *Conventional world/conventional TV with off-TV augmentation.* Different from the previous category, augmentation is now perceived to be off the TV screen, e.g., on the wall behind it, in front or next to the TV set. Implementations may involve hand-held devices [Kawakita and Nakagawa, 2014, Revelle et al., 2015, Kawamura and Otsuki, 2018], HMDs [Vinayagamoorthy et al., 2019, Saeghe et al., 2019a], or video projection-based technology [Vatavu, 2013, Jones et al., 2013, Kimura and Rekimoto, 2018, Jones et al., 2014]. This category is also located on the vertical axis in Figure 4.4.

4. *Conventional world/virtual TV.* A physical TV is not needed for this category of ARTV, where televised content is presented either in a virtual TV frame or even without a TV frame altogether. Examples include TV projected on the wall [TIME, 2017, Vatavu, 2012], and when the storyteller appears to be sitting in the living room [Zimmer et al., 2018] in the viewers' space.

5. *Augmented world/conventional TV.* In contrast to the previous categories, the aim in this case is to augment the living room directly. This often requires spatial understanding of the surfaces present in the living room for meaningful alignment of the computer-generated graphics. Viewers experience an augmented world, but watch a conventional TV screen. Nevertheless, the augmented world can offer affordances with respect to television watching, e.g., the possibility to record the TV show, open a mid-air browser window with a fact sheet about the current televised content, have Skype calls with

friends watching the same broadcast, etc. Opposed to the previous categories, the living room is now the substratum that is augmented, but the TV is still conventional, which places this category on the horizontal axis; see Figure 4.4.

6. *Augmented world/conventional TV with augmentation.* The world is augmented and so is the TV set, either by means of on-TV or off-TV augmentation. The degree of augmentation of each component, world and TV, positions implementations of this category at various locations in the ARTV Continuum illustrated in Figure 4.4.

7. *Augmented world/virtual TV.* This category is achieved when the world is augmented (as in *6*), but the physical TV is virtual (as in *4*). These characteristics position category *7* at the top of Figure 4.4. For instance, in the RoomAlive [Jones et al., 2014] prototype, the room is modeled and content is projected on its surfaces directly without using a physical TV set. The home entertainment prototype of Vatavu [2012] implemented virtual TV screens exclusively.

8. *Virtual world/conventional TV.* This category resembles the Augmented Virtuality of the RV Continuum [Milgram and Kishino, 1994]: a virtual world is augmented by a physical TV set. As a use-case, imagine watching your favorite TV show as a live video feed of your physical TV screen, while wearing a VR headset but still being physically present in the same room with friends watching the same physical TV screen.

9. *Virtual world/virtual TV.* Similar to the previous category, with the exception that the physical TV set is replaced by one or more virtual screens. This includes the scenario where a virtual TV is aligned to a virtual model of a conventional TV set in order to recreate a familiar TV viewing experience in a virtual space. This category is located at the top-right of Figure 4.4.

Other scenarios can be imagined in our conceptual space, depending on the interpolation between real and virtual on both the world and the TV axes; see the multiple instances of ARTV category *6* shown in Figure 4.4. We note that many of the categories that we discussed in this section haven't been proposed yet in the literature, which reveals the generative power [Beaudouin-Lafon, 2004] of our conceptual framework for ARTV.

It is also important to note that Figure 4.4 accommodates AVTV as well, represented

by the region located at the intersection of the Augmented Virtuality part of the world continuum with the TV axis. However, just like in Milgram and Jr.'s [1999] case, the distinction between ARTV and AVTV can only be defined in loose terms, e.g., "As we venture away from the poles of the RV continuum towards the centre, we also eventually begin to encounter the problem of deciding whether in fact what we are doing is augmenting a real world with virtual graphic objects, or whether we are modifying a virtual environment by augmenting it with real data," and "it is not always . . . simple . . . to distinguish between AR and AV" [Milgram and Jr., 1999], which is equally true for our ARTV space.

4.7   Conclusion

We examined in this paper the concept of ARTV by following a four-step method starting from understanding the foundations of ARTV in AR to highlighting key concepts, performing a literature survey of ARTV and connected areas and, finally, to proposing a continuum for conceptualizing and structuring current and future work in ARTV.

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.* While the specific implementation may vary (and, in the years to come, we are to see more innovations in this regard), what is important at this moment is to have a rigorous basis to structure future research and developments, to be consistent in our terminology, and to communicate our understanding of ARTV precisely by relating to proper frameworks. Our ARTV Continuum is an attempt toward such a systematization and toward providing the community with a common vocabulary for possible categories of ARTV systems to enable better understanding and communication of advances in ARTV.

Our goal in this paper has been to provide researchers and practitioners with a conceptual framework in which to think about, talk about, and position their ARTV work and prototypes. We look forward to seeing how the community will employ our conceptual framework to inform the design of novel interactive augmented experiences for television.

Chapter 5

Augmenting Television with Augmented Reality

5.0   Chapter overview

5.0.1   Thesis context

In the previous chapter a conceptual framework for ARTV experiences in the living room was created. The current chapter presents a pilot user study that aims to investigate viewers' perception of an ARTV experience when programme-related holographic content is added to a nature documentary TV show.

The chapter answers the following research question: How does augmenting an existing TV programme, using passive AR artefacts, affect the viewer's experience and engagement?

The main content of this chapter is adapted from: Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, pages 255–261, New York, NY, USA, 2019a. ACM. ISBN 978-1-4503-6017-3. doi: 10.1145/3317697.3325129. URL `http://doi.acm.org/10.1145/3317697.3325129`.

5.0.2   Author's contributions

Pejman Saeghe designed the study, created the prototype, recruited participants, conducted the sessions, collected the data, analysed the results, wrote and presented the paper. Ollie Pattinson and Bruce Weir collaborated on the creation of the prototype. Ollie Pattinson and Maxine Glancy collaborated on data collection and analysis. Sarah Clinch, Stephen Robert Pettifer, Bruce Weir, and Robert Stevens provided critical feedback and edits during the writing of the paper.

5.0.3   Published abstract

This paper explores the effects of adding augmented reality (AR) artefacts to an existing TV programme. A prototype was implemented augmenting a popular nature documentary. Synchronised content was delivered over a Microsoft HoloLens and a TV. Our preliminary findings suggest that the addition of AR to an existing TV programme can result in creation of engaging experiences. However, presenting content outside the traditional TV window challenges traditional storytelling conventions and viewer expectations. Further research is required to understand the risks and opportunities presented when adding AR artefacts to TV.

5.1   Introduction

The success of augmented reality (AR) in the military, medical, and industrial settings can be attributed to the affordances it provides in making task-related operations more efficient. For instance, in repairing complex machinery AR may be used to reduce task completion time [Henderson and Feiner, 2011] and in assembly lines it may be used to reduce error rate and mental effort [Tang et al., 2003]. So far, for the general public, AR has achieved relatively small success, compared to that in industro-milarity contexts, primarily in gaming and advertising. However, similar to the Internet—which was first developed, tested and used in the military for a number of years before finding its way eventually to (almost) every household and work environment—AR may find widespread use in daily life in the near future.

The British Broadcasting Corporation Research & Development department (BBC R&D) and the University of Manchester School of Computer Science have been working together to investigate the feasibility and various design considerations of applying AR in TV. One use-case is to add AR content to an existing TV programme.

A few broadcasters, media companies, and researchers have looked at the application of AR in TV. Kawakita and Nakagawa [2014] designed a system called "Augmented TV" where an animated character appeared to emerge from a TV screen when viewed using a hand-held device (HHD). However, their work is primarily concerned with technical implementation details. Lack of a user evaluations is another limitation of their work.

Revelle et al. [2015] created "Electric Agents"; an AR game layered on top of an existing educational TV programme, aimed at children between 6 and 10 years of age. The experience took the viewers back and forth between a passive TV-watching mode and an active game-playing mode. The passive mode would be interrupted by a TV character zapping words out of the protagonist's mouth. This marked the beginning of the game-playing mode where children in a pair would use HHDs looking around the room in order to find those words in both their visual and written representations. The children then used the HHDs to "throw" the matched pair back into the TV. This study shows a potential for creating engaging experiences by mixing the passive TV viewing mode with an active AR-enabled game-playing mode. However, this use-case ignores the potential for creating passive immersive experiences using AR.

Fradet et al. [2017] presented a system called "MR TV Mosaik" where a virtual mosaic is displayed next to the TV enabling: 1) visualisation of other available programmes, 2) switching to a different program via drag and drop from the virtual mosaic to the TV screen, 3) access to programme guides and other textual information about the programme, 4) virtual subtitles, and 5) 3-D objects.

Their system is focused on providing control and extra information via AR. This misses the opportunity to use the affordances of AR in a meaningful way. For instance, using the interaction provided by AR to drive the story instead of changing the channel. Furthermore, these operations can be done using non-AR technology. The lack of a user-based evaluation is another limitation of their work.

In this paper we focus on creating an ARTV experience and consider the following research question:

RQ: How does augmenting an existing TV programme, using passive AR artefatcs, affect the viewer's experience and engagement?

To answer this research question, we created a prototype and evaluated it in a user study using two questionnaires—the User Engagement Scale-Short Form (UES-SF) [O'Brien et al., 2018] and the User Experience Questionnaire (UEQ) [Laugwitz et al., 2008]—and a semi-structured interview. Overall, the scores were high for all

constructs of the UES-SF (Figure 5.1). On the UEQ, *efficiency* and *dependability* scores were close to neutral while the scores for the remaining constructs were higher (Figure 5.2).

Figure 5.1: Aesthetic Appeal (AE), Focused Attention (FA), Perceived Usability (PU), Reward (RW), and overall engagement score (UES) for UES-SF.
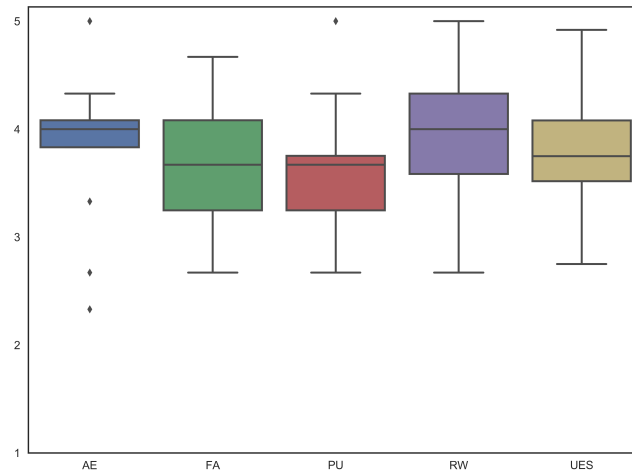
Figure 5.2: Attractiveness (AT), Perspicuity (PR), Efficiency (EF), Dependability (DP), Stimulation (ST) and Novelty (NV) score distributions for the UEQ.
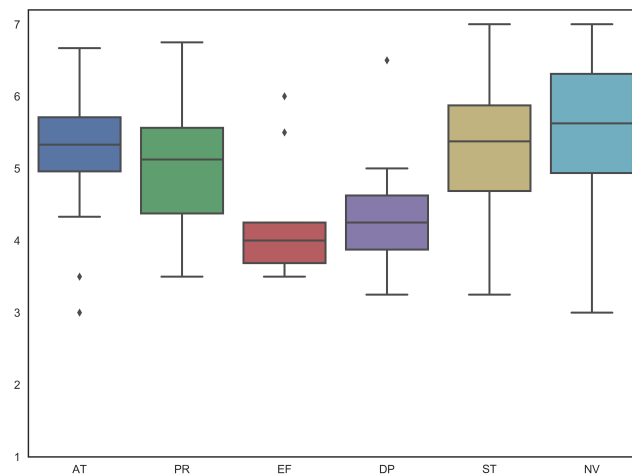
5.2   Prototype

Our prototype delivered content to the viewer via two platforms: the main content was delivered on the TV screen, while the AR content was viewable using the HMD (a

Figure 5.3: The timeline depicting how ARTV content was presented to the viewers.
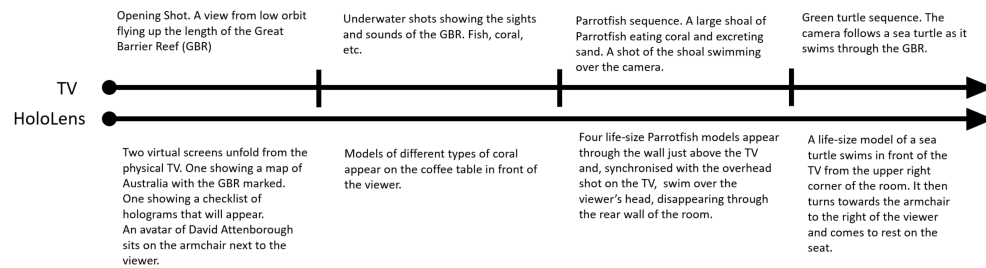
Microsoft HoloLens). Playback of TV and holographic content was synchronised using HbbTV 2.0's sync functionality [Merkel, 2011]. The orchestration of the AR content relative to the TV content can be seen in Figure 5.3.

For the main content, we chose a five-minute video clip from the BBC's Blue Planet programme; a popular documentary exploring the world's oceans. The story in the clip took place around the Great Barrier Reef (GBR) and followed the activities of a female sea turtle.

For the AR content we chose a variety of artefacts in order to evoke conversations in the interviews about ARTV possibilities. The AR artefacts used were:

1) Two virtual screens, either side of the TV, that delivered *contextual* and *feedback* information to the viewer. Context about the location of the GBR was conveyed by showing a map of Australia with the GBR area being highlighted. Feedback consisted of the list of holograms due to appear during the experience (Figures 5.4b and 5.4c).

2) A life-size virtual male human with idle animation, sitting on a real armchair in the viewer's environment, who appears to look at the TV screen and makes eye contact as the viewer looks at him (Figure 5.4a).

3) A set of holograms related to the main content. Corals appeared on a coffee table in front of the viewer. A school of parrot fish appeared above the TV and swam towards the viewer, eventually passing over their head (Figure 5.4b). A sea turtle appeared near the TV, swam a curved path towards the viewer settled next to them on the sofa (Figure 5.4c).

Figure 5.4: Screenshots from the ARTV prototype used in our exploratory study.
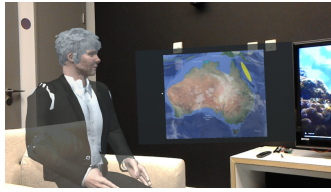
(a) A virtual person sitting on an arm-chair next to the viewer.
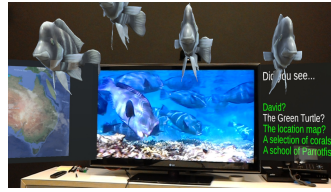
(b) Virtual fish as they swim towards the viewer.

(c) A virtual sea turtle crossing the focal point of the viewer.

5.3   User study

We conducted a user study with a sample of $n = 12$ BBC R&D members of staff. There were no inclusion criteria except for participants' availability and willingness to take part in the study. We followed the BBC's internal ethics procedures, with all participants providing informed consent. The study was conducted in a user-testing lab designed to look and feel like a living room with sofas, armchairs, a coffee table and a smart TV. Each participant attended a session that started with a brief introduction to the research. They were then given a few minutes to familiarise themselves with the HoloLens device and to adjust it for comfort. After watching the programme, they were asked to remove the HoloLens and complete the two questionnaires:

1) The User Engagement Scale-Short Form (UES-SF): consisting of twelve items measuring four constructs: *aesthetic appeal*, *focused attention*, *perceived usability*, and *reward* [O'Brien et al., 2018]

2) The User Experience Questionnaire (UEQ): consisting of twenty six items measuring six constructs: *attractiveness*, *perspicuity*, *efficiency*, *dependability*, *stimulation*, and *novelty* [Laugwitz et al., 2008]

After completing the questionnaires participants engaged in a semi-structured interview conducted by a member of the research team. The main points in the interview were concerned with the following points:

1) overall impression;

2) thoughts on individual holograms regarding each artefact's positioning, visual quality, animation, and potential user interactions;

3) suitable TV genres for this use-case;

4) the context in which the participants thought it would make sense to watch such experiences in terms of location, time of day, social context, etc.;

5) thoughts regarding interactive vs passive nature of the artefacts and the experience as a whole;

6) time of appearance of the holograms in relation to the content on TV;

7) other devices for presenting the AR content such as HHD and projection;

The total engagement required from each participant was approximately one hour.

5.4 Results

UES-SF results indicated that, in general, all measured constructs scored broadly the same, with a moderately high value of $3.5 - 4$ (Figure 5.1). Most participants scored very close to the median on *aesthetic appeal* with a few clear outliers. Spread for *focused attention* and *reward* was much bigger.

UEQ results indicate a big spread of scores for *stimulation*, *novelty* and to a slightly lesser extent for *perspicuity*. These results also indicate agreement around 4 (the neutral score on a 7 point Likert scale) for *dependability* and a tighter agreement around the same value for *efficiency*. All the other constructs score slightly higher (Figure 5.2).

For both questionnaires a larger sample size would yield more reliable readings.

Analysis of the interviews indicate that overall participants found the experience novel and engaging. The favourite hologram was the sea turtle. Some participants found the virtual person unsettling. Most frequent responses about suitable genre were sport and news. Many participants felt the scenario would work best when they are on their own, while some thought it would be interesting to share the experience with others. Some participants found the passive nature of the experience suitable and were not expecting to be able to interact with the AR content; "it did not even enter my head to try and do something, and really, I don't know if I would have wanted to". The fixed orchestration of the AR and TV content divided opinion; some liked it while others wanted to be able to pause the TV content to view and explore AR artefacts at their own pace. Most participants were happy to have their hands free, therefore preferring HMD and projection based AR over HHD. *Fear of missing out* was raised a few times. For some participants this was about missing AR content, while for others this was about

missing TV content. Two participants reported not seeing the AR sea turtle and the AR fish.

5.5   Discussion

Our findings suggest that augmenting an existing TV programme with AR artefacts can be engaging (Figure 5.1). This finding is corroborated by the results of UEQ (Figure 5.2) which suggest our prototype was perceived as novel, stimulating, aesthetically appealing and perspicuous. The neutral scores for efficiency and dependability could be interpreted in two ways:

1) These constructs are not relevant to a passive ARTV hybrid.

2) Participants found the prototype not dependable or efficient.

Another finding suggests that viewers may miss AR content that is presented outside the TV display. In our study two participants missed the virtual sea turtle and the virtual fish. This phenomenon could be interpreted in multiple ways:

1) *Inattentional Blindness (IB):* The stimuli presented outside the primary zone of attention (i.e., the TV screen) may fall under IB [Mack et al., 1998].

2) *Novice participants:* Most participant had very little, if any, prior exposure to AR and none had watched an ARTV experience before. Prior work has suggested that a lack of experience, as opposed to a lack of attention, may cause missing of visual content [Braun, 1998].

3) *Missing conventions for ARTV:* Each of the currently mature media (as well as each genre within each media) has gone through an initial phase of trial and error in order to establish conventions. For instance, fans of the horror genre know what to expect when they go to the cinema to watch a horror film. Conversely, since no conventions yet exist for ARTV, viewers could simply be missing content because they do not know what to expect.

5.6   Future work

We used various types of holograms to augment a natural documentary programme. Future research is required to understand the effects of each type of AR artefact on viewer engagement.

Experimentation with different genres of TV programmes is required to understand how augmentation affects viewer engagement regarding other genres.

Creating a story from scratch for ARTV would yield useful insights and provide the opportunity to test how the affordances of AR could be mixed with TV to create engaging experiences.

Objective measures such as eye-tracking and user interaction data can be used to provide a different and perhaps more granular insight into viewer engagement.

Other subjective questionnaires could be used for instance to measure immersion with [Jennett et al., 2008b] and cognitive load with [Hart and Staveland, 1988].

5.7   Conclusion

AR is likely to become ubiquitous in the near future. The application of AR to enhance conventional TV viewing can result in engaging, stimulating and novel experiences. However, there are risks and issues involved with delivering content in this hybrid format that needs to be understood by further research and exploration.

Chapter 6

Exploring the Display Dimension with a Lifelike Hologram

6.0 Chapter overview

6.0.1 Thesis context

Chapter 4 identified nine types of ARTV experiences. The previous chapter explored the impact of off-TV augmentations, where both TV and the living room were *real*—type 3 of our conceptual framework. This current chapter explores type 9 of our framework, where both the TV and the living room are virtual.

While the previous chapter explored various types of augmentations, this chapter focuses on the impact of changing the movement behaviour of a lifelike hologram on ARTV viewers' experience.

The chapter answers the following research question: Considering the following three options for a lifelike hologram's entry point: 1) the TV screen 2) adjacent to the TV screen 3) outside the field of view and two options regarding its subsequent behaviour: 1) moving continuously in the viewing environment 2) moving towards, and settling in, a destination close to the TV screen and within the field of view (e.g., on a coffee table) how, if at all, do changes in these options impact an ARTV viewer's experience?

The main content of this chapter is adapted from: Pejman Saeghe, Bruce Weir, Mark McGill, Sarah Clinch, and Robert Stevens. Augmenting a nature documentary with a

114

6.0.2 Author's contributions

Pejman Saeghe designed the study, created the experimental software, recruited participants, conducted the sessions, carried out the data collection, conducted interviews, analysed the results, and wrote the paper. Bruce Weir and Mark McGill collaborated in the creation of the experimental software, and along with Robert Stevens, and Sarah Clinch, provided continuous support and critical feedback throughout all stages of the study.

6.0.3 Published abstract

As the underlying technology for Augmented Reality Television (ARTV) matures, content creators need to understand how to create engaging ARTV content. One approach is to explore many possible options afforded by an underlying design space. We investigate six points on a design dimension of presenting a lifelike programme-related hologram to a nature documentary viewer. In a repeated-measures study ($n = 10$), we manipulate the *starting point* and the *movement behaviour* of a hologram and measure various aspects of participants' ARTV experiences.

Our findings highlight the importance of personal preferences and that of the perceived role of a hologram in relation to the underlying TV content; suggesting that there may not be a single way to augment a TV programme, and it may not be suitable to augment existing TV shows. Instead, creators need to tell a story over both AR and TV, and give audiences capabilities to customise ARTV content.

6.1 Introduction

Augmented Reality Television (ARTV) has been a topic of interest for researchers and broadcasters in the past few decades (e.g., Stauder and Robert [2002], Vatavu [2012], von der Au et al. [2020]). However, there has been a tendency in prior work to either overlook evaluation completely (e.g., Stauder and Robert [2002], Sotelo et al. [2017]) or to conduct limited evaluations, such as evaluating a specific implementation of a

proposed system to validate a set of design decisions (e.g., Revelle et al. [2015]), to measure system usefulness (e.g., Vatavu [2012]), or to measure the performance of the underlying system (e.g., Gómez et al. [2014]).

As the underlying ARTV technology matures, implementation of ARTV experiences will become easier. However, due to a lack of a shared language for ARTV—i.e., a common set of conventions shared by content creators and their audiences—it is challenging to create effective and engaging ARTV experiences. This may result in the creation of ARTV experiences that do not deliver any real value for the audiences, beyond their initial novelty effect, which in turn can shape the public's perception of ARTV as technology for the sake of technology. Thus, it is important to shift the focus of research from a predominantly technology-oriented one to one that is audience-centred.

One possible direction is through rapid prototyping to explore an underlying design space of a desired scenario; to investigate the impact of various design options on the ARTV audiences, and decide which, if any, would be suitable for a target audience demographic.

In this chapter, we create an ARTV prototype and investigate six points on a *display* dimension for a lifelike programme-related hologram. We focus on a scenario where a single viewer is sitting in front of a TV set, watching a nature documentary show, and refine Saeghe et al.'s [2020] display dimension to consider three display regions for a complimentary hologram: 1) on the TV screen, 2) around the TV screen, and 3) outside the field of view. Furthermore, since our lifelike hologram—a sea turtle—moves in the viewing environment, we consider the point from which it enters the viewing environment and its subsequent behaviour throughout a short ARTV episode, and ask the following research question (RQ):

RQ: Considering the following three options for a lifelike hologram's entry point:

    1) the TV screen

    2) adjacent to the TV screen

    3) outside the field of view

and two options regarding its subsequent behaviour:

    1) moving continuously in the viewing environment

    2) moving towards, and settling in, a destination close to the TV screen and within the field of view (e.g., on a coffee table)

how, if at all, do changes in these options impact an ARTV viewer's experience?

In a remote repeated-measures user study ($n = 10$) we measure various subjective aspects of viewers' ARTV experience, such as *focused attention*, *immersion*, *confusion*, *distraction*, *enjoyment*, and *interest*. We subsequently conducted an interview over Zoom with each participant.

Our findings highlight the importance of personal preferences and the importance of the perceived role of the lifelike hologram in relation to the underlying TV content and narrative. If replicated regarding other dimensions and other types of holograms, the implications of our findings suggest that there may not be a single suitable way to augment a TV programme; instead, audiences may want the ability to customise their experience based on individual needs and preferences. Furthermore, it may not be suitable to augment an existing TV show; instead, content creators need to design ARTV experiences in ways that a single coherent narrative is told over both AR and TV.

Our contributions in this chapter are:

1) the first ARTV user study to investigate the impact of changing two specific design decisions regarding the visual presentation of a lifelike hologram

2) a refinement of Saeghe et al.'s [2020] display dimension, by introducing three visual regions in the context of an ARTV experience for the living room

3) a three-tier framework to classify a hologram's movement in the viewing environment

4) a six-item questionnaire that specifically targets a viewer's ARTV experience

6.2 Related work

We overview prior ARTV-related research, and highlight a few studies that went beyond testing a fixed set of design decisions. Finally, we expand on Saeghe et al.'s [2020] *"Display"* dimension, which provides the basis for the six points of the ARTV design space that we investigated in this chapter.

6.2.1 Augmented Reality Television

Enhancing a conventional TV viewing experience was reported by Saeghe et al. [2020] as the most widely addressed theme in ARTV research, with use-cases including

provision of novel interaction techniques (e.g., Baillard et al. [2017]), extending the real estate of a TV screen (e.g., Kimura and Rekimoto [2018]), delivering programme-related additional holograms in the viewing environment (e.g., Saeghe et al. [2019a]), delivering virtual TV screens around a TV set (e.g., Vatavu [2013]), or even replacing a TV set entirely (e.g., Zimmer et al. [2018]).

ARTV for the living room was conceptualised by Vatavu et al. [2020]—inspired by Milgram and Kishino's [1994] *reality-virtuality continuum*—as a 2-D space, where the TV set and the living room were each placed, and free to move, on their own reality-virtuality continuum. This design space was used by Vatavu et al. [2020] to generate nine types of ARTV experiences, one of which consisted of viewing a virtual TV screen in a virtual living room—a category that was highlighted as one of the areas where further research was required.

6.2.2   ARTV user studies

Traditionally, user studies conducted in the ARTV field have tended to be of a limited scope; typically, focusing on testing an implementation with a set of design decisions made in advance by the researchers (e.g., Vatavu [2012], Gómez et al. [2014], Revelle et al. [2015]).

A few studies, however, have explored a wider range of possibilities afforded by an underlying design space. For instance, Jones et al. [2013] identified three primary dimensions of a design space for peripherally projected *illusions*, in the context of playing video games, and investigated eleven points on this design space. More recently, Mathis et al. [2020] used VR as a test-bed and investigated TV viewers' preferences regarding the positioning of text messages in relation to a virtual TV screen. They investigated five points in the underlying design space. In another study, McGill et al. [2020] investigated the impact of using audio glasses (Bose Frames) to deliver various channels of a 5.1 surround mix, in addition to a stereo mix (down-mixed from an original 5.1) played from the speakers on a TV set. They investigated three points regarding genre, and three points regarding the audio, in the underlying design space.

6.2.3   Entry point and movement behaviour

Saeghe et al. [2020] defined "display" as an ARTV design dimension that affords various possibilities regarding the visual presentation of AR content in relation to TV content, from the perspective of a TV viewer. In the context of a single viewer positioned in

front of a TV display, we refine this dimension by considering three broad regions where a hologram can be positioned in relation to a TV screen and the field of view of the viewer: 1) superimposed on the TV screen, 2) around the TV screen, and 3) outside the field of view. Furthermore, since a hologram can potentially move in 3-D space, we dissect its movement position into three components: 1) starting point, 2) travel path, and 3) destination.

While a non-moving hologram can be described as one with no *travel path* with its *starting* and *destination* points coinciding, at least two versions of a moving hologram can be envisaged; by changing the destination parameter. For instance, a hologram can start at a point, travel a path, and then *settle* on a destination, or it can move *continuously* in the viewing environment, without an apparent destination.

In this chapter, we investigate six points on the design space of an ARTV experience where a moving lifelike hologram is added to an underlying nature documentary show. Namely, we investigate the impact of changing the starting point—between three above-mentioned viewing regions—and its movement behaviour between settling and continuously moving in the viewing environment.

6.3 Method

We conducted a remote repeated-measures user study ($n = 10$) in VR to investigate the impact of changing the starting point and the hologram's movement behaviour in the viewing environment on viewers' ARTV experience. The study was approved by the departmental ethics committee at The University of Manchester (Reference: 2021-11427-19154).

This section provides the following information:

1) The ARTV experience (Section 6.3.1), describing ARTV content experienced by the participants.

2) Experimental conditions (Section 6.3.2), describing six experimental conditions used in this study.

3) The VR environment (Section 6.3.3), describing the virtual environment in which the study took place.

4) Procedure (Section 6.3.4), describing the typical procedure before, during, and after each session.

5) Instruments (Section 6.3.5), detailing the instruments used to measure various aspects of the viewers' experience.

6) An overview of participants' demographics (Section 6.3.6)

6.3.1   The ARTV experience

The ARTV experience consisted of a video clip (TV content) and a hologram (AR content), where the TV show's main character—a sea turtle—appeared outside the TV screen as a life-sized hologram.

The video clip (3.6 minutes in duration) was a segment from the Coral Reefs episode of the BBC's Blue Planet II programme[1]. It told a story around the daily activities of a sea turtle, such as searching for food and competing for a spot on a rock at the bottom of the sea to get its shell cleaned by small fish.

The hologram was a life-size 3-D sea turtle asset acquired from the Unity Asset Store[2]. We used the Unity game engine to build the experience and deployed it to the Oculus Quest 2 VR headset.

The original video clip was split into six shorter clips (each about 36 seconds in duration), enabling us to create six short ARTV episodes corresponding to our six experimental conditions (see Table 6.1). The order of the video clips remained unaltered between participants, to keep the narrative flow in tact. The experimental conditions were counterbalanced according to a Latin square.

Table 6.1: Six experimental conditions.

Condition	Starting Point	Movement Behaviour
1	The TV screen	Settle on the coffee table
2	The TV screen	Continuous movement
3	Adjacent to the TV screen (wall)	Settle on the coffee table
4	Adjacent to the TV screen (wall)	Continuous movement
5	Outside the field of view (ceiling)	Settle on the coffee table
6	Outside the field of view (ceiling)	Continuous movement

6.3.2 Experimental conditions

The study had a 3x2 design, consisting of 3 conditions for hologram starting point: 1) the TV screen, 2) adjacent to the TV screen, and 3) outside the field of view. and 2 conditions for hologram movement behaviour: 1) settling below the TV screen on a coffee table, within the field of view, and 2) continuously move in the viewing environment. Table 6.1 presents the six experimental conditions that result from mixing three starting points and two hologram behaviours.

In the TV screen, TV screen-adjacent, and outside field of view starting point conditions the hologram entered the environment from a TV screen, the wall next to the TV screen and the ceiling, respectively (see Figures 6.2b, 6.2c, and 6.2d, respectively).

In the settle on the coffee table conditions, immediately after entering the viewing environment, the hologram swam towards the coffee table (see 6.3.3) and rested on it. The resting animation consisted of minimal eye and mouth movements.

In the continuous movement conditions, after entering the viewing environment, the hologram swam across the living room, in an elliptical path towards the viewer, then towards the TV screen. The short length of the clips meant that regardless of its starting point, the hologram would arrive to the area near the top of the TV screen by the end of the clip.

In all conditions, once the hologram entered the viewing environment it remained visible in the viewer's field of view, without a requirement for the viewer to move their head.

6.3.3 The VR environment

We simulated a conventional living room in VR. The simulated environment was modeled after the British Broadcasting Corporation Research and Development's (BBC R&D) user testing facility in Salford, U.K. (see Figure 6.1).

Figure 6.2 illustrates the virtual living room with TV content being displayed on the virtual TV set and a holographic sea turtle, when settled on the coffee table (6.2a) and when entering the viewing environment from three different points: the TV screen (6.2b), next to the TV screen (6.2c), and the ceiling (6.2d).

---

[1]https://en.wikipedia.org/wiki/Blue_Planet_II
[2]https://assetstore.unity.com/packages/3d/characters/animals/reptiles/sea-turtles-57461

Figure 6.1: The BBC R&D's physical user testing facility, Salford UK.

6.3.4   Procedure

Before the session, participants received the participant's information sheet, and completed a consent form and a basic demographic questionnaire. They received the experimental software in the form of a .apk file, a day before their scheduled session.

During the session, participants joined a researcher on a one-to-one Zoom call. They first received a verbal introduction to the study, then were asked to wear their VR headset. For each of the six ARTV episodes that they watched during their session, they would view the ARTV episode, remove their headset and respond to a set of questionnaire items (see Table 6.2) on a web browser using a laptop.

After wearing a VR HMD, a participant would find themselves sitting on a sofa in a virtual living room facing a virtual TV display. Additionally, they would see a sofa to their right-hand side, an armchair to their left-hand side, a coffee table between them and the virtual TV set, and a plant to the right-hand side of the TV display. Further to the right of the plant, there was a window with curtains half drawn, and to the left of the TV display there was a shut door (see Figure6.2).

After watching all six episodes, the moderator conducted a semi-structured interview. The interviews were recorded and subsequently transcribed. At the end of the session, participants were briefed and given the opportunity to ask questions. They received a £10 Amazon gift voucher, on completion of the session.

Figure 6.2: The virtual living room. Also depicted are the hologram settled on the coffee table, and the three starting points.

(a) Hologram settled on the coffee table.

(b) The TV screen starting point.

(c) The TV-adjacent starting point.

(d) The ceiling starting point.

Table 6.2: Three questionnaires that were answered by participants six times, once after each ARTV episode.

Questionnaire	Number of items
Film IEQ	24
Focused Attention	3
ARTVQ	6

6.3.5 Instruments

Participants answered the same set of questions after viewing each ARTV episode. These questions consisted of three parts with 33 items in total (see Table 6.2).

Film IEQ

The *Immersive Experience Questionnaire for Film and TV (Film IEQ)* is a 24-item questionnaire used for measuring immersion in video media [Rigby et al., 2019]. Rigby et al. [2019] developed the Film IEQ by adapting Jennett et al.'s [2008a] games-oriented *Immersive Experience Questionnaire (IEQ)*, making it suitable for a passive film viewing

experience. The Film IEQ consists of the following four factors:

1) *Captivation*: twelve items targeting the viewer's sense of enjoyment, interest, and motivation.

2) *Real-world dissociation*: three items concerned with the viewer's sense of awareness of their surroundings.

3) *Comprehension*: four items targeting the viewer's understanding of the concepts and themes of the video content.

4) *Transportation*: five items concerned with the extent to which the viewer felt like they were placed in the story-world.

Each of the 24 items in the Film IEQ is rated between one and seven and the final score is obtained by adding the score of all the 24 items.

We asked the participants to answer the Film IEQ items in relation to the video content that was displayed on the virtual TV screen.

Focused attention

O'Brien et al. [2018] developed the *User Engagement Scale - Short Form* as a practical way to measure user engagement. It consists of twelve items with the following four sub-scales:

1) *Focused Attention (FA)*: captures a feeling of being absorbed and losing track of time.

2) *Perceived Usability (PU)*: capturing the negative aspects of interacting with a system and the degree of control and effort expended.

3) *Aesthetic Appeal (AE)*: the attractiveness and appeal of the interface of a system.

4) *Novelty (NO)*: the curiosity and interest in the interactive task.

While PU, AE, and NO predominantly aim to capture aspects of a user's interaction with an interactive system, in our experience FA can be used, with minor modification in wording, to measure a viewer's engagement with passive TV-like audio-visual content (see Table 6.3). Each of the FA items is rated between 1: strongly disagree, 2: disagree, 3: neither agree nor disagree, 4: agree, and 5: strongly agree. The FA score is obtained by averaging the response to the items.

Table 6.3: Modifications made to the original Focused Attention (FA) sub-scale to adapt the items for our study (changes from underlined to boldface).

Original	Modified
I lost myself in this experience.	I lost myself in this **clip**.
The time I spent using Application X just slipped away.	The time I spent **watching this clip** just slipped away.
I was absorbed in this experience.	I was absorbed in this **clip**.

Similar to how we used the Film IEQ (see 6.3.5), we asked the participants to answer the FA items in relation to the video content that was displayed on the virtual TV screen.

ARTVQ

In order to synthesize a holistic view of the viewer's ARTV experience we devised a six-item *Augmented Reality Television Questionnaire (ARTVQ)*. Three items targeted the combined impression of the video plus hologram, while the other three items focused only on the hologram.

ARTV

> *Confusion:* to what extent was the combination of TV show and a holographic sea turtle confusing?
>
> *Distraction:* to what extent was the holographic sea turtle distracting from the TV show?
>
> *Enjoyment* to what extent did you enjoy watching the combination of TV with a holographic sea turtle?

Hologram

> *Interest:* to what extent were you interested in the holographic sea turtle?
>
> *Realism:* how real did the holographic sea turtle seem to you? (a modified version of Schubert et al.'s [2001] IPQ item 11)
>
> *Presence:* to what extent did you have a sense that the holographic sea turtle was present in the virtual living room? (a modified version of Schubert et al.'s [2001] IPQ item 1)

The last two items in the ARTVQ (*Realism* and *Presence*) were adapted from the *Igroup Presence Questionnaire (IPQ)* [Schubert et al., 2001], which is used for measuring the sense of presence in virtual environments [Schubert et al., 2001], due

Table 6.4: Participants' role in the media industry. Selection of more than one option was permitted.

Role	Consumer	Researcher	Enthusiast	Producer	Technologist	Writer	Engineer
Total	7	3	3	2	2	1	0

Table 6.5: Devices typically used by participants to watch TV content. Selection of more than one device was allowed.

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

to their relevance to our study. However, since our intention here was to target the hologram instead of the virtual viewing environment, instead of capturing the extent to which a viewer felt present in the virtual environment, we aimed to capture the extent to which the viewer felt that the hologram was present in the virtual environment with them. Therefore, the original IPQ item—*"In the computer generated world I had a sense of 'being there' "*—was rephrased as follows: *to what extent did you have a sense that the holographic sea turtle was present in the virtual living room?*

Similarly, we reworded another original item from the IPQ—*"How real did the virtual world seem to you?"*—as follows: *how real did the holographic sea turtle seem to you?*

6.3.6   Participants

Adult participants were recruited using social media (LinkedIn and Twitter) and electronic mailing lists. Our inclusion criteria for participants were: 1) be aged eighteen or over, 2) be a fluent or native English speaker, 3) have normal (or corrected to normal) vision, and 4) have no clinical hearing loss. Recruitment was incentivised with a £10 Amazon voucher.

Age and gender

Ten individuals (8 male, 2 female) opted to participate; four participants in the $25-29$ and three participants in the $50-54$ age groups. The other three participants were in the $18-24$, $35-39$, and $40-44$ age groups, one participant per group.

Experiences with TV and media

Table 6.4 presents participants' role in the media industry. Selection of more than one role was allowed. Six selected more than one role. The most frequently selected

Table 6.6: Participants' role in the AR industry. Selection of more than one option was permitted.

Role	Consumer	Researcher	Developer	Enthusiast	Content creator	Technologist	Engineer
Total	7	3	3	2	2	1	1

Table 6.7: Participants' reason for using AR. Selection of more than one option was permitted.

Reason	Gaming	R&D	Work	Social platforms	Socialising	Watching videos	None
Total	4	4	4	1	0	0	1

role was *consumer* (seven participants) followed by researcher and enthusiast (three participants each). No one selected the *engineer* role.

All participants reported watching a minimum of two hours of TV per week, with one watching at least three hours, three watching at least four hours, and four watching more than five hours. The most popular devices were TV set and desktop computer, with five participants selecting them. This was followed by laptop and mobile phone, with four participants selecting them. Eight participants reported consuming TV content on multiple devices (see Table 6.5).

Familiarity with, and use of, AR

Six participants selected more than one role. The most popular role was consumer, with seven of participants selecting it. This was followed by researcher and developer (three participants each). See Table 6.6.

Four participants selected more than one reason for using AR. The most popular reasons for using AR were *gaming*, *research & development*, and *work*, with four participants selecting them. One participant selected *social platforms*. One person selected *none*. No one selected *socialising with friends and family* or *watching videos* (see Table 6.7).

Eight participants reported some weekly AR usage, with four using AR under an hour, two using at least one hour, one using at least two hours, and one using more than five hours per week. The participant that selected more than five hours had, in the demographic questionnaire, clarified that the reported value corresponds to work-hours spent to create AR applications and when they are not developing AR applications they "rarely watch AR content".

The most popular device for consuming AR was mobile phone, with seven participants selecting it. The second most popular device HMD, with two participants selecting it. One participant did not select any AR devices (this participant reported not using AR at all). No one selected tablet or projection-based AR.

Table 6.8: Participants' role in the VR industry. Selection of more than one option was permitted.

Role	Consumer	Researcher	Developer	Enthusiast	Content creator	Technologist	Engineer
Total	8	4	4	3	2	2	0

Table 6.9: Participants' reason for using VR. Selection of more than one option was permitted.

Reason	Gaming	R&D	Work	Social platforms	Watching videos	Socialising
Total	7	7	5	3	3	1

Familiarity with, and use of, VR

Seven participants selected more than one role. The most popular role was consumer, with eight participants selecting it. The second most popular roles were developer and researcher, with four participants selecting them (see Table 6.8).

Eight participants selected more than one reason for using VR. The most popular reasons for using VR were gaming and research & development, with seven participants selecting them. The second most popular reason for using VR was work, with five participants selecting it (see Table 6.9).

All participants reported some weekly VR usage, with three using VR under an hour, one using at least one hour, two at least two hours, and four more than five hours per week.

6.4   Questionnaire Responses

6.4.1   Overview

Before looking at the results per condition, we first present an overview of the aggregate questionnaire results (see Table 6.10). To get a sense of the different aspects of

Table 6.10: Min, max, median, mean, standard deviation, maximum possible value, and *score* calculated for the questionnaires used in the study (see Section 6.3.5), presented in aggregate form. *Score* is calculated by dividing the mean by the maximum possible value for each column, then converting the result into a percentage. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson [1994].

	FA	Film IEQ	Confusion	Distraction	Enjoyment	Interest	Realism	Presence
Min	1.9	83.5	0.85	1.62	1.5	1.73	2.37	3.02
Max	4.9	147.5	6.35	7.78	7.67	8.57	7.7	7.85
Median	3.6	118.08	2.77	4.62	4.67	4.73	4.7	5.68
Mean	3.55	116.67	3.02	4.45	4.67	4.73	4.7	5.52
Std. dev.	0.51	9.76	1.41	1.33	1.42	1.51	0.94	0.99
Max possible	5	168	7	7	7	7	7	7
Score	71%	69.45%	43.14%	63.57%	66.71%	67.57%	67.14%	78.85%

Table 6.11: Chi-squared, degrees of freedom (DF), and p-value from Friedman tests carried out for questionnaires.

	FA	Film IEQ	Confusion	Distraction	Enjoyment	Interest	Realism	Presence
Chi-squared	9.317	2.128	3.113	4.007	4.018	6.536	1.633	4.916
DF	5	5	5	5	5	5	5	5
P-value	0.097	0.831	0.683	0.549	0.547	0.258	0.897	0.426

participants' experience, we calculate a new value—*score* by dividing the mean by the maximum possible value of the scale and converting it to a percentage. This score is presented in the final row in Table 6.10. We propose that the score can provide a broad-stroke overview of how the ARTV experience under investigation (regardless of the changes made to the hologram's entry point and behaviour) was perceived by the participants.

The relative closeness of the score for the FA and the Film IEQ could indicate that both instruments are measuring comparable aspects of the viewing experience. Furthermore, The relatively high score for both the FA (71%) and the Film IEQ (69.45%) can be an indication of a high level of attention and immersion experienced by the participants when viewing the ARTV content.

The *score* values of the ARTVQ items (see Table 6.10) indicate that overall, the participants enjoyed watching the ARTV content (enjoyment = 66.71%), and did not find the ARTV content too confusing (confusion = 43.14%). However, they found the hologram to be distracting from the TV show (distraction = 63.57%). The distraction caused by the hologram is possibly not independent from the fact that the participants found the hologram to be interesting (interest = 67.57%) and highly 'present' in their environment (78.85%). The hologram was also perceived to be realistic (67.17%).

6.4.2 Statistical significance

Friedman tests were carried out to compare the questionnaire results for the six conditions. No significant difference was found between the conditions. Table 6.11 presents Chi-squared, degrees of freedom, and p-values.

6.4.3 Questionnaire results per condition

Table 6.12 presents ranked scores for each questionnaire per condition. Further detail regarding individual questionnaire response statistics, along with line plots can be found in Appendix B.

Table 6.12: The six experimental conditions (Cond.) with starting point (Entry), and hologram behaviour (behaviour) ranked per all questionnaire: Focused Attention (FA), Immersive Experience Questionnaire for Film and TV (Film IEQ), and the six ARTVQ items, namely, Confusion (Conf.), Distraction (Dist.), Enjoyment (Enjoy.), Interest, Realism (Real.), and Presence (Pres.). Confusion and distraction ranks have been inverted (i.e., high score=low rank). A lower number represents a better ranking. Green and lavender colour boxes indicate best and worst rankings, respectively.

Cond.	Entry	Behaviour	FA	Film IEQ	Conf.	Dist.	Enjoy.	Interest	Real.	Pres.
1	TV screen	Settle	5	4	4	1	4	2	1	4
2	TV screen	Continuous	3	3	1	6	1	1	1	3
3	TV-adjacent	Settle	4	5	2	2	2	5	6	5
4	TV-adjacent	Continuous	2	2	5	4	5	4	4	6
5	Outside fov	Settle	6	6	6	3	6	6	4	2
6	Outside fov	Continuous	1	1	3	6	3	3	3	1

FA and Film IEQ scores have identical rankings across all experimental conditions except for conditions 1 and 3 (Table 6.12). Top three highest ranking conditions for FA and Film IEQ (i.e., conditions 6, 4, and 2) consist of a continuous movement of the hologram in the viewing environment; this can be an indication that a continuous movement is likely to draw more attention and induce higher levels of immersion, when compared to conditions where the hologram settles in an area adjacent to the TV screen. Amongst conditions with continuous movement, the amount of attention and the sense of immersion decreases as the hologram entry point approaches the TV screen; with the highest and lowest values at outside the field of view and the TV screen, respectively. When the movement changes to settle, the outside the field of view entry point drew the least amount of attention and induced the least amount of immersion amongst all six conditions (Table 6.12).

Condition 2 (entry from the TV screen with continuous movement) performed consistently well, obtaining the first place for confusion (i.e., the least amount of confusion), enjoyment, interest, and realism. Conversely, condition 5 (entry from outside the field of view and settling on the coffee table) performed consistently badly, obtaining the last place for focused attention, immersion (Film IEQ), confusion (i.e., the highest amount of confusion), enjoyment, and interest (see Table 6.12).

6.5   Interviews

We used a *deductive content analysis* technique [Mayring, 2014] to analyse interview transcripts. Initially, a code-book with 14 codes were generated from the semi-structured interview protocol. Through an iterative process, the codes were subsequently grouped into the following categories: 1) suitable entry points, 2) hologram's behaviour in the

viewing environment, 3) the holographic sea turtle, 4) interaction, 5) presentation of story-related holograms outside the TV screen, 6) watching TV content in VR, and 7) the virtual living room.

6.5.1 Suitable entry points

We answer the following question regarding a hologram's entry point:

RQ: *In the context of an ARTV experience, which areas are more suitable for a lifelike programme-related AR artefact to enter the viewing environment?*

Overall, the responses highlight that the "right" entry point may depend entirely on the role of the hologram in the story. For instance: "it really depends on . . . the context of the story you're trying to tell" [P6] and "I would prefer [the hologram] to come out of the actual story." [P7]

We group the responses in the following three sub-categories: 1) coming out from the TV screen, 2) emerging from the environment, 3) other methods to introduce a hologram, such as fading-in and having the hologram already present before the TV show starts.

Coming out of the TV screen: Five participants reported coming out of the TV screen as their favourite entry point for the holographic artefact. However, the success of this condition appears to "very much depend on what you're showing on the TV screen" [P8] at the time when the hologram enters the viewing environment (or exists the TV display).

In cases where the hologram came through the TV screen when there was a sea turtle displayed on the TV screen, two participants thought it "felt right" [P2], and "felt more real and amusing" [P1], whereas another participant found it distracting and jarring, due to the hologram occluding the TV screen: "there's always going to be a couple of moments where you cannot see what's going on [on the TV screen] . . . that's quite distracting." [P5]

Other comments suggest that this entry point would be suitable only when the 2D turtle in the TV programme is itself perceived to be entering the viewers' environment in holographic form: "If the sea turtle would come out of the screen in the same direction and same rotation . . . that would be really nice . . . because then you have this sense of this sea turtle [from] the TV show [is] now in my living room." [P10]

When there was not a sea turtle present on the TV screen this entry point was considered to be "not really pleasant, because it just didn't match with the [TV] content" [P10].

Emerging from the environment: Three participants reported the wall adjacent to the TV screen and two reported the ceiling as their favourite entry points.

While one participant reported that the wall adjacent to the TV screen and the ceiling "felt unnatural" and "like a glitch" [P1]; the rest of the participants reported these conditions to be less distracting than when the hologram came out from the TV screen. For instance: "it was least disruptive when it . . . entered . . . almost adjacent to the television." [P3]

Other methods to introduce a lifelike hologram: As opposed to entering the environment dynamically from a designated entry point, P4 suggested either fading it in or having the hologram present before the TV programme starts: "I would prefer to see it suddenly appear, like fade-in on the table, or be already on the table . . . or in an aquarium."

6.5.2 Hologram's behaviour in the viewing environment

We answer the following question regarding a hologram's behaviour in the viewing environment:

RQ: *In the context of an ARTV experience, how should a programme-related holographic artefact behave in the viewing environment?*

Similar to *hologram's entry point* (Section 6.5.1), the responses here also suggest that the "right" behaviour "really depends on the context." [P1]

We group the responses into the following five sub-categories: 1) landing on the coffee table, 2) continuous motion across the living room, 3) hologram as part of the story, 4) reactive to the environment, and 5) controllable by the viewer.

Landing on the coffee table: Five participants preferred having the sea turtle land on the coffee table, instead of having it move around the living room continuously. It tended to "invite inspection . . . my instinct was . . . to interact and find out more" [P3]. It was also perceived to "give you a lot of options to watch it more closely" [P1] and to "provide the least amount of disruption." [P2]

Continuous motion across the living room: Three participants preferred the continuous motion of the hologram[3]. The main reason for disliking the continuous motion was distraction from the TV screen. For instance, by causing the viewer "to pay the least amount of attention to the [TV] clip and wondering what I missed." [P2]

---

[3]Two participants reported not having a preference between continuous movement and settling on the coffee table.

However, P10 "perceived [the continuous animation] as more pleasant, mainly because ... a TV show is an ongoing process ... with the turtle [on the] table, in a quiet position, I feel like the [TV show has stopped] ... the [continuous] animation was way more comfortable."

Hologram as part of the story: Three participants suggested that the hologram's behaviour should be either similar to the behaviour of its counterpart on the TV screen, or it should be driven by the needs of the narrative. For instance: "[would] be great if [the hologram is] part of the story [and] if there was a reason [for its existence and behaviour], [for example,] if she was being chased and she came out of the screen at you." [P7]

Reactive to the environment: Two participants highlighted that the hologram needs to react to the viewer and the elements in the viewing environment: "I'd want it to behave like the actual creature, [for example,] go off looking for food maybe, or come up to you and start interacting" [P4] and "I prefer the holographic sea turtle to interact with the physical entities within my surroundings." [P9]

Controllable by the viewer: P4 suggested that the hologram "needs to be controllable by the [viewer] ... I wanted to control it."

6.5.3   The holographic sea turtle

We answer the following question regarding the holographic sea turtle:

RQ: *How was the holographic sea turtle perceived by the viewers, and how did it affect their experience?*

While the negative comments can all be grouped under the umbrella of distraction, the positive comments highlighted three main points: 1) an increased sense of immersion and engagement, 2) a bridge to the story-world, and 3) getting a better sense of the animal's size and proportions.

Distraction: All participants reported distraction as the main negative effect of the hologram. For instance, P2 reported feeling that they may have missed details from the show "because [they were] paying closer attention to the detail of the turtle model" and P7 said they were "more interested in seeing what was happening [regarding the hologram] rather than actually listening to the show."

An increased sense of immersion and engagement: Four participants reported an increase in their sense of engagement with the story. For instance, P5 said that the

hologram "makes it more immersive, intriguing, and probably more memorable" and P10 said they "felt more immersed into the environment."

Furthermore, P6 reported that the addition of a hologram would encourage them to re-watch the previously seen TV content: "I've never really found a reason to go back and [re-]watch [a nature documentary] . . . if there were [holograms] on top of it . . . I'd go back [and re-watch the show]."

A bridge to the story-world: P2 articulated the role of the hologram as connecting the viewer to the story-world, by creating a bridge between the TV display and the viewer: "[the hologram] provides a bridge to [increase my connection] to the space that's being depicted . . . I think it's subtle . . . and possibly even subconscious. But I think it provides a bridge to get into the content."

Getting a better sense of the animal's size and proportions: Getting a sense of size, proportions, and how the actual animal swims was pointed out by P1 as the main positive aspect of having a holographic sea turtle: "you have the object in context with you, you see its size; you can see how they are made up from all the angles. That's the best thing about it."

6.5.4   Interaction

We answer the following question regarding viewers' interaction with a hologram:

RQ: *In the context of an ARTV experience, how, if at all, would the viewers want to interact with a programme- related hologram?*

While two participants preferred to consume content in a passive mode, the rest of the participants ($n = 8$) reported a willingness to interact. The three themes that emerged from the responses were: 1) interaction for the sake of interaction, 2) having an interactive mode triggered by the viewer, and 3) interaction to get more information.

Interaction for the sake of interaction: Three participants reported a desire to interact with the hologram, as if it was a *real* creature. For instance, P6 "was very tempted to go out and try and touch it" and P3 reported that "as soon as it landed in front of [them], all [they] wanted to do was [to] pick it up."

An interactive mode: P5 reported that they would expect to interact with the hologram either when the TV programme is paused, or when "the TV show had some space" such that while they were interacting they would not miss out on the story.

P4 reported that as soon as they start their interaction, they expect the programme to transform from a *"directed experience"* to one where the viewer can manipulate

the objects; and back to the original "directed experience" once the viewer is finished interacting.

Interaction to get more information: P1 suggested to use the hologram as a way to deliver extra information about the real animal: "when I click on it [I want to see] where they live, how endangered they are, is there a place near me where I can go and watch the real animal ... whatever helps me to learn more about the turtle."

6.5.5 Presentation of story-related holograms outside the TV screen

We answer the following question regarding viewers' thoughts in relation to programme-related holographic content:

RQ: *What did the participants think about having story-related holograms outside the TV display?*

Participants were asked to generalise from what they saw during the experiment and share their thoughts regarding the possibilities provided by combining story-related holograms with a TV show.

We present the comments in three categories: 1) enriching a conventional TV programme, 2) a new form of storytelling, and 3) eliminating the TV display.

Enriching a conventional TV programme: Five participants reported that story-related holograms would be suitable for enriching a conventional TV viewing experience. For instance, P9 reported that "it can encourage people to watch TV more and to engage in TV content more" and P6 suggested that holograms "bring people into the experience ... by bringing the experience out to them."

A new form of storytelling: Two participants suggested that they "could easily see the technology being used to add components to enhance the narrative" [P2], and that "it creates new opportunities for storytelling." [P8]

Eliminating the TV: P1 suggested that a mechanism should be provided for the viewer to select how much of the narrative's components they want to view in holographic mode. They suggested that when all the relevant components are presented outside the TV screen, the viewer can ignore the TV screen and follow the story mid-air in holographic mode and get much more from the experience.

6.5.6 Watching TV in VR

We answer the following question regarding viewers' experience of watching TV in VR:

RQ: *How did the participants find the experience of watching TV content in VR?*

The comments in this section reflect participants' thoughts in relation to viewing TV in VR in general and without holographic augmentations.

Eight participants only made positive comments, characterising watching TV in VR either comparable or better than the traditional way of watching TV. For instance, P1 said "you feel much more immersed. It's like experiencing the video in a theater"; P2 said "[it's] less distracting than watching [TV] in the physical world"; P4 said that watching TV is "basically one of the reasons why I bought a VR headset."

Two participants pointed that out wearing prescription glasses while wearing a VR headset and the weight of the VR headset was the main barriers for them to watch TV in VR over a long periods.

6.5.7   The virtual living room

We answer the following question regarding the viewers' comments in relation the virtual viewing room:

RQ: *What did the participants think about the virtual living used in the experiment?*

Seven participants made only positive comments, two of which reported preferring the virtual living room to their own physical space, because it was "more spacious" [P3] than their own living room and because the sofa was "nice, comfortable, more comfortable than the chair I'm sat on." [P6]

Two participants highlighted that the living room wasn't real enough to look completely real. For instance, P10 said that it looked "like a typical lab environment . . . it is too clean to be fully realistic."

P4 reported that being "completely locked off inside a virtual space can be a bit awkward [when] trying to use your physical objects". They suggested a feature such as Oculus' Passthrough—where the VR user's view is augmented using real-time camera feed from their physical environment—as a potential fix to this issue.

6.6   Discussion

In this section we discuss our observations and learning regarding 1) the presentation of lifelike creatures using AR during a TV viewing, 2) watching TV in VR and VR's potential as a platform to deliver ARTV experiences, and 3) ARTV experiences in social

settings. In each section, we highlight a few of the challenges and opportunities that future work can consider.

6.6.1 Lifelike creatures as AR companions

In this paper we focused on a lifelike hologram of a sea turtle which, from the perspective of a viewer, entered the viewing environment and either settled near the TV screen or swam around continuously. We investigated the impact of these two different behaviours and the hologram's *entry point* on the viewer (see Section 6.3.2). Addition of such artefacts is likely to create engaging and immersive experiences, however, to create successful experiences that last beyond an initial novelty phase, content creators and broadcasters may want to consider addressing challenges regarding distraction, confusion, personal preferences, and the role of holograms in the underlying narrative. We formulate some of these challenges as research questions to be considered by future work:

1) What are the ways in which a viewer's distraction and confusion—caused by the introduction/existence of lifelike holographic artefacts—can be minimised, while maximising their immersion, enjoyment, and interest?

2) What are the ways in which viewers' preferences can be elicited, stored, and applied to modify the behaviour of lifelike holograms?

3) How can modifications based on personal preferences be implemented in cases with more than one primary viewer?

4) What are the changes that need to take place in the production and distribution pipelines to enable the creation of ARTV content where AR assets play an integral role in the narrative?

6.6.2 Watching videos in VR: a potential direction for practical ARTV

The focus of the ARTV community (including researchers and broadcasters) has been to use AR technology in combination with a physical TV set (see Section 6.2.1 for examples). In cases where VR has been involved, the reason has been to use VR as a testbed to facilitate conducting user studies, with the goal of implementing the final system outside VR (see Section 6.2.2). However, this attitude will likely change in

the future, as VR finds its audiences as a medium to watch 2-D imagery (e.g., TV content, and film). Three participants reported using VR for watching videos, and one participant said that watching TV content was one of the main reasons for them to buy a VR headset (Section 6.5.6). We propose that future work should consider the challenges and opportunities of augmenting TV content using holograms inside virtual environments, as a medium in its own right, and not just as a way of facilitating data collection for user studies.

6.6.3   TV viewing is typically social, so is ARTV likely to be

Our investigation in this work was limited to the impact of changes in hologram presentation for a single viewer. Similar to TV consumption, ARTV content will likely be consumed both by a single viewer and by multiple viewers in a social setting. This introduces challenges and opportunities for broadcasters and content creators. For instance:

1) To what extent the experiences need to be personalised according to each individual viewer's preferences vs. presenting a single experience where the viewers can view (or participate in) given their own devices?

2) How should interruptions from passers-by and bystanders, such as a roommate suddenly coming into the living room during an ARTV experience, be managed? For instance, how should the content adapt to the changes made in the environment?

6.7   Conclusion

We created a proof-of-concept ARTV programme in VR, where a lifelike sea turtle was presented as a hologram in the viewing environment while the viewer was watching a nature documentary clip about sea turtles. In a repeated-measures user study ($n = 10$), we investigated the impact of changing the starting point and the movement behaviour of the holographic sea turtle on the viewers' immersion, focused attention, confusion, distraction, enjoyment, interest in the hologram, perceived realism of the hologram, and its perceived presence in the viewing environment.

We found that the highest level of enjoyment, interest in the hologram, and perceived realism of the hologram, coincided with the highest level of distraction and the lowest

level of confusion, in a condition where the hologram started from the *TV screen* and *moved continuously in the field of view*; The highest level of confusion and the lowest level of enjoyment and interest were experienced when the hologram started its path from *outside the field of view* and *settled* adjacent to the TV set (i.e., on a coffee table).

Additionally, our findings from the interviews suggest that the *"right"* behaviour for this type of holographic content likely depends on personal preferences in addition to the perceived role of the hologram in the underlying narrative and TV content.

This study highlights the importance of personal preferences regarding ARTV content and provides a method for investigating the broader ARTV design space.

Chapter 7

Conclusion

This chapter contains discussion, limitations, future work, and conclusion.

7.1 Discussion

As consumer AR devices become readily commercially available, researchers and broadcasters are investigating the ways in which AR can be used to enhance/transform a conventional TV viewing experience—a concept referred to as Augmented Reality Television (ARTV). For ARTV to become a mass medium of communication, challenges regarding production, distribution, and consumption of ARTV content need to be addressed [Saeghe et al., 2020]. However, while the challenges regarding production and distribution of ARTV content remain largely technical, the consumption of ARTV content introduces challenges that relate to viewing habits and expectations of TV audiences. This thesis focused on the consumption of ARTV content, aiming to help content creators and broadcasters in understanding some of the design possibilities afforded by ARTV, and the ways in which these design decisions can impact the viewers' experiences of ARTV content.

The following provides a summary and the final versions of the design space concepts that were identified throughout this thesis.

7.1.1 ARTV Design Dimensions

Six design dimensions emerged from an analysis of prior work, through a systematic literature review; these were *abstraction*, *interaction*, *time*, *context*, *display*, and *editorial control* (see Chapter 2). To capture a wider spectrum of the theoretical design

space underlying ARTV experiences, we proposed a higher level classification, consisting of two overarching categories of *content* and *non-content*, and the relationships between various elements within and between each category (see Table 3.9). A *cheat sheet* created based on this categorisation was subsequently used in a user study to aid participants in conceptualising ARTV scenarios (see Appendix A.1.3). The analysis of these scenarios, in turn, resulted in discovering new concepts and the categorisation was subsequently updated. The rest of this section presents a summary of the final update to these dimensions.

ARTV design space dimensions were grouped into four overarching categories: *content*, *people*, *space*, and *objects*.

Content:

Consisting of AR content and/or TV content, this category contains the following dimensions and aspects:

- Genre: for instance, a children's show, an educational programme, a game show, an escape room experience, sports, a dance show, and documentaries.

- Visual display: consisting of unframed; extended, multiple, and single frame; AR delivering TV content; and 360° displays.

- Dependency: independent, dependent, additional.

- Time: asynchronous, synchronous intermittent, and synchronous continuous.

- Broadcast mode: live and on-demand.

- Modality: auditory, visual, haptic, and olfactory.

- Representation: photorealistic, cartoonish, and everything in between.

People:

Consists of all the dimensions related to people:

- Viewers: single vs. multiple, primary viewer vs. bystander vs. passer-by.

- For multiple viewers: concurrent vs. non-concurrent, and co-located vs. at-a-distance viewing.

- Social group: friends and family vs. strangers, and the social interactions between people while consuming content.

- Interaction with content: display-level, game-level, and story-level; editorial control.

- People's familiarity with space and objects.

- Demographics.

- Movement: moving vs. static ARTV consumption.

Space:

Consists of all the dimensions that relate to the viewing environment:

- The venue: public vs. private; indoors vs. outdoors.

- Purposeful vs. non-purposeful augmentation of space.

- Generation/modification of content based on the features of the viewing environment.

Objects:

Consists of all the dimensions that relate to objects:

- Purposeful vs. non-purposeful augmentation of objects.

- Generation/modification of content based on features of the objects present in the viewing environment.

- Integration of objects in the ARTV experience.

While it would be difficult to imagine extending the aspects of some of the dimensions (e.g., modality may not be able to be expanded beyond five aspects—corresponding to humans' five senses), some dimensions (e.g., genre and broadcast mode) may be expanded in the future. Broadcast mode, for instance, may need to be extended to accommodate for emerging ARTV experience, since neither live nor on-demand aspects are completely fit-for-purpose in this respect; especially considering the challenges regarding converting a live ARTV broadcast to an on-demand consumption model, it

may be necessary to invent another model that sits between live and on-demand that is more suitable for interactive ARTV content.

Note that there could be different ways of categorising some of the dimensions, based on how one interprets them. For instance, regarding viewers' familiarity with the viewing environment and the objects present in that environment, the above classification assumes these as a dimension belonging to the "people" category. However, one could also classify them as dimensions belonging to the space and the objects categories by modifying the wording. For instance, positioning the space dimension in the space category would result in *space: familiar vs. unfamiliar*, and similarly for the object category *object: familiar vs. unfamiliar*. The important point remains that the designers of ARTV content should consider the degree to which their audiences are familiar with the viewing environment and the objects within it, while engaging with an ARTV experience. For instance, an ARTV experience may be designed for the viewers at home (i.e., a familiar space with familiar objects), or it could be designed for a commercial escape room, where participants are typically unfamiliar with the space and the objects within it.

7.1.2 ARTV: A New Storytelling Vehicle

The overarching contribution of this thesis is to provide content creators and broadcasters, the target audience of this thesis, with a set of methods and practical tools to create content that takes advantage of the affordances of AR in the context of audio-visual content viewing. The thesis provided a set of design space dimensions operationalised using a cheat cheat (see Chapter 3); a conceptual framework for ARTV experiences in the living room (see Chapter 4), highlighting nine specific types of ARTV experiences (see Figure 4.4); and observations regarding the impact of augmentations on the viewers (see Chapters 5 and 6).

Our findings suggest that while augmenting conventional TV content can create engaging experiences (see Chapter 5), it may be best to incorporate the augmentations into the underlying narrative, which may only be possible if the content is created with the augmentations in mind, rather than adding augmentations after TV content has already been created (see Chapter 6). We found that when many augmentations are present, some viewers may miss parts of the content or the underlying narrative (see Chapter 5), and that there may not be a best way to present off-screen augmentations; instead, viewers may have various preferences, for instance regarding where, when, and how they want to consume the augmentations (see Chapter 6).

An observation of this thesis is that the consumption side of content viewing has multiple aspects, including two related topics of viewing conventions and viewers' attention: conventional media platforms, such as TV and cinema benefit from existing (and perhaps evolving) conventions that have been developed through decades of experimentation. This results in the audiences' knowledge of what to expect when consuming content. In its most basic form, it consists of the TV (or cinema) viewers' understanding that they need to look in the direction of a 2D screen (and listen to the sounds), to engage with the content and comprehend the underlying narrative; the director will then direct their visual attention on the 2D screen, according to the requirements of the story and director's own style of storytelling. These conventions get more nuanced, of course with regards to each specific genre. For instance, the way in which a news item is presented would likely be different from the way in which a murder mystery unfolds.

Currently there is a lack of such conventions for ARTV content—viewers do not know what to expect. Since content can be displayed anywhere in the viewing environment, the viewers may not even know where to look. This, more than ever, emphasises the role of content creators as directors of viewers' attention. While this provides a challenge for content creators and viewers alike, it also provides novel opportunities to expand our storytelling capacities through experimentation that aims to develop new sets of conventions for content that can be interactive, can be displayed anywhere, and can be a mixture of 2D displays and 3D artefacts. For instance, consider an ARTV show where the experience transforms throughout its presentation. Imagine a programme that starts with the viewer(s) watching content in their living room on a physical TV set, then some of the elements of the story appear as on-screen augmentations, for instance to enable some form of interaction from the viewer(s). Then necessitated by the underlying narrative, a protagonist leaves the screen and appears as an off-screen augmentation, for instance to escape from something or someone in the story-world. The viewer(s) follow the protagonist to another area of the room where an important element of the story unfolds on a virtual 2D screen. Then the living room is transformed to the story world, using augmentations that match the theme of the show to maximise viewers' immersion in the story. Finally, as the protagonist overcomes their challenges, the story resolves and the living room goes back to normal.

7.2 Limitations

Overall, our investigations were constrained by time and budget, which in turn imposed limitations on the studies and the findings. For instance, after identifying an underlying design space for ARTV, and providing a method to support practitioners, we conducted an exploratory study which was then followed by a deeper investigation into one specific design dimension: visual display. This limits our concrete findings regarding each individual design dimension. However, the method provided in this thesis can be used to target each individual design dimension (or combine a few dimensions) for deeper explorations.

The rest of this section discusses limitations regarding participants, study design, and using a design space.

7.2.1 Participants

In Chapter 5, we recruited internally at the BBC R&D. It is likely that our participants were biased in their attitude towards new technology, as typically people in such jobs are more interested in novel technology than the general public. Although their knowledge of novel technology, such as AR headsets may vary from that of the general public, our participants' TV viewing behaviour is likely similar to that of the general public.

In Chapter 3 we recruited from a pool of professionals in the immersive field. Since recruitment was conducted via social media (Twitter and LinkedIn) and mailing lists, our access was limited to our social networks. Recruiting from different pools would result in different ARTV scenarios, which in turn could result in identification of other dimensions. For this reason, we proposed the method of recruiting practitioners and asking them to write ARTV scenarios as an effective way of identifying new dimensions.

In Chapter 6 we recruited participants who had access to an Oculus Quest 2. This limits our findings to a specific user group. However, while there are differences in device characteristics (e.g., field of view, weight, tethered vs. untethered), we do not expect between-device differences to impact content consumption in VR, especially as VR headsets advance towards providing more comfortable experience for their users.

The limited number of participants in studies presented in Chapters 5 ($n = 12$), 3 ($n = 10$), and 6 ($n = 10$) limits our ability to make concrete assertions about the impact of programme-related holograms and their associated design decisions. This limitation, particularly as regards Chapters 3 and 6 was a consequence of a global pandemic (COVID-19), and in general reflect a relatively small pool of participants that we were

able to recruit from.

7.2.2   Study design

The study in Chapter 6 consisted of a within-subject design. This likely resulted in the novelty of the experience wearing off after the first couple of episodes. A between-subject design of a comparable study—with a larger sample size—is likely to provide more precise insight. However, a between subject design typically requires a larger sample size. As discussed above, we were unable to recruit a large number of participants. Furthermore, to reduce the length of the sessions (to under one hour) and avoid fatigue and discomfort for the participants, we split a nature documentary clip into six short episodes. This intermittent presentation of the show likely distracted from the narrative.

7.2.3   Using a design space

While our aim has been to provide a tool to support practitioners in navigating the underlying ARTV design space, there are limitations associated with this approach. For instance, given that there are likely other possible design dimensions, not yet mapped in our design space, asking practitioners to focus solely on our design space may limit their imagination. However, we found that practitioners used our cheat sheet in a flexible way, taking advantage of the prompts to choose a specific aspect for further exploration and coming up with aspects that were missing from the cheat sheet. Which in turn resulted in updating the design space and associated cheat sheet.

7.3   Future work

In this section we provide discussion and propose potential avenues for future work regarding exploration of other dimensions and exploring ARTV experiences beyond the living room.

7.3.1   Exploring other dimensions

In Chapter 6, we investigated two aspects of a hologram's movement behaviour in the context of our display dimension. There are many other dimensions that are open for

explorations. For instance: How does genre impact viewers' preferences of ARTV content? A dichotomy to explore could be the difference between an information-focused genre, such as news, and a narrative-focused genre, such as a nature documentary. For instance, ARTV may be a suitable candidate to maximise information delivery in the context of a news programme, for example by presenting statistics, charts, etc. outside the TV display during a newscast. However, in the context of a drama programme, providing extra information, for instance about an actor or a location, while the programme is running may distract the viewers from the narrative. In this context, ARTV may be a suitable candidate to increase immersion, for example by augmenting the viewing environment to better match the mood portrayed in the drama, or by having elements from the drama to "come out" or to "jump out" of the TV screen to elicit a certain emotional response.

Another dimension to explore is multiple primary viewers. This dimension is especially important when considering interaction. Who has the right to modify (interact with) content, and how do multiple primary viewers negotiate their right to overwrite each other's decisions?

7.3.2 ARTV beyond the living room

Given the expected availability of everyday AR glasses in the future, high speed connectivity to the Internet, and the availability of streaming services, it is reasonable to expect to see ARTV experiences that deliver content "on-the-go". Taking ARTV beyond a conventional living room experience can be a next step in the field of ARTV.

This introduces new challenges, for instance regarding the impact of bystanders and passersby—people who may be able to view TV content but not augmentations—and public spaces on a primary viewer's experience. How does the presence of bystanders (and passersby) impact a viewer of ARTV? This question is particularly interesting in scenarios where an ARTV viewers' role is transformed from a mere viewer to an active participant. For instance, in the context of an ARTV game show. How do cultural norms impact an ARTV participant's behaviour (and interactions) in the presence of bystanders and passerby?

Another relevant question in this context is regarding privacy. In scenarios where the underlying system (or indeed the ARTV viewer them self) has the capability to readily augment and modify the environment, to what extent should the ARTV programme and its viewer be able to augment (modify the appearance of) a public space and the people (and object) in it? Especially, given that the ARTV viewer is likely able to record and

store an ARTV experience, the role of bystanders and passersby, their awareness of their role, and their consent regarding their participation in an ARTV experience need to be addressed.

Another example is in the context of ARTV experience in moving vehicle. As the society moves towards accepting autonomous vehicles, the next "conventional TV viewing environment" may be inside of an autonomous vehicle. Addressing the challenges of presenting ARTV content in an autonomous vehicle, such as motion sickness, can ensure the availability of content when the underlying technology becomes readily available.

7.4   Conclusion

In this thesis, we explored three research questions:

RQ1 What is the underlying design space of ARTV?

RQ2 How can we support practitioners to navigate this design space?

RQ3 What are viewers' perceptions of ARTV?

To answer RQ1, we first conducted a systematic literature review ($n = 42$ papers) and identified six design dimensions commonly used in ARTV work (Chapter 2). We then refined this design space based on the findings of a user study ($n = 10$), where we asked practitioners to write ARTV scenarios and subsequently interviewed them (Chapter 3). In conclusion, we were able to identify an underlying design space for ARTV, and found it to be a useful tool to support practitioners.

Regarding RQ2, we created a "cheat sheet"—presenting design dimensions as questions—and evaluated the utility of the cheat sheet and its usefulness in conceptualising novel ARTV scenarios, in a user study with $n = 10$ practitioners (Chapter 3). Our findings indicate that the cheat sheet can be useful in conceptualising novel ARTV scenarios. We observed that an exercise of asking practitioners to write ARTV scenarios can be an effective way to refine and update the design space. Our work suggests that the design space presented in this way can be a useful addition to the broadcasters' toolkit.

Regarding RQ3, since a conventional TV viewing experience is traditionally associated with the living room, we first created a conceptual framework for ARTV experiences in the living room (Chapter 4). Focusing on nature documentaries, we

then examined viewer's perceptions in this context, in two user studies. In the first ($n = 12$), we added holograms—displayed via a Microsoft HoloLens device—to an existing nature documentary, displayed on a physical TV set. Our findings indicate that while viewers found the content engaging, for some viewers, the presence of holograms caused a degree of anxiety and fear of missing parts of the content and underlying narrative. In the second ($n = 10$), we focused on investigating the extent to which AR holograms can be used to bring living creatures out of the TV and into the viewing environment, and the impact this has on the viewers. We investigated this in the context of an ARTV experience where both the TV and the living room are virtual, and explored six points on the display dimension of the hologram. Our findings highlight the role of personal preferences and the perceived role of the hologram in the underlying narrative and TV content.

Overall, our findings highlight a need to carefully consider the impact of additional augmentations on the viewers. Furthermore, viewers should be given the ability to customise their ARTV experiences, based on personal preferences.

References

R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6): 34–47, Nov 2001. ISSN 1558-1756. doi: 10.1109/38.963459.

Ronald T. Azuma. A Survey of Augmented Reality. *Presence: Teleoper. Virtual Environ.*, 6(4):355–385, August 1997. doi: 10.1162/pres.1997.6.4.355. URL `http://dx.doi.org/10.1162/pres.1997.6.4.355`.

Ronald T. Azuma. The Most Important Challenge Facing Augmented Reality. *Presence: Teleoper. Virtual Environ.*, 25(3):234–238, December 2016. ISSN 1054-7460. doi: 10.1162/PRES_a_00264. URL `https://doi.org/10.1162/PRES_a_00264`.

Ronald T. Azuma. Making Augmented Reality a Reality. page JTu1F.1. Optical Society of America, 2017. URL `http://www.osapublishing.org/abstract.cfm?URI=AIO-2017-JTu1F.1`.

Ronald T. Azuma. The road to ubiquitous consumer augmented reality systems. *Human Behavior and Emerging Technologies*, 1(1):26–32, 2019. doi: 10.1002/hbe2.113. URL `https://doi.org/10.1002/hbe2.113`.

Caroline Baillard, Matthieu Fradet, Vincent Alleaume, Pierrick Jouet, and Anthony Laurent. 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, pages 67:1–67:2, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5548-3. doi: 10.1145/3139131.3141196. URL `http://doi.acm.org/10.1145/3139131.3141196`.

Chris Baume. The Mermaid's Tears, September 2017. URL `https://www.bbc.co.uk/rd/blog/2017-09-mermaids-tears-object-based-audio`.

Rik Bauwens, Hendrik Lievens, Maarten Wijnants, Chris Pike, Iris Jennes, and Werner Bailer. Interactive Radio Experiences. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, page 273–278, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450360173. doi: 10.1145/3317697.3323347. URL `https://doi.org/10.1145/3317697.3323347`.

BBC. His Dark Materials: Discover your Daemon, December 2019. URL `https://www.bbc.co.uk/rd/blog/2019-07-personalised-documentary-data-instagramifaction`.

Michel Beaudouin-Lafon. Designing Interaction, Not Interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '04, page 15–22, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581138679. doi: 10.1145/989863.989865. URL `https://doi.org/10.1145/989863.989865`.

Frank Bentley, Max Silverman, and Melissa Bica. Exploring online video watching behaviors. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, page 108–117, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450360173. doi: 10.1145/3317697.3323355. URL `https://doi.org/10.1145/3317697.3323355`.

Toni Bibiloni, Antoni Oliver, Cristina Manresa-Yee, and Javier Varona. Usability testing of a Hypervideo-based Interactive TV application. In *VI International Conference on Interactive Digital TV IV Iberoamerican Conference on Applications and Usability of Interactive TV*, page 127, 2015.

Mark Billinghurst, Adrian Clark, and Gun Lee. A Survey of Augmented Reality. *Foundations and Trends in Human-Computer Interaction*, 8(2–3):73–272, March 2015. doi: 10.1561/1100000049. URL `http://dx.doi.org/10.1561/1100000049`.

Oliver Bimber and Daisuke Iwai. Superimposing Dynamic Range. *ACM Trans. Graph.*, 27(5), December 2008. ISSN 0730-0301. doi: 10.1145/1409060.1409103. URL `https://doi.org/10.1145/1409060.1409103`.

Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA, 2005.

Jay David Bolter, Richard Grusin, and Richard A Grusin. *Remediation: Understanding new media*. MIT Press, 2000.

Jochen Braun. Vision and attention: the role of training. *Nature*, 393(6684):424, 1998.

Matthew Brooks. Creating a Personalised and Interactive Episode of BBC
    Click with StoryKit, April 2019. URL `https://www.bbc.co.uk/rd/blog/`
    `2019-04-object-based-media-click-interactivity-tv`.

Grigore C. Burdea and Philippe Coiffet. *Virtual Reality Technology*. John Wiley &
    Sons, 2003.

Teresa Chambel, V. Michael Bove, Sharon Strover, Paula Viana, and Graham Thomas.
    Immersive Media Experiences: Immersiveme 2013 Workshop at ACM Multimedia.
    In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13,
    page 1095–1096, New York, NY, USA, 2013. Association for Computing Machinery.
    ISBN 9781450324045. doi: 10.1145/2502081.2503831. URL `https://doi.org/`
    `10.1145/2502081.2503831`.

Teresa Chambel, Rene Kaiser, Omar A. Niamut, Wei Tsang Ooi, and Judith A. Redi.
    AltMM '16: Proceedings of the 1st International Workshop on Multimedia Alter-
    nate Realities. New York, NY, USA, 2016. Association for Computing Machinery.
    ISBN 9781450345217. URL `https://dl.acm.org/doi/proceedings/10.1145/`
    `2983298`.

Shivakumar Chandrasekaran and Umesh Kesavan. Augmented reality in broadcasting.
    In *2017 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*,
    pages 81–83. IEEE, 2017.

Hyunwoo Cho, Sung-Uk Jung, Suwon Lee, Young-Suk Yoon, Sangheon Park, and
    Hyung-Keun Jee. ARStudio: AR Broadcasting System based on Multiple RGB-D
    Cameras. In *2018 International Conference on Information and Communication
    Technology Convergence (ICTC)*, pages 1140–1143. IEEE, 2018.

Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. A Review of
    Overview+detail, Zooming, and Focus+context Interfaces. *ACM Comput. Surv.*,
    41(1), January 2009. ISSN 0360-0300. doi: 10.1145/1456650.1456652. URL
    `https://doi.org/10.1145/1456650.1456652`.

John Corpuz. 20 Remote Control Apps for Android Devices, July
    2019. URL `https://www.tomsguide.com/uk/us/pictures-story/`
    `494-android-tv-remote-apps.html`.

Jasmine Cox, Rhianne Jones, Chris Northwood, Jonathan Tutcher, and Ben Robinson. Object-Based Production: A Personalised Interactive Cooking Application. In *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '17 Adjunct, page 79–80, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450350235. doi: 10.1145/3084289.3089912. URL `https://doi.org/10.1145/3084289.3089912`.

Alan B. Craig. *Understanding Augmented Reality: Concepts and Applications*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2013. ISBN 9780240824086.

George A Cushen and Mark S Nixon. Real-time semantic clothing segmentation. In *International Symposium on Visual Computing*, pages 272–281. Springer, 2012.

Tiago De Gaspari, Antonio Carlos Sementille, Daniel Zuniga Vielmas, Ivan Abdo Aguilar, and João Fernando Marar. ARSTUDIO: A Virtual Studio System with Augmented Reality Features. In *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '14, pages 17–25, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-3254-5. doi: 10.1145/2670473.2670491. URL `http://doi.acm.org/10.1145/2670473.2670491`.

Marco de Sá and Elizabeth F. Churchill. Mobile Augmented Reality: A Design Perspective. In Huang W., Alem L., and Livingston M., editors, *Human Factors in Augmented Reality Environments*. Springer, New York, NY, USA, 2013. doi: 10.1007/978-1-4614-4205-9_6. URL `https://doi.org/10.1007/978-1-4614-4205-9_6`.

F. de Sorbier, Y. Takaya, Y. Uematsu, I. Daribo, and H. Saito. Augmented reality for 3D TV using depth camera input. In *Proceedings of the 16th International Conference on Virtual Systems and Multimedia*, pages 117–123, 2010. URL `https://doi.org/10.1109/VSMM.2010.5665956`.

Federico Debandi, Roberto Iacoviello, Alberto Messina, Maurizio Montagnuolo, Federico Manuri, Andrea Sanna, and Davide Zappia. 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*, volume 364, page 012048. IOP Publishing, 2018.

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

Jon Driver. A selective review of selective attention research from the past century.
    British Journal of Psychology, 92(1):53–78, 2001.

Jennica Falk, Johan Redström, and Staffan Björk. Amplifying Reality. In *Proceedings of
    the 1st International Symposium on Handheld and Ubiquitous Computing*, HUC '99,
    page 274–280, Berlin, Heidelberg, 1999. Springer-Verlag. ISBN 3540665501. doi:
    10.5555/647985.743879. URL http://dx.doi.org/10.5555/647985.743879.

Matthieu Fradet, Caroline Baillard, Anthony Laurent, Tao Luo, Philippe Robert, Vincent
    Alleaume, Pierrick Jouet, and Fabien Servant. MR TV Mozaik: A New Mixed Reality
    Interactive TV Experience. In *Adjunct Proc. of ISMAR 2017*, pages 155–159. IEEE,
    2017.

Jon Francombe. Vostok-K Incident: Immersive Audio Drama on Per-
    sonal Devices, October 2018. URL https://www.bbc.co.uk/rd/blog/
    2018-10-multi-speaker-immersive-audio-metadata.

David Geerts, Rinze Leenheer, Dirk De Grooff, Joost Negenman, and Susanne Heijs-
    traten. In Front of and behind the Second Screen: Viewer and Producer Perspectives
    on a Companion App. In *Proceedings of the ACM International Conference on
    Interactive Experiences for TV and Online Video*, TVX '14, page 95–102, New York,
    NY, USA, 2014. Association for Computing Machinery. ISBN 9781450328388. doi:
    10.1145/2602299.2602312. URL https://doi.org/10.1145/2602299.2602312.

David Geerts, Evert van Beek, and Fernanda Chocron Miranda. Viewers' Visions of the
    Future. In *Proceedings of the 2019 ACM International Conference on Interactive
    Experiences for TV and Online Video*, TVX '19, page 59–69, New York, NY, USA,
    2019. Association for Computing Machinery. ISBN 9781450360173. doi: 10.1145/
    3317697.3323356. URL https://doi.org/10.1145/3317697.3323356.

Sabiha Ghellal and Irma Lindt. Interactive Movie Elements in a Pervasive Game.
    Personal Ubiquitous Comput., 12(4):307–315, February 2008. ISSN 1617-
    4909. doi: 10.1007/s00779-007-0186-8. URL https://doi.org/10.1007/
    s00779-007-0186-8.

Simon Gibbs, Constantin Arapis, Christian Breiteneder, Vali Lalioti, Sina Mostafawy, and Josef Speier. Virtual studios: An overview. *IEEE multimedia*, 5(1):18–35, 1998.

David Gómez, Ana M Bernardos, and José R Casar. A system to enable level-of-detail mobile interaction with augmented media objects. In *2014 Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 346–351. IEEE, 2014.

Lauren Goode. Get Ready to Hear a Lot More About 'XR', May 2019. URL `https://www.wired.com/story/what-is-xr/`.

O. Grau and V. Vinayagamoorthy. Stereoscopic 3D Sports Content without Stereo Rigs. *SMPTE Motion Imaging Journal*, 119(7):51–55, Oct 2010. ISSN 2160-2492. doi: 10.5594/J17314.

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

Nick Hanson. Creating a Personalised and Interactive Episode of BBC Click with StoryKit, August 2019a. URL `https://www.bbc.co.uk/rd/blog/2019-07-personalised-documentary-data-instagramifaction`.

Nick Hanson. Where Next For Interactive Stories?, January 2019b. URL `https://www.bbc.co.uk/rd/blog/2019-01-interactive-drama-stories-branching-narrative`.

Nick Hanson. Storytelling of the Future, March 2019c. URL `https://www.bbc.co.uk/rd/blog/2019-02-storytelling-interactive-digital-drama`.

Sandra G Hart and Lowell E Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

Steven Henderson and Steven Feiner. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE transactions on visualization and computer graphics*, 17(10):1355–1368, 2011.

Jens Herder, Philipp Ladwig, Kai Vermeegen, Dennis Hergert, Florian Busch, Kevin Klever, Sebastian Holthausen, and Bektur Ryskeldiev. Mixed Reality Experience-How to Use a Virtual (TV) Studio for Demonstration of Virtual Reality Applications. 2018.

Gregory Hough, Ian Williams, and Cham Athwal. Measurements of live actor motion in mixed reality interaction. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 99–104. IEEE, 2014.

Simon Hudson and JR Brent Ritchie. Promoting destinations via film tourism: An empirical identification of supporting marketing initiatives. *Journal of travel research*, 44(4):387–396, 2006.

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

Kaitlyn Irvine. XR: VR, AR, MR - What's the Difference?, October 2017. URL `https://www.viget.com/articles/xr-vr-ar-mr-whats-the-difference/`.

Aric Jenkins. The Fall and Rise of VR: The Struggle to Make Virtual Reality Get Real, June 2019. URL `https://fortune.com/longform/virtual-reality-struggle-hope-vr/`.

Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, 2008a. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2008.04.004. URL `https://www.sciencedirect.com/science/article/pii/S1071581908000499`.

Charlene Jennett, Anna L Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9):641–661, 2008b.

Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 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, page 637–644, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450330695. doi: 10.1145/2642918.2647383. URL `https://doi.org/10.1145/2642918.2647383`.

Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 869–878, New

York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654. 2466112. URL http://doi.acm.org/10.1145/2470654.2466112.

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

Yuki Kawamura and Kazuhiro Otsuki. 3D free-viewpoint augmented reality on a second display synchronized with a broadcast program. In *Proceedings of the 4th International Conference on Communication and Information Processing*, pages 273–277. ACM, 2018.

Seoksoo Kim. Virtual studio system for augmented reality & Chroma key processing. In *13th International Conference on Advanced Communication Technology (ICACT2011)*, pages 762–765. IEEE, 2011.

Soonchoul Kim, Moonhyun Kim, Kuinam J Kim, Bumsuk Choi, and Jinwook Chung. A scheme of AR-based personalized interactive broadcasting service in terrestrial digital broadcasting system. *Cluster Computing*, 20(3):2349–2354, 2017.

Naoki Kimura and Jun Rekimoto. 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*, CHI '18, pages 427:1–427:10, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3174001. URL http://doi.acm.org/10.1145/3173574.3174001.

Bettina Laugwitz, Theo Held, and Martin Schrepp. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*, pages 63–76. Springer, 2008.

Hojun Lee, Gyutae Ha, Sangho Lee, and Shiho Kim. A mixed reality tele-presence platform to exchange emotion and sensory information based on MPEG-V standard. In *2017 IEEE Virtual Reality (VR)*, pages 349–350. IEEE, 2017.

Jie Li, Zhiyuan Zheng, Britta Meixner, Thomas Röggla, Maxine Glancy, and Pablo Cesar. Designing an Object-Based Preproduction Tool for Multiscreen TV Viewing. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, 2018. Association for Computing

Machinery. ISBN 9781450356213. doi: 10.1145/3170427.3188658. URL `https://doi.org/10.1145/3170427.3188658`.

Yang Li, Dong-Xiao Li, Liang-Hao Wang, and Ming Zhang. Real time stereo rendering for augmented reality on 3DTV system. In *2012 International Conference on Systems and Informatics (ICSAI2012)*, pages 2125–2129. IEEE, 2012.

Ying Li. Augmented Reality for remote education. In *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, volume 3, pages V3–187. IEEE, 2010.

Geoffrey R Loftus and Michael EJ Masson. Using confidence intervals in within-subject designs. *Psychonomic bulletin & review*, 1(4):476–490, 1994.

Valentin Lohmüller and Christian Wolff. Towards a Comprehensive Definition of Second Screen. In *Proceedings of Mensch Und Computer 2019*, MuC'19, page 167–177, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450371988. doi: 10.1145/3340764.3340781. URL `https://doi.org/10.1145/3340764.3340781`.

Víctor Manuel López Jaquero, Radu-Daniel Vatavu, Jose Ignacio Panach, Oscar Pastor, and Jean Vanderdonckt. A Newcomer's Guide to EICS, the Engineering Interactive Computing Systems Community. *Proc. ACM Hum.-Comput. Interact.*, 3(EICS), June 2019. doi: 10.1145/3300960. URL `https://doi.org/10.1145/3300960`.

Artur Lugmayr, Bjoern Stockleben, Thomas Risse, Juha Kaario, and Bogdan Pogorelc. New Business, Design and Models to Create Semantic Ambient Media Experiences. *Multimedia Tools Appl.*, 66(1):1–5, September 2013. ISSN 1380-7501. doi: 10.1007/s11042-012-1239-1. URL `https://doi.org/10.1007/s11042-012-1239-1`.

Lotfi Maalej, Yosr Mlouhi, Imed Jabri, and Tahar Battikh. Artificial vision and augmented reality applied to the analysis of sports broadcast videos. In *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, pages 1–9. IEEE, 2015.

Arien Mack, Irvin Rock, et al. *Inattentional blindness*, volume 33. MIT press Cambridge, MA, 1998.

Steve Mann. Mediated reality. *Linux J.*, 1999(59es):5–es, mar 1999. ISSN 1075-3583.

Steve Mann, Tom Furness, Yu Yuan, Jay Iorio, and Zixin Wang. All Reality: Virtual, Augmented, Mixed (X), Mediated (X,Y), and Multimediated Reality, 2018. URL `https://arxiv.org/abs/1804.08386`.

Bernard Marr. What Is Extended Reality Technology? A Simple Explanation For Anyone, August 2019. URL `https://www.forbes.com/sites/bernardmarr/2019/08/12/` `what-is-extended-reality-technology-a-simple-explanation-for-anyone/` `?sh=783bf0c67249`.

Florian Mathis, Xuesong Zhang, Mark McGill, Adalberto L. Simeone, and Mohamed Khamis. Assessing Social Text Placement in Mixed Reality TV. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 205–211, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3399402. URL `https://doi.org/10.1145/3391614.3399402`.

Philipp Mayring. Qualitative content analysis: theoretical foundation, basic procedures and software solution. 2014.

Mark McGill, Florian Mathis, Mohamed Khamis, and Julie Williamson. Augmenting TV Viewing Using Acoustically Transparent Auditory Headsets. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 34–44, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393650. URL `https://doi.org/10.1145/3391614.3393650`.

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

Jennifer McNally and Beth Harrington. How Millennials and Teens Consume Mobile Video. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '17, page 31–39, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450345293. doi: 10.1145/3077548.3077555. URL `https://doi.org/10.1145/3077548.3077555`.

L. M. J. Meesters, W. A. IJsselsteijn, and P. J. H. Seuntiens. A survey of perceptual evaluations and requirements of three-dimensional TV. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(3):381–391, March 2004. ISSN 1558-2205. doi: 10.1109/TCSVT.2004.823398.

Klaus Merkel. Hybrid broadcast broadband TV, the new way to a comprehensive TV experience. In *Electronic Media Technology (CEMT), 2011 14th ITG Conference on*, pages 1–4. IEEE, 2011.

Paul Milgram and Herman Colquhoun Jr. *A Taxonomy of Real and Virtual World Display Integration*. Springer-Verlag, Berlin, Heidelberg, 1999.

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

Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 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*, volume 2351, 1995. doi: 10.1117/12.197321. URL `https://doi.org/10.1117/12.197321`.

Paul Milgram, Herman Colquhoun, et al. A taxonomy of real and virtual world display integration. *Mixed reality: Merging real and virtual worlds*, 1(1999):1–26, 1999.

David Moher, Alessandro Liberati, Jennifer Tetzlaff, and Douglas G Altman. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4):264–269, 2009.

Mukesh Nathan, Chris Harrison, Svetlana Yarosh, Loren Terveen, Larry Stead, and Brian Amento. CollaboraTV: Making Television Viewing Social Again. In *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*, UXTV '08, page 85–94, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605581002. doi: 10.1145/1453805.1453824. URL `https://doi.org/10.1145/1453805.1453824`.

Netflix. Black Mirror: Bandersnatch, December 2018. URL `https://www.netflix.com/gb/title/80988062`.

Nielsen. Streamers Show Strong Ties with Traditional TV, 2018. URL `https://www.nielsen.com/us/en/insights/news/2018/streamers-showstrong-ties-with-traditional-tv.html`.

Lyndon Nixon, Lotte Belice Baltussen, and Johan Oomen. 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, page 37–40, Aachen, DEU, 2015. CEUR-WS.org.

Heather L. O'Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112:28–39, 2018. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2018.01.004. URL `https://www.sciencedirect.com/science/article/pii/S1071581918300041`.

Virpi Oksman, Elina Noppari, Antti Tammela, Maarit Mäkinen, and Ville Ollikainen. Mobile TV in Everyday Life Contexts: Individual Entertainment or Shared Experiences? In *Proceedings of the 5th European Conference on Interactive TV and Video*, EuroITV '07, page 215–225, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 9783540725589. URL `https://doi.org/10.1007/978-3-540-72559-6_23`.

David Oyarzun, Andoni Mujika, Aitor Álvarez, Aritz Legarretaetxeberria, Aitor Arrieta, and María del Puy Carretero. High-realistic and flexible virtual presenters. In *International Conference on Articulated Motion and Deformable Objects*, pages 108–117. Springer, 2010.

H. M. Ozaktas and L. Onural. *Three-Dimensional Television: Capture, Transmission, Display*. Springer, Heidelberg, Germany, 2010. URL `http://dx.doi.org/10.1007/978-3-540-72532-9`.

J. A. Paradiso and J. A. Landay. Guest Editors' Introduction: Cross-Reality Environments. *IEEE Pervasive Computing*, 8(3):14–15, July 2009. ISSN 1558-2590. doi: 10.1109/MPRV.2009.47. URL `https://doi.org/10.1109/MPRV.2009.47`.

Hyerim Park and Woontack Woo. Metadata design for location-based film experience in augmented places. In *2015 IEEE International Symposium on Mixed and Augmented Reality-Media, Art, Social Science, Humanities and Design*, pages 40–45. IEEE, 2015.

Hyerim Park and Woontack Woo. Metadata design for AR spacetelling experience using movie clips. In *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pages 388–391. IEEE, 2017.

Bogdan Pogorelc, Radu-Daniel Vatavu, Artur Lugmayr, Björn Stockleben, Thomas
    Risse, Juha Kaario, Estefania Constanza Lomonaco, and Matjaž Gams. Semantic
    Ambient Media: From Ambient Advertising to Ambient-Assisted Living. *Multi-
    media Tools Appl.*, 58(2):399–425, May 2012. ISSN 1380-7501. doi: 10.1007/
    s11042-011-0917-8. URL `https://doi.org/10.1007/s11042-011-0917-8`.

Bogdan Pogorelc, Artur Lugmayr, Björn Stockleben, Radu-Daniel Vatavu, Nina Tah-
    masebi, Estefanía Serral, Emilija Stojmenova, Bojan Imperl, Thomas Risse, Gideon
    Zenz, and et al. Ambient Bloom: New Business, Content, Design and Models to
    Increase the Semantic Ambient Media Experience. *Multimedia Tools Appl.*, 66(1):
    7–32, September 2013. ISSN 1380-7501. doi: 10.1007/s11042-012-1228-4. URL
    `https://doi.org/10.1007/s11042-012-1228-4`.

Irina Popovici and Radu-Daniel Vatavu. Understanding Users' Preferences for Aug-
    mented Reality Television. In *Proceedings of the 2019 International Symposium
    on Mixed and Augmented Reality*, ISMAR '19, pages 269–278, 2019a. URL
    `https://doi.org/10.1109/ISMAR.2019.00024`.

Irina Popovici and Radu-Daniel Vatavu. Consolidating the Research Agenda of
    Augmented Reality Television with Insights from Potential End-Users. In *Pro-
    ceedings of 2019 IEEE International Symposium on Mixed and Augmented Re-
    ality Adjunct (ISMAR-Adjunct)*, ISMAR '19, pages 73–74, 2019b. URL `https:
    //doi.org/10.1109/ISMAR-Adjunct.2019.00033`.

Irina Popovici and Radu-Daniel Vatavu. 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, pages 199–204, New York, NY, USA, 2019c. ACM. ISBN 978-1-4503-6017-3.
    doi: 10.1145/3317697.3325121. URL `http://doi.acm.org/10.1145/3317697.
    3325121`.

Irina Popovici, Radu-Daniel Vatavu, Pu Feng, and Wenjun Wu. AR-TV and AR-
    DiàNshì: Cultural Differences in Users' Preferences for Augmented Reality Tele-
    vision. In *ACM International Conference on Interactive Media Experiences*,
    IMX '21, page 50–60, New York, NY, USA, 2021. Association for Comput-
    ing Machinery. ISBN 9781450383899. doi: 10.1145/3452918.3458801. URL
    `https://doi.org/10.1145/3452918.3458801`.

Michael I Posner. Orienting of attention. *Quarterly journal of experimental psychology*, 32(1):3–25, 1980.

Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001*, pages 89–102. Springer, 2001.

Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. Electric Agents: Combining Collaborative Mobile Augmented Reality and Web-Based Video to Reinvent Interactive Television. *Comput. Entertain.*, 12(3):1:1–1:21, February 2015. ISSN 1544-3574. doi: 10.1145/2702109.2633413. URL `http://doi.acm.org/10.1145/2702109.2633413`.

Jacob M. Rigby, Duncan P Brumby, Sandy J. J. Gould, and Anna L Cox. 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, page 35–46, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450360173. doi: 10.1145/3317697.3323361. URL `https://doi.org/10.1145/3317697.3323361`.

Eduardo Rodrigues, Lucas Silva Figueiredo, Lucas Maggi, Edvar Neto, Layon Tavares Bezerra, João Marcelo Teixeira, and Veronica Teichrieb. Mixed Reality TVs: Applying Motion Parallax for Enhanced Viewing and Control Experiences on Consumer TVs. In *2017 19th Symposium on Virtual and Augmented Reality (SVR)*, pages 319–330. IEEE, 2017.

Asreen Rostami, Chiara Rossitto, and Annika Waern. Frictional Realities: Enabling Immersion in Mixed-Reality Performances. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '18, page 15–27, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450351157. doi: 10.1145/3210825.3210827. URL `https://doi.org/10.1145/3210825.3210827`.

Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '19, pages 255–261, New

York, NY, USA, 2019a. ACM. ISBN 978-1-4503-6017-3. doi: 10.1145/3317697. 3325129. URL http://doi.acm.org/10.1145/3317697.3325129.

Pejman Saeghe, Sarah Clinch, Bruce Weir, Steve Pettifer, and Robert Stevens. 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*. ACM, 2019b.

Pejman Saeghe, Gavin Abercrombie, Bruce Weir, Sarah Clinch, Stephen Pettifer, and Robert Stevens. Augmented Reality and Television: Dimensions and Themes. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 13–23, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393649. URL https://doi.org/10. 1145/3391614.3393649.

Pejman Saeghe, Mark McGill, Bruce Weir, Sarah Clinch, and Robert Stevens. Evaluating and updating a design space for augmented reality television. In *ACM International Conference on Interactive Media Experiences*, IMX '22, page 79–94, New York, NY, USA, 2022a. Association for Computing Machinery. ISBN 9781450392129. doi: 10.1145/3505284.3529965. URL https://doi.org/10. 1145/3505284.3529965.

Pejman Saeghe, Bruce Weir, Mark McGill, Sarah Clinch, and Robert Stevens. Augmenting a nature documentary with a lifelike hologram in virtual reality. In *ACM International Conference on Interactive Media Experiences*, IMX '22, page 275–280, New York, NY, USA, 2022b. Association for Computing Machinery. ISBN 9781450392129. doi: 10.1145/3505284.3532974. URL https://doi.org/10. 1145/3505284.3532974.

Hideo Saito, Toshihiro Honda, Yusuke Nakayama, and Francois De Sorbier. Camera pose estimation for mixed and diminished reality in FTV. In *2014 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4. IEEE, 2014.

Andrea Sanna, Fabrizio Lamberti, Francesco De Pace, Roberto Iacoviello, and Paola Sunna. ARSSET: Augmented Reality Support on SET. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 356–376. Springer, 2017.

Ingo Schiller, Bogumil Bartczak, Falko Kellner, and Reinhard Koch. Increasing Realism and Supporting Content Planning for Dynamic Scenes in a Mixed Reality System incorporating a Time-of-Flight Camera. *JVRB-Journal of Virtual Reality and Broadcasting*, 7(4), 2010.

Albrecht Schmidt, Hans-Werner Gellersen, and Michael Beigl. Matching Information and Ambient Media. In *Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, CoBuild '99, page 140–149, Berlin, Heidelberg, 1999. Springer-Verlag. ISBN 354066596X. URL https://doi.org/10.1007/10705432_13.

O Schreer, W Waizenegger, W Fernando, H Kodikara Arachchi, A Oehme, A Smolic, B Yargicoglu, A Akman, and U Curjel. Mixed reality technologies for immersive interactive broadcast. 2016.

Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001.

Jenny Scribani. What is Extended Reality (XR)?, January 2019. URL https://www.visualcapitalist.com/extended-reality-xr/.

Jonathan Simsch and Jens Herder. SpiderFeedback: Visual Feedback for Orientation in Virtual TV Studios. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, ACE '14, pages 12:1–12:8, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2945-3. doi: 10.1145/2663806.2663830. URL http://doi.acm.org/10.1145/2663806.2663830.

Robert Sitnik, Slawomir Pasko, Maciej Karaszewski, and Marcin Witkowski. Internet virtual studio: low-cost augmented reality system for WebTV. In *The Engineering Reality of Virtual Reality 2008*, volume 6804, page 68040E. International Society for Optics and Photonics, 2008.

José Somolinos. Augmented Reality disrupts Television, Feburary 2019. URL https://medium.com/@josesomolinos/augmented-reality-disrupts-television-9d79ed01f27.

Rafael Sotelo, Jose Joskowicz, and Nicolás Rondan. Experiences on hybrid television and augmented reality on ISDB-T. In *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5. IEEE, 2017.

Maximilian Speicher, Brian D. Hall, and Michael Nebeling. What is Mixed Reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 537:1–537:15, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300767. URL http://doi.acm.org/10.1145/3290605.3300767.

J. Stauder and P. Robert. Let the sunshine on your screen: introducing augmented reality into interactive television. In *Proceedings. IEEE International Conference on Multimedia and Expo*, volume 1, pages 837–840 vol.1, Aug 2002. doi: 10.1109/ICME.2002.1035912.

Ivan E. Sutherland. A Head-Mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), page 757–764, New York, NY, USA, 1968. Association for Computing Machinery. ISBN 9781450378994. doi: 10.1145/1476589.1476686. URL https://doi.org/10.1145/1476589.1476686.

Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 73–80. ACM, 2003.

TIME. Facebook Reveals Augmented Reality, Virtual Reality and More at Annual F8 Conference, April 2017. URL https://www.youtube.com/watch?v=YtUye84PuFY.

TWiT Tech Podcast Network. Augmented Reality on Your Television, July 2018. URL https://www.youtube.com/watch?v=lpj90PkfPJ8.

Radu-Daniel Vatavu. Point & Click Mediated Interactions for Large Home Entertainment Displays. *Multimedia Tools and Applications*, 59(1):113–128, 2012. doi: 10.1007/s11042-010-0698-5. URL http://dx.doi.org/10.1007/s11042-010-0698-5.

Radu-Daniel Vatavu. 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, pages 143–152, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1951-5. doi: 10.1145/2465958.2465972. URL http://doi.acm.org/10.1145/2465958.2465972.

Radu-Daniel Vatavu. 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, pages 13–22, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3526-3. doi: 10.1145/2745197.2745207. URL `http://doi.acm.org/10.1145/2745197.2745207`.

Radu-Daniel Vatavu, Pejman Saeghe, Teresa Chambel, Vinoba Vinayagamoorthy, and Marian F Ursu. Conceptualizing Augmented Reality Television for the Living Room. In *ACM International Conference on Interactive Media Experiences*, IMX '20, page 1–12, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379762. doi: 10.1145/3391614.3393660. URL `https://doi.org/10.1145/3391614.3393660`.

Vinoba Vinayagamoorthy, Penelope Allen, Matt Hammond, and Michael Evans. Researching the User Experience for Connected Tv: A Case Study. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, page 589–604, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450310161. doi: 10.1145/2212776.2212832. URL `https://doi.org/10.1145/2212776.2212832`.

Vinoba Vinayagamoorthy, Rajiv Ramdhany, and Matt Hammond. Enabling Frame-Accurate Synchronised Companion Screen Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '16, page 83–92, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450340670. doi: 10.1145/2932206.2932214. URL `https://doi.org/10.1145/2932206.2932214`.

Vinoba Vinayagamoorthy, Maxine Glancy, Paul Debenham, Alastair Bruce, Christoph Ziegler, and Richard Schäffer. 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, pages 179–184, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5115-7. doi: 10.1145/3210825.3213562. URL `http://doi.acm.org/10.1145/3210825.3213562`.

Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 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, CHI '19, pages 532:1–532:12, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300762. URL `http://doi.acm.org/10.1145/3290605.3300762`.

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

John Waterworth and Kei Hoshi. *Human-Experiential Design of Presence in Everyday Blended Reality*. Springer, Switzerland, 2016. URL `https://www.springer.com/gp/book/9783319303321`.

Brian Westover. How to connect Alexa to your Samsung Smart TV, December 2019. URL `https://www.tomsguide.com/uk/us/samsung-smart-tv-alexa-how-to,news-29557.html`.

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

Jacob O. Wobbrock and Julie A. Kientz. Research Contributions in Human-Computer Interaction. *Interactions*, 23(3):38–44, April 2016. ISSN 1072-5520. doi: 10.1145/2907069. URL `https://doi.org/10.1145/2907069`.

Claes Wohlin. Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, EASE '14, pages 38:1–38:10, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2476-2. doi: 10.1145/2601248.2601268. URL `http://doi.acm.org/10.1145/2601248.2601268`.

Yongxin Yan and Xiaolei Zhang. Research and analysis of the Virtual Reality with FLARToolKit. In *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, pages 1614–1617. IEEE, 2011.

Zenith Media. Media Consumption Forecasts 2018. URL `https://www.zenithmedia.com/product/media-consumption-forecasts2018`.

Yanxiang Zhang and Ziqiang Zhu. Live Mixed Reality Video Production for Educational Presentation. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 363–372. Springer, 2018.

Christian Zimmer, Nanette Ratz, Michael Bertram, and Christian Geiger. War children: Using ar in a documentary context. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 390–394, 2018. doi: 10.1109/ISMAR-Adjunct.2018.00112.

Appendix A

Supplementary Material for Chapter 4

A.1   Video scripts, handouts, and cheat sheet

A.1.1   Task 1

First task was identical between study and control groups. After sending us a signed consent form, participants received:

1) a link to a YouTube video describing the first task,

2) a PDF file containing the script of the above mentioned YouTube video, and

3) a handout containing basic definitions, additional notes, and instructions regarding time, word-length, and format constraints of the ARTV scenario.

ARTV Design Space Evaluation

Introduction 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.

ARTV Design Space Evaluation

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)

A.1.2   Task 2 - control group

For their second task, participants in the control group received the same handout as before (see Appendix A.1.1).

A.1.3   Task 2 - study group

For their second task, participants in the study group received *the ARTV design space cheat sheet*.

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

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 vs 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:

A) Within content:

(Between AR content and TV content)

1. 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?
2. Time: what is the relationship between AR and TV contents regarding their timelines?
   a. Are they synchronized?
   b. If they are synchronized, are both AR and TV elements presented simultaneously, or is one element paused while the other element is being presented?
3. Display: how is content presented?
   a. In a single or multiple frames
   b. An extended frame, for instance a TV screen visually extended with content on the surrounding walls
   c. Unframed
   d. Mixture of framed and unframed

B) Between ARTV content and non-content:

Between content and space:

1. Does content augment the space in a meaningful way?
2. Is content being modified or generated based on the features of space?

Between content and people:

Can viewers interact with content? What is the effect of their interaction?
1. Selection and manipulation of content
2. Change of story path
3. Accomplishment of tasks in a game-like fashion

Between content and objects:

1. Does content augment objects?
2. Is content modified or generated based on features of objects in the environment?
3. Can objects in the environment be integrated into the ARTV experience?

C) Within non-content:

Between people and people:

If there are multiple viewers:
1. Are the viewers co-located?

2. Are they viewing ARTV concurrently?
3. Do the viewers know each other prior to the viewing experience?
4. Are they meant to talk or interact with each other during the viewing?

Between people and space:

1. How familiar are viewers with the space?
2. Can they manipulate space?
3. Can they move around in the space?

Between people and objects:

1. How familiar are the viewers with the objects in the environment?
2. Can they influence or manipulate these objects?

Final 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.

A.1.4   ARTV design space cheat sheet - update

ARTV design space cheat sheet updated in light of our findings during the analysis of scenarios and interviews.

ARTV Design Space

(Updated)

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 images (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 vs 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 primary 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?

Appendix B

Supplementary Material for Chapter 7

B.1   Questionnaire response statistics and plots

This section presents the statistics regarding individual questionnaire responses and line plots with 95% within-subject confidence intervals.

Table B.1: Min, max, median, mean, standard deviation, maximum possible value, and *score* calculated for focused attention per six experimental conditions. *Score* is calculated by dividing the mean by the maximum possible value for each column, then converting the result into a percentage. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
Min	2.61	2.2	2.61	3.33	1.9	2.43
Max	4.9	4.13	4.2	3.91	4.03	4.31
Median	3.31	3.61	3.44	3.68	3.61	3.75
Mean	3.43	3.56	3.5	3.65	3.4	3.73
Std. Dev.	0.59	0.56	0.45	0.2	0.65	0.53
Max Possible	5	5	5	5	5	5
Score	68.6%	71.2%	70%	73%	68%	74.6%

Figure B.1: Focused attention score across the six experimental conditions. Error bars show the 95% within participant confidence intervals [Loftus and Masson, 1994].
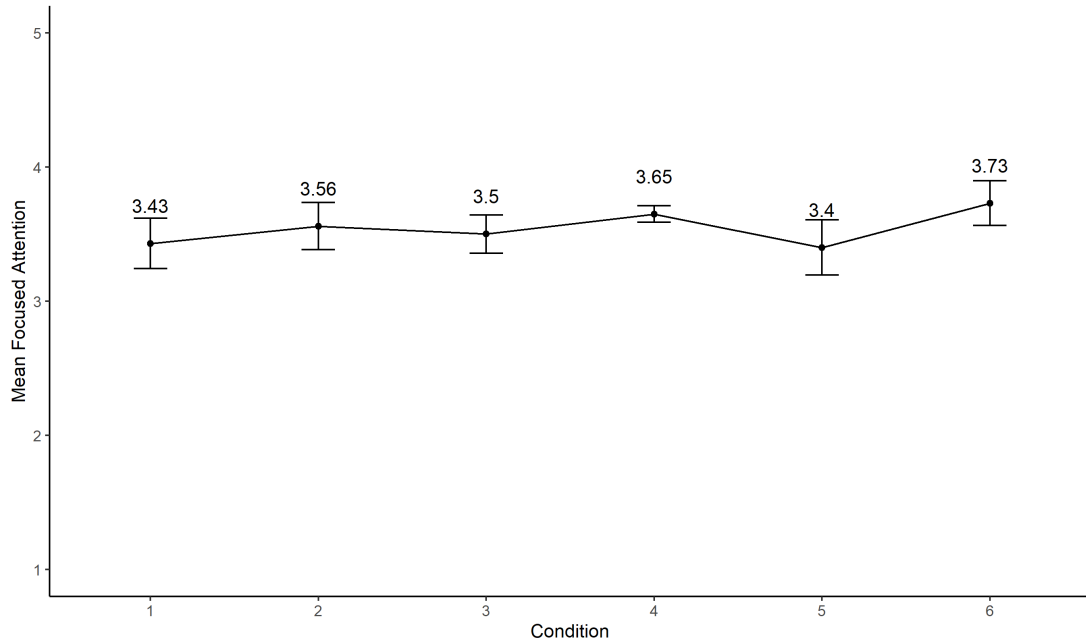

Table B.2: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Film IEQ per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
Min	104.33	83.5	105.83	111.17	95.5	105.17
Max	138.5	130.17	122.5	125.5	124.5	147.5
Median	114.33	120.42	114.92	118.08	119.33	118.08
Mean	116	116.7	115.6	117.2	114.6	119.9
Std. Dev.	11.13	13.74	5.25	4.39	10.32	11.81
Max Possible	168	168	168	168	168	168
Score	69.05%	69.46%	68.81%	69.76%	68.21%	71.37%

Table B.3: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Confusion per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
Min	0.85	1.02	1.35	1.35	1.35	0.85
Max	5.68	5.35	4.85	5.35	6.35	4.35
Median	2.77	2.68	2.43	2.93	3.43	3.02
Mean	3.1	2.5	2.8	3.2	3.6	2.9
Std. Dev.	1.62	1.25	1.4	1.51	1.59	1.1
Max Possible	7	7	7	7	7	7
Mean Ratio	44.29%	35.71%	40%	45.71%	51.43%	41.43%

Figure B.2: Film IEQ score across the six experimental conditions. Error bars show the 95% within participant confidence intervals [Loftus and Masson, 1994].
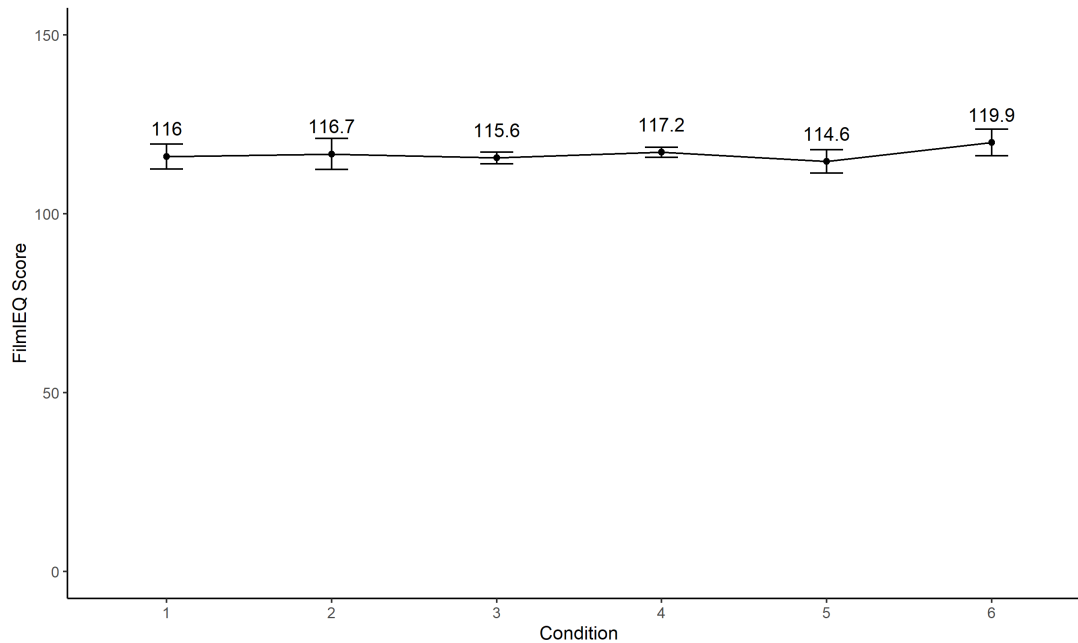

Table B.4: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Distraction per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
Min	2.45	2.45	1.62	1.95	2.45	3.45
Max	6.45	7.78	6.95	5.62	6.62	6.62
Median	3.7	4.62	4.03	4.7	4.78	4.78
Mean	3.9	4.8	4.1	4.6	4.5	4.8
Std. Dev.	1.34	1.51	1.57	1.08	1.44	1
Max Possible	7	7	7	7	7	7
Mean Ratio	55.71%	68.57%	58.57%	65.71%	64.29%	68.57%

Table B.5: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Enjoyment per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
Min	2.83	3	3.83	1.5	2	2.67
Max	7.5	7.67	6.67	6	6	7
Median	4.33	5.17	5	4.58	3.58	4.75
Mean	4.5	5.4	5.2	4.4	3.9	4.6
Std. Dev.	1.51	1.81	0.84	1.27	1.37	1.31
Max Possible	7	7	7	7	7	7
Mean Ratio	64.29%	77.14%	74.29%	62.86%	55.71%	65.71%

Figure B.3: ARTV Questionnaire scores related to the TV+hologram: Confusion, Distraction, and Enjoyment across the six experimental conditions. Error bars show the 95% within participant confidence intervals [Loftus and Masson, 1994].
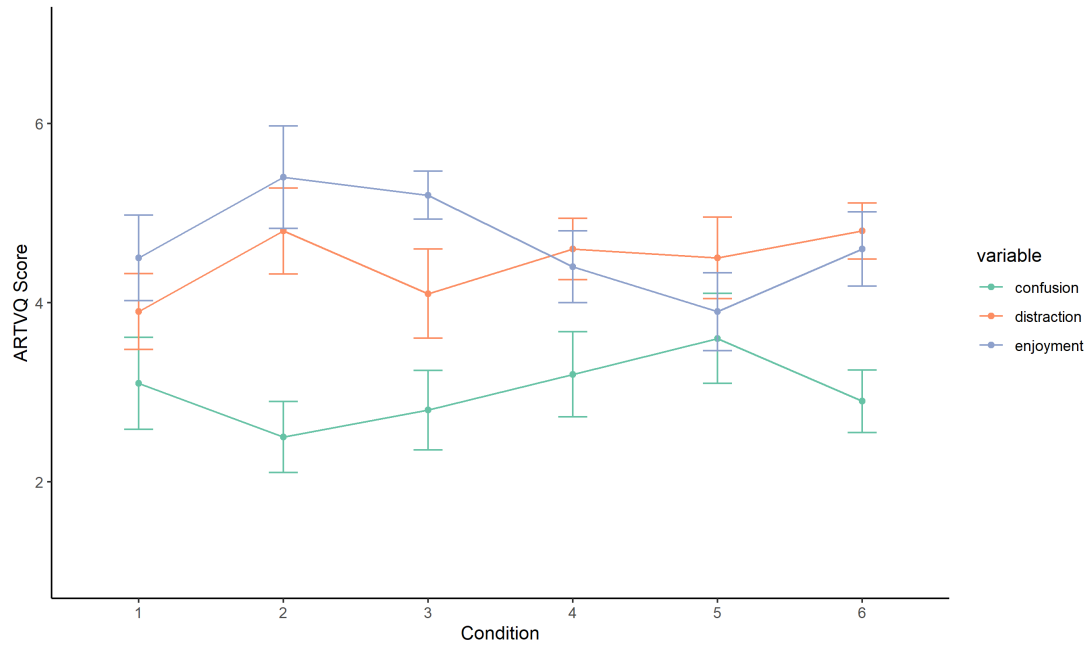

Table B.6: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Interest per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	**Condition 2**	**Condition 3**	**Condition 4**	**Condition 5**	**Condition 6**
Min	2.73	2.9	1.73	2.57	2.4	2.57
Max	8.57	7.57	6.57	5.73	6.9	6.9
Median	5.07	5.73	4.48	4.57	3.98	4.65
Mean	5	5.8	4.3	4.4	4.2	4.7
Std. Dev.	1.77	1.52	1.52	1.05	1.63	1.24
Max Possible	7	7	7	7	7	7
Mean Ratio	71.43%	82.86%	61.43%	62.86%	60%	67.14%

Table B.7: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Realism per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	**Condition 2**	**Condition 3**	**Condition 4**	**Condition 5**	**Condition 6**
Min	3.7	3.03	3.03	3.7	2.37	3.7
Max	7.7	6.03	5.87	5.37	7.03	6.03
Median	4.78	4.78	4.37	4.7	4.7	4.7
Mean	4.9	4.9	4.4	4.6	4.6	4.8
Std. Dev.	1.15	0.87	0.91	0.49	1.39	0.71
Max Possible	7	7	7	7	7	7
Mean Ratio	70%	70%	62.86%	65.71%	65.71%	68.57%

Table B.8: Min, max, median, mean, standard deviation, and *mean ratio* of the responses for Presence per six experimental conditions). *Mean ratio* is calculated by dividing the mean score for each column by the maximum possible value for that column, then multiplying the result by 100. All values have been calculated after adjusting for the within subject variation according to Loftus and Masson.

	Condition 1	**Condition 2**	**Condition 3**	**Condition 4**	**Condition 5**	**Condition 6**
Min	3.02	3.18	3.85	3.18	3.68	3.85
Max	6.18	7.02	7.18	6.02	7.85	7.18
Median	5.68	5.68	5.27	5.43	5.85	5.85
Mean	5.4	5.5	5.3	5.2	5.8	5.9
Std. Dev.	0.92	1.21	0.86	0.83	1.17	0.96
Max Possible	7	7	7	7	7	7
Mean Ratio	77.14%	78.57%	75.71%	74.29%	82.86%	84.29%

Figure B.4: ARTV Questionnaire scores related to the hologram: Interest, Realism, and Presence across the six experimental conditions. Error bars show the 95% within participant confidence intervals [Loftus and Masson, 1994].