Video Media in the Hyperconnected Age: Investigating Emergent Viewing Practices

Jacob Mark Rigby

Supervisors:
Dr. Duncan Brumby - Prof. Anna Cox - Dr. Sandy Gould

Examiners:
Dr. Enrico Costanza (internal) - Prof. Marianna Obrist (external)

A dissertation submitted in partial fulfilment
of the requirements for the degree of:

Doctor of Engineering
of
University College London

UCL Interaction Centre, Department of Computer Science, University College London

2018

DECLARATION

I, Jacob Mark Rigby, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Recent technological developments have changed the way video is consumed. The uptake of fast internet connections and ubiquitous mobile devices mean that people can watch via on demand services, and that viewers often media multitask with phones and tablets during viewing.

This thesis examines on-demand viewing and media multitasking with mobile devices in detail. Two situated studies extend our understanding of these behaviours through video observation and diary studies. It was found that using mobile devices while viewing was common, though subject to different usage patterns and individual differences. Self-reported media multitasking propensity correlated with observational data, suggesting that some people consistently media multitask more than others. People valued the freedom and choice provided by on-demand services, which drove their popularity. Viewing occurred in a range of contexts and on a variety of devices. However, some were concerned that it was difficult to limit their viewing.

In order to quantify viewer experience, a questionnaire was developed to measure immersion. This was used in two lab experiments investigating specific behaviours that were previously observed: watching on screens of different sizes; and being interrupted by notifications while watching. It was found that both watching on small screens and interruptions from notifications negatively affected immersion.

The findings of this research affect viewers, content producers, and TV networks. To preserve and improve viewing experiences, stakeholders should be mindful of both positive and negative effects when considering personal usage and the development of new viewing technologies.

IMPACT STATEMENT

This thesis investigates new ways in which people are watching video media, specifically looking at how technology is affecting viewers' behaviour and experiences. Two specific viewing practices were investigated in detail: using mobile devices while viewing; and accessing media through on-demand services. The results showed that these practices can offer new and improved experiences for viewers, such as watching in different locations and on different devices. However, they could also lead to negative effects in some circumstances, such as distraction, and watching more content than originally intended.

The findings of this work have impacts both within and outside of academia. Within academia, this research has greatly expanded our current understanding of how people are augmenting their video viewing experiences with new technology. Through collection of situated data via observation and diary studies, a more fine-grained view of people's viewing habits has been developed than in previous research. This is useful in informing the design of future technologies related to video consumption, as well as better understanding existing ones. Furthermore, the questionnaire developed to measure viewer immersion (Film IEQ) provides a standardised way to understand how such technologies may affect viewing experiences, which has been demonstrated in two lab experiments in this thesis. In order to disseminate this knowledge to the wider scientific community, five academic papers have been published from this thesis at the time of submission (a full list of which appears later in this document).

Outside of academia, this work is directly relevant to the film and television industry. Content producers can take into account the actionable findings

of this work to make their content more suitable to modern viewing habits — e.g., a film production team could create two different cuts of a movie, one tailored for small screens (e.g. mobile phones) which accounts for the limitations surrounding viewing on small screens, and one for larger screens (e.g. TV) which could include details that could otherwise be missed on a small screen. Furthermore, this thesis presents findings that viewers themselves can take onboard, such as managing their mobile device notifications to preserve emergent viewing experiences, and limiting their amount of on-demand viewing to prevent any negative feelings that might be felt from binge watching.

ACKNOWLEDGEMENTS

First and foremost, my supervisors, Duncan, Anna, and Sandy. Their support and honest feedback throughout my doctorate has been instrumental in me getting this far. I was very lucky to have a supervisory team that complemented each other so well. Their help has allowed me to ensure my work is held to the high standards necessary for academic research. I would not have succeeded without it!

Special thanks to Paul Marshall, who took the time and effort to examine me for my first year viva, upgrade viva, and a surprise third viva. These experiences helped me to learn how to present and defend my work, as well ensuring it was moving in a sensible direction.

My examiners, Enrico Costanza and Marianna Obrist, for taking the time to read my thesis and examine me for my viva. What I thought would be a scary experience turned out to be quite pleasant.

My friends and colleagues at UCLIC, who have been a fantastic source of support and feedback during the last few years. Time is short, and I wish I'd gotten to know more of you better.

Finally, thanks to my parents for their unconditional support, and to Alice for putting up with me and my ways throughout this period of highs and lows.

This research was supported by EPSRC grant EP/G037159/1 through the UCL Doctoral Training Centre in Virtual Environments, Imaging and Visualisation (VEIV), and in part by BBC R&D.

PUBLICATIONS

The following publications are based on work presented in this thesis:

Jacob M. Rigby, Duncan P. Brumby, Sandy J.J. Gould, and Anna L. Cox. 2018. "I Can Watch What I Want": A Diary Study of On-Demand and Cross-Device Viewing. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video* (TVX '18). ACM, New York, NY, USA. DOI: http://doi.org/10.1145/3210825.3210832

Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J.J. Gould. 2018. Old Habits Die Hard: A Diary Study of On-Demand Video Viewing. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI EA '18). ACM, New York, NY, USA, Paper LBW016. DOI: http://doi.org/10.1145/3170427.3188665

Jacob M. Rigby, Duncan P. Brumby, Sandy J.J. Gould, and Anna L. Cox. 2017. Media Multitasking at Home: A Video Observation Study of Concurrent TV and Mobile Device Usage. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video* (TVX '17). ACM, New York, NY, USA, 3-10. DOI: http://doi.org/10.1145/3077548.3077560

Jacob M. Rigby, Duncan P. Brumby, Sandy J.J. Gould, and Anna L. Cox. 2017. Film, interrupted: investigating how mobile device notifications affect immersion during movies. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '17). ACM, New York, NY, USA, Article 93, 8 pages. DOI: http://doi.org/10.1145/3098279.3122136

Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J.J. Gould. 2016. Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (MobileHCI '16). ACM, New York, NY, USA, 714-721. DOI: `http://doi.org/10.1145/2957265.2961843`

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 PROBLEM STATEMENT

The way people consume television and film has been transformed in recent years. Due to various technological advances, it has changed from a passive, highly scheduled experience to one which puts the viewer in control. The internet has had a very large impact — the introduction of high-speed internet services to residential areas and the ever-increasing speeds of mobile internet have allowed for the streaming of video in real time, something which was not possible with previous low-bandwidth connections. This has led to the development of video-on-demand systems (e.g. YouTube, Netflix, and BBC iPlayer), where media can be consumed at the viewer's leisure instead of being at the mercy of the TV schedule.

Alongside the adoption of high-speed internet connections has been the rapid uptake of mobile devices, most noticeably mobile phones and tablets (Ofcom, 2015b). These devices give easy access to many forms of media, including film and television, but can also be a source of distraction by constantly demanding attention from the user through large numbers of notifications (Sahami Shirazi et al., 2014; Kushlev et al., 2016). The portability, power, and always-on internet connections of these devices has allowed computer systems to permeate all areas of our lives, becoming fully-fledged multimedia devices.

Since watching television became a common household activity in the mid-20th century, viewing has been a mostly passive experience. However, there have always been some low-tech interactive elements: families would play along with quiz shows, discuss the news, and speculate about what will happen in soap operas. In recent decades, viewers could interact with television shows using their household telephone, calling in to place a vote, talk to the presenters, and even play games. There were also some social elements, where families and friends might gather to watch shows together. With the spread of new broadcasting and network technologies such as satellite, cable, digital broadcast TV, and the internet, more direct implementations of interactive TV have become technologically feasible and more widespread as a result.

One instance of modern-day interactive TV is through "second screening", where an additional screen (typically that of a mobile device) is used while watching. Now that tablet and smart phone ownership is so common, this type of "media multitasking" (the simultaneous use of multiple media [Ophir et al. 2009]), has become especially popular, and affords opportunities to use these devices to support or extend viewing experiences. This can be either self-motivated (e.g. using Google to search for an actor, or tweeting one's own opinion about a sporting event) or through the use of specially designed companion content intended to be used while watching (Nandakumar and Murray, 2014). However, easy access to vast amounts of content, applications, and services via the internet can also lead to people using their devices for activities unrelated to the content they are watching (e.g., texting friends, using social media, or online shopping [Holz et al. 2015]). Furthermore, people can now use these devices to watch television and film content

thanks to increasingly popular on-demand and catchup services (e.g. Netflix[1], BBC iPlayer[2], and Amazon Video [3]), moving the televisual experience away from the traditional living room TV and into a multitude of new environments. This gives consumers a wide choice of devices and contexts for consuming video content, ranging from very small phones screens for watching on public transport, to large TV screens in their living rooms for watching when at home.

Media multitasking with mobile devices and using on-demand services are two practices have quickly become commonplace (Holz et al., 2015; The Nielsen Company, 2016b), but our current understanding of these behaviours could be developed further. Firstly, while we know people participate in these behaviours, the extent of this tends to be based on self-reports rather than on real-world data. Secondly, people's motivations for engaging in these behaviours are not well understood. For instance, on-demand video services allow for users to choose, what, when, where, and on which device they want to watch, but what motivates these decisions? Thirdly, the impacts of such viewing behaviours are not well understood. In terms of media multitasking with mobile devices, we know from other domains that multitasking generally comes at a cognitive cost, and that interruptions are usually detrimental to the task being performed. However, task performance and errors cannot easily be measured in a mostly passive scenario such as watching video, so how can we tell if multitasking and interruptions in this domain come with similar negative effects?

[1] `http://www.netflix.com/` [Accessed 30th October 2017]
[2] `http://www.bbc.co.uk/iplayer` [Accessed 30th October 2017]
[3] `http://www.amazon.co.uk/Amazon-Video/b?node=3010085031` [Accessed 17th July 2018]

This dissertation will examine the emergent viewing practices of media multitasking and on-demand viewing, specifically focusing on the UK. The research questions that will be addressed are:

RQ 1: How prevalent are technology-driven emergent viewing practices?

RQ 2: What are the motivations for participating in these viewing practices?

RQ 3: How do these viewing practices affect viewer experience?

Developing a more detailed understanding of the prevalence, motivations, and effects of emergent viewing behaviours is not only of interest to academics, but also to a number of other parties. Firstly, the viewers themselves — watching film and television is a popular leisure activity, fostering enjoyment, relaxation, learning, and even social interaction (Rubin, 1983), and so it is important to preserve these benefits. Secondly, content producers — the people who write, create and produce video media — are now having to compete with mobile devices for attention, but on the other hand they can also use new technologies as a vehicle to enhance and extend media experiences. Finally, TV networks and advertisers — traditional revenue models are being disrupted by the move away from broadcast TV, and mobile devices can be used as a method of advertising avoidance. However, these could also provide new opportunities for delivering adverts and new revenue streams.

1.2 THESIS STRUCTURE

The structure of this of thesis is as follows. Chapter 2 reviews relevant literature and provides motivation for the subsequent work. It details how technology usage has increased dramatically in all areas of life including television

watching, and how new behaviours have emerged around this. It also explores how these behaviours have been investigated and measured in prior studies to examine their impact, and the difficulties faced that are particular to the domain of film and television. A number of areas where further research would be beneficial were identified, specifically with regard to identifying these new viewing behaviours and measuring their impacts on viewers.

Chapter 3 details an in the wild study conducted to observe participants watching TV in their home. This study used video recording to give a detailed, real-world account of TV watching and mobile device usage. It was found that device usage while watching was common, accounting for up to 23% of participants' TV time. Furthermore, extensive use of on-demand services was also observed, accounting for 26% of viewing time across households and up to 68% of viewing for individual households. A number of related studies have used self-reporting methods, such as interviews, diary studies, and questionnaires to establish behaviour while watching TV (Foehr, 2006; Rideout et al., 2010; Vanattenhoven and Geerts, 2012; Voorveld and van der Goot, 2013). However, these methods can lack granularity, and participants can forget important events or simply not wish to reveal them, leading to results that may not accurately reflect real behaviour. A limited amount of observational studies have been conducted in this domain, taking different approaches (Voorveld and Viswanathan, 2014; Rooksby et al., 2014; Holz et al., 2015). However, this work has sometimes contradicted the findings of prior self reported studies, and each other. The contribution of the study presented in Chapter 3 is the use of real-world video data over a long period of time, which was also compared with self-reported media multitasking propensity.

Chapter 4 details a two-week diary study, during which nine households

were asked to record all of their on-demand viewing. This included what they watched, where they watched, which device they used, and at which time they watched. This was motivated by the study in Chapter 3 finding that over a quarter of viewing time was via on-demand services, as well as various industry and media reports of increased use of on-demand video services (e.g. Ofcom 2017; BARB 2017). In order to extend the theme of collecting situated data, a diary study was considered the most practical way to collect information about on-demand viewing behaviour. This was because one of the key advantages of on-demand services is that they allow of viewers to watch in locations other than the living room, and on devices other than the television. This was not observable in the video observation study in Chapter 3 due to the focus on the living room environment, therefore discounting a number of diverse scenarios where viewing could occur. This study found that viewing was affected by a number of contextual factors which impacted the choices people made, such as which device to view on, which location, etc. While people exhibited many behaviours that were not generally possible with traditional broadcast TV, such as watching mobile devices, and in a variety of locations, a large portion of viewing still confirmed to "traditional" ideas of TV viewing — evening viewing in the living room on a large screen. The contribution of the study presented in Chapter 4 is to develop our understanding of exactly how people are interacting with on-demand services, to give a fine-grained view that compliments our current broad understanding.

Chapter 3 and Chapter 4 allowed for the development of a detailed understanding of both concurrent TV and device usage and on-demand service usage respectively. A number of common viewing phenomena were observed in these studies, though it was difficult to fully understand these and how

they effect people's viewing experiences in isolation. In order to remedy this, a means of exploring these further in a controlled environment was sought. Chapter 5 describes the development of standardised tool to measure the effect of technological interventions, resulting in the Immersive Experience Questionnaire for Film and Television (Film IEQ). Questionnaires were explored as a way of measuring the effects of interventions and phenomena on people's viewing experiences, as they offer a means of quantifying subjective experience in a standardised way. No pre-existing tool to measure a person's broad video viewing experience was found, and so related literature was examined. The field of computer games research provides the Immersive Experience Questionnaire (IEQ) (Jennett et al., 2008), a widely-used tool used to measure player experience in terms of the immersion they feel. The IEQ consists of a number of subscales measuring different concepts that make up immersion. While a number of these could be directly applicable to TV and film, some elements were not, and so the IEQ was modified to be better suited to this domain. As a result of this, the Film IEQ was developed as a way of assessing the impact of technological interventions on viewer experience, which allows us to develop a deeper understand of particular behaviours. This was then used in following lab studies to investigate behaviours in isolation.

The first study utilising the Film IEQ is reported in Chapter 6, which details a controlled lab experiment to investigate the effect of screen size on viewer immersion. The diary study in Chapter 4 found that participants watched on a variety of screens, and that 29% of recorded viewing sessions included viewing on a handheld mobile device. However, throughout the diary entries and interviews conducted, participants generally expressed a preference for larger screens where possible. In order to examine this in detail,

participants were exposed to content on three commonly used devices: a smart phone, laptop monitor, and large TV screen. Immersion levels were measured using the Film IEQ after each condition. The results showed that screen size had an effect on immersion, with very small screens leading to lower immersion. Therefore, larger screens are better for engendering a more immersive experience.

Chapter 7 explores the impact of device notifications on viewer immersion by means of a controlled lab experiment. In Chapter 3, frequent short mobile device interactions were observed, suggesting the kind of notification-driven checking behaviour present when participating in an instant messaging conversation. We know that people receive many notifications throughout the day (Sahami Shirazi et al., 2014; Pielot et al., 2014), and that interruptions are often detrimental to performance in other domains (Monk et al., 2002; González and Mark, 2004). This provided motivation for the study in Chapter 7, where interrupted and uninterrupted viewing were compared. Immersion was measured using the Film IEQ, and the results showed that interruptions had a negative effect on immersion.

Chapter 8 provides a general discussion of the work presented in this thesis. The contributions are outlined, as well as the implications for stakeholders in this area. Limitations of this work are also explored, as well as possible future directions for research.

CHAPTER 2

BACKGROUND AND MOTIVATION

2.1 TECHNOLOGY: CATALYST FOR EMERGENT VIEWING PRACTICES

Since the home computer began to gain traction in the 70s and 80s, the general public's appetite for personal computing devices has continued to grow. In 2002, more than half of all households in the UK owned a personal computer and 44% had access to the internet (Office for National Statistics, 2017), and in 2011 80% of households owned a computer and 77% had internet access. While internet Access has grown steadily to 90% of households in the present day, desktop and laptop computers have become less popular in favour of mobile devices.

To obtain a detailed overview of technology usage and telecommunications, one can look to industry bodies and market research agencies such as Ofcom[1], Nielsen[2], and the Broadcasters Audience Research Board (BARB)[3], who perform large-scale research outside the remit of typical academic studies in this domain. Ofcom's annual Communications Market Report utilises surveys from large representative samples of the UK public to report the state of the telecommunications industry. Due to their regularity and scope, these studies can provide an excellent overview of general trends. In 2017, Of-

[1] `http://www.ofcom.org.uk/` [Accessed 18th July 2018]
[2] `http://http://www.nielsen.com/` [Accessed 18th July 2018]
[3] `http://http://www.barb.co.uk/` [Accessed 18th July 2018]

com found that ownership of desktop computers fell to 29% in 2017 (Ofcom, 2017). Conversely, with the introduction of smartphones and tablets, mobile device ownership has grown rapidly — 76% of adults now own a smartphone, and 58% of homes now have a tablet computer. Furthermore, these mobile devices are now the most common way of Accessing the internet, and are used for over two hours a day on average, rising to over four hours among 18-24 year-olds (Ofcom, 2016a).

The increase in popularity of mobile devices has gone hand-in-hand with faster and faster internet connections, both domestic and mobile. As of 2017, the average UK internet download speed was 36 Mbps, more than double that of 2014 (Ofcom, 2017). This has allowed for media content to be delivered over the internet instead of traditional over-the-air or cable broadcasting routes, and out of this have emerged numerous Video On-Demand (VOD) services able to stream content directly to consumers, both free and subscription based.

Mobile devices have become ubiquitous in all areas of life. Prior research has shown that people's smartphones are usually nearby (Ichikawa et al., 2005; Dey et al., 2011; Wiese et al., 2013). Just like most other everyday activities, mobile devices are increasingly present while watching television and film. This can offer the possibility of adding value to the viewing experience, for example through specially designed companion apps intended to be used while watching a particular programme. Furthermore, viewers can use their device to look up information about what they are watching themselves, such as to research an actor or to refresh their memory about the plot. However, the presence of mobile devices also leads to the potential for distraction. Viewers are offered the possibility of performing tasks and activities com-

pletely unrelated to what they are watching on TV (Holz et al., 2015). This kind of multitasking behaviour can potentially lead to disengagement from the video content (Angell et al., 2016).

2.2 On-Demand Viewing

Starting in the middle part of the 20th century, TV ownership in the UK has steadily grown, and quickly replaced the radio as the centrepiece of the home. This trend continued for decades, with TV ownership growing year on year[4]. UK terrestrial television originally consisted on a single BBC channel in 1936, followed by a second BBC channel in 1964. Both were publicly funded. Competition in the form of commercial channels (funded by advertisements) started to be introduced in 1955 with the launch of ITV, and after a long period of little choice, Channel 4 followed in 1982 and Channel 5 in 1997. This coincided with the advent of cable and satellite TV services in the 80s and 90s, offering a hitherto unheard of amount of channels, but only to paying subscribers. This meant that the large amount of households who did not have a cable or satellite subscription had five channels to choose from. Since then, the slow introduction of Digital Video Broadcasting technology, as well as a forced analogue to digital switchover between 2007 and 2012, has brought a larger selection of free channels into people's homes without the need for a paid subscription.

Until fairly recently, TV viewing typically happened in a linear fashion - shows were scheduled to be broadcast at particular times, and viewers chose what they wanted to watch from the TV schedule and tuned in at those times,

[4] `http://www.barb.co.uk/resources/tv-ownership/` [Accessed 7th Nov 2017]

often once a week. This type of "appointment viewing" brought about a sense of occasion to watching our favourite programmes. Families and friends could meet up and watch together, and discussing shows with others was unlikely to result in "spoilers" (finding out about key plot details before watching it for oneself) because everyone was at the same place in the story. For both TV and film, the networks and production companies tightly controlled access to content. For TV, viewers were spoon-fed content dictated by the TV schedule. In terms of movies, the cinema has traditionally been the only (legal) way to see the latest releases, with home media releases on VHS or DVD following after a period of time.

This lack of choice and a strict adherence to a TV schedule has all changed with widespread use on-demand services, which has been a revolution in the media industry. Faster communications infrastructure has allowed for video content to be supplied directly to consumers via the internet, giving unprecedented choice and flexibility. Not only do they now have instant access to vast content libraries, but they can be consumed on a large range of devices whenever and wherever. In the UK, the BBC iPlayer catchup TV service is the most popular service (Ofcom, 2017), with ITV hub (another catchup service) and Youtube also being popular. Additionally, paid services like Netflix and Amazon are also widely used. It was also found that most users of these services also have a traditional pay TV subscription, and so are supplementing their pre-existing content with additional content from streaming services. This finding was also corroborated in a global survey study by The Nielsen Company (2016b). However, they also found that one third of the respondents were considering cancelling their traditional service in favour of online-only TV.

However, live broadcast TV still has a place. Viewers still value it for keeping up with the news and live events, as well as for social viewing with family and friends and to provide background noise when doing other tasks (Ofcom, 2017). There is also evidence to show that live broadcast are significant drivers of viewer engagement, with up to 68% of Twitter messages about programmes occurring during broadcast (Nielsen Holdings, 2014). This leads to a varied viewing landscape, with viewers mixing both traditional live broadcast TV and on-demand services.

Prior HCI research has examined on-demand viewing and related emergent viewing technologies. Barkhuus and Brown (2009) conducted in-depth interviews to understand how TV watching was changing as a result of new technologies. In particular, they focused on personal video recorders (PVR) and internet downloads, as this study was conducted in 2009, before internet on-demand video services were common. They found that most participants who used a PVR system had moved away from watching live TV almost entirely, preferring to queue up recordings from their downloaded library. This freedom from the TV schedule was particularly valued by those with non-standard work schedules.

Irani et al. (2010) conducted a diary study of people's viewing habits. This study examined the temporality of viewing in 14 households, which included the use of time shift and early on-demand services. They found that viewing was typically based around the rhythms of individuals' lives, households, and peers. The ability to choose when to watch could help align televisual schedules, allowing members of a household to watch together. There was also much discussion in households about what to watch and about the content of a show. Irani et al. also found that TV content was used as a background to

other tasks, and to fill gaps of unscheduled time.

A study by Vanattenhoven and Geerts (2015) also looked at how different ways of consuming media occurred around the house via qualitative interviews, including on-demand content. They noted that viewing depended on the context of other things happening in the household. They found that on-demand viewing typically involved "heavier" content requiring more focus (e.g., films and TV series), and took place in the evening. In contrast, broadcast TV typically involved "lighter" content (e.g., news), which was watched while doing other tasks.

Nogueira et al. (2017) analysed a large dataset from a Portuguese on-demand TV operator. While the insights from this work are largely concerned with the technicalities of delivering video to consumers, it does offer some high-level insights into viewer behaviour. Nogueira et al. found that users interacted with this service throughout the day, though evening viewing was most popular. They also found that users exhibited a large amount of "zapping" behaviour when selecting content, similar to "channel surfing", taking on average 2.5 minutes to settle on something. However, their data only offers insights into viewing on the TV, and only details a single video service.

2.2.1 MOBILE VIEWING

Not only has on-demand video allowed for people to watch whatever content whenever they like, but has also facilitated choice in terms of where they watch. Often, this necessitates the need to watch on a mobile device (e.g. when watching on public transport, or watching in a different room in the house). In a study by Ofcom (2015a), 22% of respondents were increasingly consuming video on screens other than the television. Motivations for viewing

on mobile devices appear to be varied (Ofcom, 2016b), including trying to fit around other household members' viewing schedules, ease of access to content and ease of navigation, and portability.

Though apparently popular, mobile viewing has not been widely studied in HCI literature. An early study by Knoche and McCarthy (2005) explored design requirements for mobile TV. Through a user study of a novel mobile TV experience, they found that their sample of four participants generally reacted favourably when using it on their daily commute. However, as this study was conducted in 2005, the authors also detailed a number of technical challenges in delivering content that have since been overcome. Furthermore, this work focused on emulating a traditional TV experience as much as possible, and did not envisage the availability and variety of on-demand services that we enjoy today.

O'Hara et al. (2007) explored motivations for mobile viewing, such as passing time and being able to be present with family members while watching something else. Again, this was an early study conducted in 2007 before mobile video was really mainstream, and before the powerful devices and fast connections that we have today.

A more recent study by McNally and Harrington (2017) examined how millennials consume mobile video, and found yet more motivations. These included procrastination, stimulation, keeping up-to-date with something, seeking information, responding to a notification, wanting background noise, being lonely, wanting to share in emotional experiences, and to revisit an experience or event. It was also found that media choice depended on a number of factors, such as the level of stimulation provided, engagement required, and video length.

Bury and Li (2015) conducted a survey study in 2013 into different ways of consuming TV. They found that mobile viewing was unpopular, with 70% of respondents never having used mobile devices for viewing. Those that participated in mobile viewing mainly did so when travelling and commuting. However, this seems to have changed in recent years, with mobile viewing growing in popularity (Ofcom, 2017). This study also clearly shows a general shift away from live TV viewing to online viewing.

2.2.2 BINGE WATCHING

The availability of large amounts of content means that it is now possible to access entire seasons of shows at any time. This is a large departure from traditional appointment viewing, which meant having to wait an entire week for the next episode of a show. This allows viewers to consume large amounts of content in a short space of time, or even a single sitting. This type of behaviour has always been present to a small degree with the availability of VHS and DVD box sets, and has been referred to as "marathon viewing", likening it to a feat of endurance similar to its distance running namesake.

More recently, the term "binge watching" has become the widely used term for excessive viewing behaviours, bringing with it comparisons to other binging behaviours such as binge drinking and binge eating. The term has now become inexorably linked with the on-demand streaming of TV shows, and even with particular streaming services such as Netflix. Binge watching behaviour is very popular, and though definitions differ, various reports have investigated its prevalence — a US survey by Deloitte, LLP (2017) found that 73% of respondents have binge watched at some point, with 29% doing so weekly. Similarly, in the UK Ofcom (2017) found that 79% of people they

asked have binge watched at some point, with 35% saying they do it at least weekly. They also found that this behaviour is more common in young people, with 62% of 16-24 year olds doing so at least weekly. The Nielsen Company (2016b) also found that being able to watch multiple episodes of content at a time was a significant driver of on-demand service usage, with 66% saying it is a motivating factor for them.

Even though the practice of binge watching is common, it is often cast in a negative light (Matrix, 2014). This is perhaps due to the negative association with the word "binge" (Jenner, 2015) and other binging behaviours, such as binge drinking or binge eating. Studies have found evidence of negative feelings and effects surrounding binge watching behaviour — de Feijter et al. (2016) found that binge watching often caused people to watch for longer than they intended, and that other tasks were often postponed. As a result, people wanted to watch less, but were often unable to control their viewing. Walton-Pattison et al. (2016) also found that binge watching often meant putting off other tasks, and that participants were aware that binge watching might cause feelings of regret. Ofcom (2017) found that a third of binge watchers missed out on sleep because of it, over a quarter neglected household chores, and over a fifth said they they felt guilty for not doing something else.

There is also disagreement as to the exact definition of binge watching. Some define it as watching two or more episodes (Pittman and Sheehan, 2015; Ofcom, 2017), some as three or more episodes (de Feijter et al., 2016; Walton-Pattison et al., 2016), and some leave it unclear (Matrix, 2014).

2.3 HUMAN MULTITASKING AND MEDIA MULTITASKING

Concurrent TV and mobile device usage is now commonplace in our homes. To understand this type of media multitasking, we can look to the large amount of work on human multitasking from psychology and HCI domains. Multitasking is the engagement in more than one task concurrently, and is ubiquitous in everyday life — consider the everyday occurrences of checking emails while writing a document, or talking on the phone while out shopping. Research on multitasking and interruptions has typically focused on the workplace (e.g. O'Conaill and Frohlich (1995); González and Mark (2004); Mark et al. (2005)) and safety critical environments (e.g. Latorella (1998); Brumby et al. (2009)), where errors and mistakes can be financially detrimental or even fatal. This has led to a large body of literature examining the nature and effects of multitasking and interruptions in these environments, as well as strategies for mitigating them.

2.3.1 EFFECTS OF MULTITASKING AND INTERRUPTIONS

There are a number of negative effects that have been associated with multitasking. Due to limitations of the human cognitive system (Meyer and Kieras, 1997) concurrent task execution requires interleaving (Burgess et al., 2000; Salvucci and Taatgen, 2008). The need for attention to be switched from one task to another results in a "switch cost" (Rogers and Monsell, 1995) where the task set must be reconfigured, incurring a response time overhead. This is described as "mental gear changing" by Monsell (2003), where attention must be shifted to different stimuli and conceptual criteria, new goal states and how

to reach them must be retrieved and committed to working memory, a different response set must be enabled and the criteria for them adjusted. Furthermore, resuming a task that was previously being engaged in after switching to another results in a “resumption lag", which is the time interval between the end of a secondary task and the resumption of the primary task (Altmann and Trafton, 2004).

Task switches are initiated by interruptions, which can be self-motivated (internal interruptions) or motivated by the environment (external interruptions) (González and Mark, 2004), and both types are equally prevalent. Interruptions have been shown to have a negative effect on both performance and error rate when performing tasks (Altmann et al., 2014). Previous work has investigated ways in which to minimise the cost of interruptions by exploiting the concept of *breakpoints*, which are the natural boundaries between two units of task execution (Newtson, 1973; Iqbal and Bailey, 2007). This allows for differentiation between tasks, and also for larger tasks to be subdivided into subtasks, and it has been shown that if interruptions are deferred to these times then cognitive cost can be reduced when recovering from an interruption (Adamczyk and Bailey, 2004; Iqbal and Bailey, 2005; Bailey and Konstan, 2006). Although breakpoints are subjective in nature, previous research shows that different observers of a task are able to identify many of the same breakpoints, showing that people have some kind of shared, innate system for subdividing tasks (Zacks et al., 2001).

The relevance of an interruption to the primary task being performed has also been shown to affect how people recover from interruptions. Interruptions containing information relevant to the task have been shown to be less disruptive because they allow the user to maintain the context of the primary

task (Czerwinski et al., 2000b; Iqbal and Bailey, 2008). This has led to the development of systems to classify interruptions based on their relevance, so that less relevant interruptions can be deferred to more suitable times (Iqbal and Bailey, 2008; Arroyo and Selker, 2011). However, exactly what makes an interruption relevant to the task at hand is not always clear, and it has been shown that seemingly relevant interruptions can be more detrimental to performance that seemingly irrelevant ones (Gould, 2014).

It has been shown that individuals multitask either because they either wish to (or have to) in order to increase efficiency (Burgess et al., 2000) — they have a number of tasks which must be completed in an allotted time, such as when involved in multiple projects at work (González and Mark, 2004), or looking after children while doing housework (Bittman and Wajcman, 2000) — or because they perceive some other utility in doing so, such as the alleviation of boredom (e.g. during travel time (Lyons and Urry, 2005)). Despite the potential for negative effects, sometimes multitasking can seem like a rational choice (Janssen et al., 2015), e.g. a doctor moving to a higher priority patient, a worker feeling that they are not progressing on their current task, or switching to obtain new, pertinent information.

2.3.2 Media Multitasking

Media multitasking is the concurrent consumption of multiple types of media (Ophir et al., 2009), and has become especially prevalent in the recent years due to the rapid uptake of mobile devices. Communications Market Reports by Ofcom have shown found that 53% of UK adults regularly media multitasked in 2013 (Ofcom, 2013), and a 2014 report showed that 99% of adults media multitask at some point during the week, for an average of 2 hours and

3 minutes every day (Ofcom, 2014), showing that this behaviour is growing very quickly. Media multitasking when watching television is particularly common (Foehr, 2006; Roberts and Foehr, 2008; Brasel and Gips, 2011).

2.3.2.1 Effects of Media Multitasking

Although there is some work which shows potential benefit in media multitasking specifically (Lui and Wong, 2012), there are a number of negative effects that have been associated with the phenomenon. When considering consuming TV with other media, the amount of time the viewer is spending away from the programme should be considered, as it could have a detrimental effect on how much the viewer is following the programme they are watching. A study into the distribution of visual attention of second screen usage by Holmes et al. (2012) found that around 30% of visual attention was given to the tablet used in the study, and the average gaze length given to the TV also decreased. Neate et al. (2015) found that the issue of distraction was a concern to viewers and that they liked to have some control over it, such as through the use of audio notification of new second screen content.

There is also evidence that people who media multitask a lot do not generally perform well at it. By conducting a series of experiments to investigate cognitive control when media multitasking, Ophir et al. (2009) found that heavy media multitaskers (HMMs) actually perform worse on tasks when media multasking than light media multitaskers (LMMs), even though they do it more often. Lottridge et al. (2015) examined this in a realistic work situation in the form of an essay writing exercise, and found that it depended on the type of distraction. The writing quality and complexity were tested, and it was found that with regard to quality HMMs performed worse overall,

and worse than LMMs when presented with irrelevant distractors, but better overall and better than LMMs when presented with relevant distractors. With regard to writing complexity, HMMs again performed better overall and better than LMMs with relevant distractors. This shows that the type of distraction is relevant when assessing the positive or negative influence of media multitasking. Related to this, media multitasking has also been linked to poor academic performance (Hembrooke and Gay, 2003) and knowledge acquisition (Lee et al., 2012), which again has implications when trying to follow a television show requiring sustained attention.

Media multitasking has also been linked with mental health and self-esteem issues. Pea et al. (2012) found that media multitasking was associated with feelings of negative well-being in 8–12 year-old girls, and Becker et al. (2013) found that increased media multitasking activity was associated with more self-reported symptoms of social anxiety and depression. Becker et al. (2013) concluded that media multitasking could present a risk factor for mood- and anxiety-related mental health problems.

2.3.2.2 Why do People Media Multitask?

Even though there are a number of potential downsides to media multitasking, people still choose to do it regularly. Gil de Zúñiga et al. (2015) found that people used second screens when watching the news in order to seek information and to participate in online discussions via social media. Social media has also been shown to be a popular use of second screens in other work (Courtois and D'heer, 2012; Schirra et al., 2014; Holz et al., 2015). Wang and Tchernev (2012) found that while media multitasking could result in negative cognitive effects (e.g. misguided attempts at increased productivity), it

can satisfy emotional needs such as entertainment or relaxation. Brasel and Gips (2017) investigated in-the-moment motivations for switching between content by looking at visual cues. They found that lower-level visual cues such as motion and luminance can encourage switches towards media, and high-level perceptual cues such as faces and people can discourage further switching away from media. They also found that media breaks, such as the switch from show to commercials and from commercials to show, can drive switching behaviour.

There is also evidence that media multitasking propensity may be a trait. Ophir et al. (2009) developed the Media Multitasking Index to classify people as high or low media multitaskers. The results of their experiments, where participants completed a number of tasks to asses different elements of cognitive control, showed that high media multitaskers struggled to filter out irrelevant stimuli during the tests. They unexpectedly found that participants classified as high media multitaskers actually performed worse on task-switching tests, suggesting that this group find it difficult to focus an a single stimulus when competing stimuli are present. This was further investigated by Sanbonmatsu et al. (2013), who also found that levels of media multitasking were negatively correlated with actual mutitasking ability. Furthermore, they found that participants self-reported multitasking ability was greatly inflated when compared to actual ability. This could suggest that high media multitaskers are more likely to engage in multitasking, because they believe they are good at it and will generate reward, but in reality their performance will be worse. However, it has been shown that people tend to overestimate themselves when asked to judge their own characteristics and abilities in general (Dunning et al., 2004).

Loh and Kanai (2014) built on the work of Ophir et al. (2009) by analysing brain images of both high and low media multitaskers. This work found that high media multitaskers had smaller grey matter density, further suggesting that frequent media multitaskers could have psychological and physiological differences when compared to those who media multitask infrequently.

2.3.2.3 Companion Apps

One focus of media multitasking has been the “companion app", referring to an application that is designed to accompany a TV programme in a way that enriches the viewer’s experience, most often by utilising a second screen (typically a mobile device). Such applications may be designed to be used at the same time as the programme is viewed, or at a time when the user is away from the programme. Cesar et al. (2008) describes how second screen can support four types of functions for content: control, enrich, share and transfer. Companion apps in various forms now exist for many programmes. In the past, websites accompanying the program provided some of the same enriching functionality as companion apps, often serving as a reference source not specifically intended to be used while viewing.

Recent years have seen an explosion of high-quality, multi-part television series. The large variety of critically-acclaimed serials such as HBO’s *The Wire* and *Game of Thrones*, AMC’s *Breaking Bad*, and BBC’s *Sherlock* have led some critics to declare a new “golden age" of television[5]. Each episode in this type of series can change the plot dramatically, which requires viewers to dedicate themselves to watching the entire series to fully appreciate the

[5] `http://www.cbsnews.com/news/welcome-to-tvs-second-golden-age/` [Accessed 18th July 2018]

intricate plot lines. This can make it difficult for casual viewers to immerse themselves if entering the storyline part way through.

Sometimes viewers need reminding about certain details of the plot of a show, especially if it has been a week since they saw the last episode. Serials like *Game of Thrones* are notoriously confusing for some, featuring multiple story threads — some intertwining and some independent — with many different factions and characters. Murray et al. (2012) recognised this problem and created the Story-Map iPad companion app, which used the TV series *Justified* as a test series. In their companion app, viewers could view different characters and their relationships in the form of a graph, then view important clips regarding those relationships. Viewers could also recap on story arcs and view important showdown scenes. This work was built upon through the *Game of Thrones Companion* app (Silva et al., 2015), which provided highly synchronised and contextualised information to viewers through spatial visualisations. Preliminary usability testing showed that when compared with the existing *HBO Go* companion app provided by the HBO network (which provides only general, decontextualised information), viewers were more accurately able to recall character relationships and even identify previously unseen characters when using the app developed by Silva et al. The authors note the importance of synchronising the information to the relevant plotlines and settings, and acknowledge the potential for similar data visualisations to be utilised in other genres.

While highly designed companion experiences have been well explored in the literature, uptake of such applications is still relatively low, perhaps due to the effort required to set up the dual screen experience (Neate et al., 2017). Conversely, incidental device usage without the need for special companion

apps is very common (Holz et al., 2015; Dias, 2016), but has not received as much attention by researchers. Such a common behaviour would therefore benefit from further investigation.

2.4 Defining and Measuring Media Experiences

People watch video media for a number of reasons (Rubin, 1981), and these are mostly because it is a positive experience for them in some way. However, defining what makes a media experience "good" or "bad" can be a difficult because of its subjective nature, but such a mechanism is necessary in order to assess the effect of various technological interventions in a systematic way. Operationalising experience in the context of film and TV has previously looked at *presence*, or a sense of being located inside the media instead of in the real world. A range of media have been shown to exhibit presence, including virtual reality systems (Sanchez-Vives and Slater, 2005), games (Tamborini and Skalski, 2006), television (Lombard et al., 2000; Bracken and Pettey, 2007) and books (Schubert and Crusius, 2002; Gysbers et al., 2004). Measurement of presence has typically relied on self-reporting through questionnaires and monitoring physiological responses (IJsselsteijn et al., 2000; Lessiter et al., 2001; Lombard et al., 2009).

2.4.1 Experience and Screen Size

In the context of television and film, viewer experience has commonly been investigated in relation to screen size. This is very relevant today, as mobile device ownership is now very common. Coupled with increasing use of on-demand and catchup TV services, this has allowed people to watch TV on

a wide variety of devices other than the traditional living room television. However, this raises questions about the relationship between the screen size of the device and the experience the viewer has.

A series of studies have assessed the effect of screen size on various aspects of viewer experience. In a lab study by Lombard et al. (1997), participants watched content on either a 46- or 12-inch screen and completed a questionnaire afterwards. The results suggested that in some genres screen size had an effect on the responses. Furthermore, participants reacted more strongly to clips which contained shorter shots and sudden movements.

In an experiment by Reeves et al. (1999), participants viewed short clips that portrayed different emotions which were displayed on either a 56-, 13-, or 2-inch screen. Arousal was measured by monitoring skin conductance, and attention was measured using a heart rate monitor. The results showed that screen size could increase both attention and arousal, and that for the very large screen arousal was greatly increased when viewing exciting content.

Lombard et al. (2000) investigated how screen size impacted presence when watching point-of-view footage, which was measured using both self-reported questionnaires and skin conductance. Footage was played on either a 12-inch or 46-inch screen. It was found that participants experienced sensations suggesting presence when watching the footage, and participants watching the larger screen experienced these feelings to a greater extent. IJsselsteijn et al. (2001) also found that a larger screen elicited a greater sense of self-reported presence when viewing motion footage on a large screen.

Contrary to the above research, a study by Bracken and Pettey (2007) found that watching content on a 2.5-inch iPod screen led to a greater sense of presence then watching on a 32-inch TV screen. However, the authors note

that the level of presence may have been affected but the fact that watching video on an iPod was a novel experience for most of the participants at the time, and that the iPod audio was heard through headphones whereas the TV audio was not.

In summary, the above research shows that screen size often has an affect on various viewer experience metrics, including both self-reported and objective measures. Furthermore, larger screens typically lead to more intense responses.

2.4.2 IMMERSION

While presence typically only refers to the feeling of being located in a mediated world, the concept of immersion has been used to refer to a sense of being highly engrossed in a mediated experience across multiple dimensions. The term can be used in the context of multiple media, and is generally seen as a favourable quality for media to possess. However, there is no standard definition, and the term is sometimes used almost interchangeably with concepts such as presence, involvement, and engagement. A widely cited definition by Murray (2017) describes immersion in mediated experiences as follows:

> A stirring narrative in any medium can be experienced as a virtual reality because our brains are programmed to tune into stories with an intensity that can obliterate the world around us... The experience of being transported to an elaborately simulated place is pleasurable in itself, regardless of the fantasy content. We refer to this experience as immersion. Immersion is a metaphorical term derived from the physical experience of being submerged in wa-

> ter. We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality, as different as water is from air, that takes over all of our attention, our whole perceptual apparatus. (p. 124)

From this, a number of elements make up Murray's definition of immersion, including a sense of transportation to another reality, and a high level of attention to the media. This sense of transportation could be interpreted as presence, which has been the focus of research in a number of areas, including virtual reality systems (Sanchez-Vives and Slater, 2005), games (Tamborini and Skalski, 2006), television (Lombard et al., 2000; Bracken and Pettey, 2007) and books (Schubert and Crusius, 2002; Gysbers et al., 2004). However, while a sense of presence may be an important component of immersive experiences, it may not fully describe immersion. For instance, consider playing a game of Tetris, or watching a quiz show on TV — it is possible to be immersed in these media without necessarily feeling a strong sense of presence.

The high level of attention could be interpreted as engagement with the media, and could encompass the concept of *flow*, a state of intense involvement in an activity (Csikszentmihalyi, 1996). Flow describes a state of high concentration on a task, where the difficulty is perfectly matched to the skill of the person carrying out the activity, and actions become almost automatic. Examples include playing a sport (being "in the zone") and musical improvisation. Busselle and Bilandzic (2009) describe how states of flow can be experienced when viewing or reading media, through focus on comprehending the media and constructing a mental model of the narrative.

2.4.3 Immersion in Non-Participatory Media

Expanding on her definition of immersion, Murray (2017) described immersion as being participatory, i.e., when immersed, one should be able to perform tasks in the virtual world as if it were real. For media such as games, this seems like a reasonable expectation of an immersive experience — the player has agency to make decisions about the actions they wish to take. How is it then, that it is common for people to report feeling immersed in a book or film? Biocca (2003) defined the "the book problem" to ask how it was that people can report high levels of presence when reading books, even though books are low fidelity and do not involve sensorimotor stimuli, which is considered a large part of presence in the virtual reality domain. Biocca argues that the level of presence experienced does not sit on a two pole continuum between the physical space and the virtual space, and proposed a 3-pole model instead. This introduced the notion of "mental imagery space", and Biocca suggested that the brain uses imagery to fill in the missing pieces.

Definitions of immersion are sometimes linked to task performance, where states of flow (Csikszentmihalyi, 1996) (implying a high level of task performance) can feature. If immersion is partially defined in relation to task performance, how can one feel immersed in a book or a film if there is no real task to be performed, other than reading or watching? Sherry (2004) argues that these very act of interpreting the media do allow for states of flow to occur. In film, conventions of shot composition and editing are understood by viewers, making interpreting the message easy, and deviating from these established practices makes interpreting messages hard. Similarly, some books and other printed texts are more accessible than others. Furthermore, Sherry (2004) argues that there are varying skill levels with regard to interpreting media. For

instance, experimental films may be more difficult to understand than typical Hollywood fare, but this could be improved by watching more or taking film appreciation classes. Also, some media requires prior knowledge to be fully understood and interpreted, for instance watching the final episode of a serial drama. Someone viewing without watching previous episodes is likely to struggle to full understand the plot than someone who has.

2.4.4 IMMERSIVE EXPERIENCES IN COMPUTER GAMES

Immersion and related concepts have been defined for many different types of media. The field of computer games research has had success in attempting to define immersion, and a number of definitions exist to operationalise gaming experience. Much like presence, the term has often been used loosely and without any agreed definition, and so care should be taken to differentiate between these definitions (Cairns et al., 2014a).

A number of different concepts that have been used to operationalise player experience in addition to immersion, including flow (Sweetser and Wyeth, 2005), presence (Weibel et al., 2008), puppetry (Calvillo-Gámez and Cairns, 2008). Immersion especially is frequently seen as a highly desirable characteristic for a game to possess in gaming reviews an other media. Brown and Cairns (2004) sought to better define the concept by interviewing gamers about how they perceived immersion, found three progressive levels of immersion using a grounded theory approach: *engagement* — actually investing time and effort into the game in the first place; *engrossment* — becoming interested in the game world and appreciating it, becoming emotionally involved in the game and being less aware of real-world surroundings; and finally *total immersion* — the state of being entirely separated from reality and

existing in the virtual world.

Jennett et al. (2008) further worked towards defining the concept of immersion, as well as investigating how it could be measured. They identified three features of immersion: lack of awareness of time; loss of awareness of the real world; and involvement and a sense of being in the task environment. They also describe how some concepts are related to immersion but distinct:

1. *Flow*, a state of high concentration where the experience is optimal and ability is matched to the difficulty of the task — distinct from immersion because immersion because sub-optimality does not preclude immersion;

2. *Cognitive absorption*, a state of deep involvement in computer software — distinct from immersion because immersion is concerned with games, not software in general, and because it is possible to feel absorbed in software without being immersed;

3. *Presence*, the state of ‘being there" in the game — distinct from immersion because abstract games (e.g. Tetris) can be immersive without the user feeling present.

In order to measure their multi-faceted definition of immersion, Jennett et al. (2008) developed and validated the Immersive Experience Questionnaire (IEQ) to explore different aspects of immersion. It features 31 items, which covered five factors of immersion: Cognitive Involvement; Challenge; Control; Real-World Dissociation; and Emotional Involvement. This questionnaire has been widely used as a tool for quantitatively assessing experience in a number of studies, including investigating the effects of different

controllers (Cairns et al., 2014b), how challenge affects immersion (Cox et al., 2012), the effect of touch screen size on immersion (Thompson et al., 2012), and brain-computer interface games (Gürkök, 2012). The IEQ has also been adapted to other contexts such as public speaking anxiety (Wortwein et al., 2015), and games without graphics for visually-impaired players (Engström et al., 2015), suggesting a level of adaptability to other media.

While there are quantifiable methods of defining and measuring immersion in the context of computer games, methods for doing so in the context of film and television have not been widely researched. This leaves an opportunity for the development of tools to measure similar phenomena in the film and television domains, perhaps borrowing concepts from computer games research.

2.5 Conclusions and Research Directions

This chapter has examined literature pertaining to TV usage in the modern age, specifically looking at how on-demand video and media multitasking are changing televisual experiences. Furthermore, methods of quantifying viewer experience has also been reviewed.

In terms of concurrent TV and device usage, this literature review has uncovered a large body of research in the area of multitasking and interruptions, and the effects on cognitive performance they have. The majority of work has been in safety-critical settings and the workplace, due to the serious consequences that errors and reduced performance can have in these environments. While the effects of media multitasking and interruptions while watching TV are not as immediately life threatening or financially detrimental as in

these environments, there still a number of negative effects associated with this behaviour. Despite this, people are media multitasking more and more, though they may not be aware of these negative effects. Media multitasking behaviour is being actively encouraged, for example through the use of companion apps, TV programmes actively encouraging social media use by displaying relevant hashtags, and through devices such as Google's Chromecast[6] which encourages media selection on a device while content is playing on the TV. The situation is effectively summarised by Wang and Tchernev (2012), who posit that "cognitive needs are not satisfied by media multitasking even though they drive media multitasking in the first place. Instead, emotional gratifications are obtained despite not being actively sought. This helps explain why people increasingly multitask at the cost of cognitive needs". There is a need to for more work to be done to examine the effects of these types of modern viewing practices on the TV watching experience.

Investigating the effects of multitasking have typically looked at how performance and error rates are affected. However, watching television is a mostly passive task, for which there are few meaningful performance measures. Something that can be measured is the viewer's experience, in a similar way to the methods used in computer games research. Some research has looked at the viewer's level of presence, and other studies have measured arousal and skin conductance, but these do not give a broad view of viewer experience. The concept of immersion does however offer a multi-faceted view of experience in the field of computer games, and so could potentially be adapted to video viewing. Modern viewing practices, including the constant switching of attention brought about by mobile device usage, may have

[6] `http://www.google.co.uk/chrome/devices/chromecast/` [Accessed 19th July 2018]

an effect on how the viewer experiences, enjoys, or engages with the content they are watching, and so tools and methods of measuring this would be a valued contribution in this area.

In terms of on-demand viewing, this chapter has shown that use of on-demand services is definitely on the increase, but exact usage patterns and motivations are not understood. We know that people value the freedom and choice these platforms provide, allowing them to watch when and wherever they like, but the impacts of this cross-device, cross-location viewing, as well as other behaviours surrounding this, has not been sufficiently investigated from an HCI perspective.

When looking at methods used to examine these types of behaviours, much of the previous research has relied on self reporting in the form of surveys, interviews and focus groups. While these methods can be very useful to get a sense of general habits and practices, they are often not sufficient to give an accurate picture of very specific behaviours. For example, a person may be able to say they use their mobile device while watching TV regularly, but be unable to say exactly how many times they picked up their phone on a given day. For this reason, a mixed methods approach to investigating some of these behaviours presents a good course of action, combining quantitative and qualitative methods, and in the wild and lab based approaches. This will enable the construction of a rich picture of the prevalence, motivations and effects of such behaviours.

The following chapter begins the empirical work in this thesis, and uses an in the wild study to ascertain the prevalence of concurrent mobile device and TV usage.

CHAPTER 3

AN IN THE WILD STUDY OF CONCURRENT TV AND MOBILE DEVICE USAGE

The following publication is based on work featured in this chapter:

Jacob M. Rigby, Duncan P. Brumby, Sandy J.J. Gould, and Anna L. Cox. 2017. Media Multitasking at Home: A Video Observation Study of Concurrent TV and Mobile Device Usage. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video* (TVX '17). ACM, New York, NY, USA, 3-10. DOI: `http://doi.org/10.1145/3077548.3077560`

3.1 INTRODUCTION

Previous research has shown that using mobile devices while watching television has become a common activity (Foehr, 2006; Roberts and Foehr, 2008; Brasel and Gips, 2011). In the UK, communications regulator Ofcom found that 53% of UK adults regularly media multitasked in 2013 (Ofcom, 2013), and a 2014 report showed that 99% of adults media multitask at some point during the week, for an average of 2 hours and 3 minutes every day (Ofcom, 2014). Multitasking in the living room may not have safety implications as in aviation (Dismukes et al., 2001; Loukopoulos et al., 2003) or driving (Caird et al., 2008), or be directly detrimental to productivity as in workplace en-

vironments (Mark et al., 2016), but nonetheless this changing behaviour is of interest to a number of groups. For example, TV networks who wish to retain their audiences through increased engagement, content producers who wish to create better TV experiences for viewers, and advertisers who may be concerned with viewers using media multitasking as an advertising avoidance strategy.

Studies conducted to better understand the prevalence of media multitasking behaviour have typically relied on self reporting from participants (e.g. (Foehr, 2006; Rideout et al., 2010; Vanattenhoven and Geerts, 2012; Voorveld and van der Goot, 2013)). However, it is possible that people are poor at estimating just how much time they are spending on digital devices and so misrepresent the extent to which they media multitask. In response to this, a small number of observational studies have also been conducted, for instance using direct observation (Voorveld and Viswanathan, 2014) or sensor-based telemetry (Holz et al., 2015). While these methods are more accurate than self-report data, fine-grained video data has the potential to offer a more detailed and nuanced impression of behaviour in situated contexts (e.g. (Brown et al., 2013; Pizza et al., 2016)).

In the study presented in this chapter, video observation was used to establish a detailed and accurate understanding of mobile device usage and TV consumption in the home. To do this, the behaviour of four households was video recorded over a 72 hour period. Video observation was chosen for a number of reasons. Firstly, to remove the possibility of people misremembering their behaviour — device usage can be a short occurrence that happens frequently, such as checking one's phone for new messages, and accurately recounting all of these via self report methods could be difficult. Secondly,

to allow for accurate reporting of durations that participants were performing particular tasks (e.g. watching TV, using their devices). Post-hoc video analysis allows for frame-by-frame precision, and therefore an accurate account of exact durations. Finally, video analysis allows researchers to observe other events in the living room that participants may not think are important, and therefore do not report.

3.2 Related Work

As mentioned above, a number of previous studies have investigated TV media multitasking practice, often using self reporting methods to collect data (e.g. Foehr (2006); Rideout et al. (2010); Vanattenhoven and Geerts (2012); Voorveld and van der Goot (2013)). Such methods facilitate easy data collection and can give a general view of many people's behaviours and habits. However, self-reported data can be inaccurate and lack granularity. This has led to a need for observational studies to be performed to obtain an accurate view of everyday media multitasking.

In order to better understand when people used their devices while watching TV and exactly what they were doing, Voorveld and Viswanathan (2014) conducted an analysis of observational data obtained by directly observing participants from the USA. They found that media multitasking was most prevalent when watching sport and channel surfing, during morning and afternoon, and when individuals were watching television alone. Observations were made every 10 seconds to give a fine-grained view, but were not video recorded and so could not be played back for further post-hoc analysis. Activities performed on mobile devices were also not recorded in detail.

Another situated study was performed by Holz et al. (2015), who used device logging system installed on participants' phones and tablets. Various information was logged, including app launches, websites visited. This was cross-referenced with the TV programme being watched at the time, which was established by using audio fingerprinting. It was found that the majority of device usage was unrelated to the programme they were watching, but that device usage did differ based on the type of show being watched. Furthermore, device usage seemed to correspond to the the events in the show — for example, device usage dropped towards the end of the show when watching crime dramas when the story is resolved. While this study also gave a very fine-grained view of device usage, it was not video recorded and so the physical behaviour of participants could not be studied.

Rooksby et al. (2014) also investigated parallel TV and device usage by using a device logging system, which was augmented with video observation. This work was further expanded on (Rooksby et al., 2015), but the results focus more on the social implications of how media multitasking affects home life, presented as a small number of vignettes. Furthermore, the participants had to manually turn on the cameras every time they wished to record data, meaning naturalistic data may have been omitted and the fact they were being recorded would have been fresh in their minds.

While the research by Holz et al. (2015) and Rooksby et al. (2014) is valuable in establishing media multitasking habits in the home, it leaves open an important question of what drives these behaviours. Is it the case that media multitasking behaviour reflects situational factors, such as becoming bored with the television programme or wanting to look up some relevant information, or is it that some people are more inclined to media multitask

than others?

Ophir et al. (2009) argue that a person's propensity to media multitask is not driven by situational levels of engagement but more reflects a stable individual trait — some people just prefer to media multitask while others do not. To support this claim, Ophir et al. developed the Media Multitasking Index (MMI), a measure used to establish individual media multitasking preferences. Research using the MMI has investigated cognitive differences between media multitaskers (Alzahabi and Becker, 2013; Loh and Kanai, 2014; Lottridge et al., 2015). However, little research has been done to investigate this specifically in the context of concurrent TV and phone usage in the home. In other words, are those people that self-report a high MMI actually more inclined to use a device while watching television?

The study presented in this Chapter further investigated individuals' media multitasking behaviour through means of video observation over three evenings. Two surveillance cameras were used. One recorded participants' seating areas and televisions to allow for a greater understanding of physical behaviour and other non-phone and tablet tasks that may occur, and another recorded the television to allow us to see when the TV was turned on and what was being watched. The participants also completed an MMI questionnaire to measure general media multitasking preferences, which could then be compared with real-world multitasking behaviour.

3.3 PARTICIPANTS

Five households were recruited through opportunity sampling. Each household was required to have a dedicated TV set. At least one person in each

household was required to be a regular TV watcher (at least one hour per evening), who was also required to have a smartphone as their primary device. Households were paid £75 for three evenings of continuous participation.

Household A consisted of a male and female couple, aged 67 and 56 respectively, living in a house in Worcestershire, England. Their TV was located in their living room.

Household B consisted of three cohabiting professional females aged 26, 27, and 29, living in a shared flat in Oxford, England. Their TV was located in their living room area, which adjoined the kitchen and dining area.

Household C consisted of a male and female couple, aged 58 and 59 respectively, living in a house in Worcestershire, England. Their TV was located in their living room area, which adjoined the kitchen and dining area.

Household D consisted of two parents (39 and 45 years old) and their three children (17, 12, and 9 years old) living in a house in Oxford, England. Their TV was situated in their living room.

The final household, **household E**, consisted of two parents in their thirties and their three young children (all under 8 years old) living in a house in Worcestershire, England. Their TV was situated in their living room. Due to technical issues, large portions of the data collected from this household were unusable. For this reason, household E was excluded from this study. The mean age of the remaining participants was 37 $(SD = 19.88)$.

3.4 SETTING AND PREPARATION

This observational study took place in participants' homes in the room where they typically sat to watch a dedicated TV (i.e., not on a tablet, smart phone or computer). Logging software in the form of the AWARE framework client[1] was installed on the participants' smartphones and tablets and set up to log important data:

- Screen on/off status
- App launches
- App installations
- Keyboard data
- Times when calls are made or taken, and a unique identifier for the other party (personal data was hashed for privacy)
- Times when text messages are sent or received, and a unique identifier for the other party (personal data was hashed for privacy)
- Battery level
- Network information (e.g. if the participant is connected to wifi)

3.5 MATERIALS

For each household participating, a small mains-powered surveillance camera (see Figure 3.1) was used to record a view of the TV for the purposes of identifying when the TV was turned on and for programme detection, and another identical camera was angled towards the seating area to record the participants themselves. Video footage was recorded onto micro SD cards.

[1] http://www.awareframework.com/ [Accessed 17th July 2018]

Figure 3.1: Surveillance camera used to record participants and televisions.

Participants were required to use their own smartphones and tablets for this study, with one Android smartphone being the minimum. Participants were expected to use their own dedicated televisions for viewing; this study did not record viewing on other devices.

The study utilised a pre-session questionnaire to collect demographic and technology usage data, and the Media Multitasking Index questionnaire from Ophir et al. (2009) to indicate individual media multitasking preferences in general.

3.6 Ethical Considerations

Logging personal devices presented a number of ethical and privacy considerations. In order to minimise privacy concerns, only data relevant to this study was collected — the AWARE framework used for device logging is highly customisable, and allows researchers to specify exactly what is logged from a range of sensors common to many smartphones and tablets. Furthermore, while keyboard data was collected to allow web searches to be reported, password data was not. Also, times at which communications occurred by text

message and voice call were recorded but no data sent or received was retained (though it is unavoidably possible to infer sent data from the keyboard logs), and the source or target of this was stored in the form of an irreversible hash key. All data collected was anonymised and stored remotely in a protected MySQL database.

The presence of surveillance equipment in people's homes presented some ethical and privacy issues. While ethical clearance was given to recruit households with participants under the age of 18 with parental consent, it was possible that visitors under the age of 18 could become part of the study. It was also possible that adults could unknowingly participate. For these reasons, each household was required to display a poster informing visitors of the study taking place in a prominent position near the property entrance.

3.7 PROCEDURE

Once participants had been recruited, a suitable time was arranged with them for the researcher to visit their property and install the surveillance cameras, as well as helping with the installation of the logging client if necessary. The clocks across all devices were also synchronised. During this session, the participants were shown the information sheet and given the opportunity to ask questions, then asked to sign a consent form. Finally, they were asked to fill in the questionnaire about demographics and technology usage. Once everything was set up, three evenings' worth of data were logged. These were consecutive evenings where possible, though as some participants said they would not be in the house during that time, some were not consecutive. A time was also agreed for the researcher to collect the equipment. During

this time, the participants were asked to fill out the MMI questionnaire. The researcher also assisted with the removal of logging software (if necessary) and the participants were paid. Participants were also given an opportunity to make any closing comments about the study and their behaviour.

3.8 RESULTS

In total, 24 hours' worth of footage for each camera was collected per household, resulting in a total of 192 hours of footage (96 hours for the seating cameras and a further 96 hours for the TV cameras). The cameras automatically split the footage into consecutive 30 minute sections. As this study is only concerned with behaviour during TV time, the recordings showing the TV were first reviewed in order to discard sections where the TV was off. The corresponding footage of the seating areas were also discarded for these times. Once all of the sections with TV activity were identified, the corresponding clips were combined into a one file per evening to keep file sizes manageable — one file for the camera looking at the TV and one for the seating. Both video feeds for each evening were then synchronized and coded using Chronoviz[2]. During the coding process, the video was first annotated to show when the TV was on, then all further codes were performed inside this time (again, this study is only concerned with what happens during TV time). Within this, the video was annotated to show when each participant was present and when they were using any mobile devices. Further to this, any other notable or interesting events were also annotated, such as use of on-demand services.

The total amount of time participants' TVs were turned on across all

[2] `http://www.chronoviz.com/` [Accessed 17th July 2018]

Household	Duration
A	19:07:52
B	07:23:05
C	17:48:07
D	10:04:49

Table 3.1: Duration TV was on in households

households was 54:23:53 (M = 13:35:58, SD = 05:45:08). This equates to about 57% of the total video recordings. As can be seen in Table 3.1, there was considerable variability in the total amount of time that the TV was on in each household (*range*: 07:23:05 - 19:07:52).

Results for individual participants can be seen in Table 3.2. It can be seen that for individual participants, mean total time present when the TV was on was 06:38:48 (SD = 05:49:10, all times HH:MM:SS), mean total device usage when present was 00:41:21 (SD = 00:54:24), and mean number of uses was 7.67 ($SD = 8.26$). It can also be seen in Table 3.2 that there were large individual differences in media multitasking habits between participants. Some participants did not use their devices at all (C1, D3 and D4) while others used their devices for nearly a quarter of the time they were watching (C2). Furthermore, some participants favoured shorter, more frequent uses while others exhibited fewer but longer uses.

To illustrate this, we can look at the two participants with the highest percentage of TV time spend using their devices, C2 and B2. Participant B2 used their device 28 times for an average of 1 minute and 27 seconds, at a rate of .115 uses per minute, whereas C2 only used their device 10 times, but for an average 11 minutes and 16 seconds at a rate of .014 uses per minute. This can be seen in Figure 3.2, which shows a snapshot of behaviour over a circa one-hour period. However, these individuals did not necessarily sustain the same usage pattern uniformly over the course of their viewing.

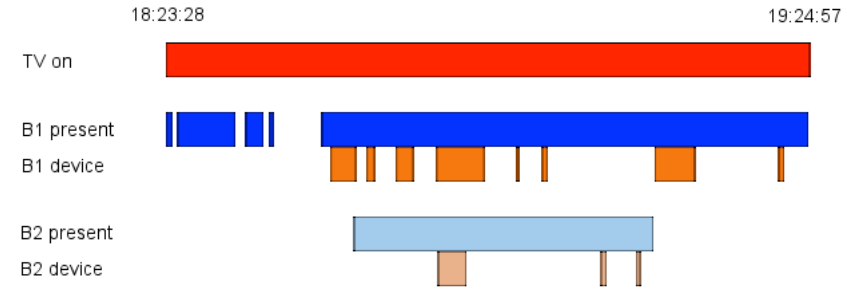
Participant	Age	Time present when TV on (% of TV on time)	Device use while present (% of time present)	Mean time per use	Number of uses	Uses per hour of time present	MMI
A1	67	17:27:05 (91.22%)	00:51:12 (4.89%)	00:06:24	8	.46	2.27
A2	56	17:30:05 (91.48%)	02:24:02 (13.72%)	00:13:06	11	.63	2.82
B1	27	04:41:28 (63.53%)	00:40:46 (14.48%)	00:02:24	17	3.62	2.7
B2	26	04:04:10 (55.11%)	00:40:24 (16.55%)	00:01:27	28	6.88	4.02
B3	29	02:02:41 (27.69%)	00:05:47 (4.72%)	00:00:58	6	2.93	2.16
C1	58	08:04:50 (44.39%)	00:00:00 (0%)	00:00:00	0	0	0
C2	59	11:39:30 (65.49%)	02:39:45 (22.84%)	00:15:58	10	.86	2.63
D1	39	03:14:43 (32.19%)	00:25:52 (13.29%)	00:03:42	7	2.16	2.68
D2	45	01:30:09 (14.90%)	00:05:52 (6.51%)	00:01:57	3	2	1.04
D3	17	03:59:27 (39.59%)	00:00:00 (0%)	00:00:00	0	0	1.18
D4	10	01:22:03 (13.56%)	00:00:00 (0%)	00:00:00	0	0	1.37
D5	12	04:09:30 (41.25%)	00:22:32 (9.03%)	00:11:16	2	.48	6.45*

Table 3.2: Results for all participants, grouped by household (all times HH:MM:SS). *Note: value marked * denotes anomalous value removed from analysis*.

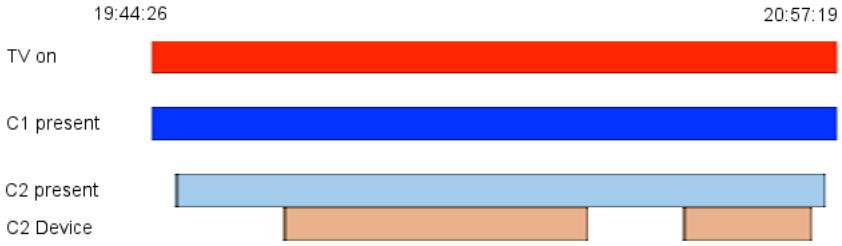
Unfortunately, due to technical problems the data collected from the logging software was mostly incomplete for all of the participants. Therefore it is not presented here.

3.8.1 THE TV AS A MEETING POINT

Previous literature has shown how the living room and television are used as a meeting place where family and friends gather to be with one another, both to watch programmes together and also to do other tasks while not actively watching (Lull, 1980; Kubey, 1986; Logan et al., 1995; D'heer et al., 2012).



(a) Example of frequent but short device uses form household B.



(b) Example of longer device uses from household C.

Figure 3.2: Comparison of different device usage patterns over a short period.

Figure 3.3 shows some examples of participants using the TV as a background to other activities performed together. Times when there were more than one person present while the TV was on accounted for 28:59:13 (53% of total TV time) across households.

3.9 Other Observations

Three households were recorded completing some kind of work in front of the TV, which can be seen in Figure 3.4. Participant A2 was recording completing some accounting work, B1 and B3, both teachers, were recorded marking work, and the children from household D were recorded doing homework.

(a) Participant wraps present in front of TV.

(b) Participants sleeping in front of TV.

(c) Participant knitting in front of TV.

(d) Participants reading together in front of TV.

Figure 3.3: Some examples of participants performing tasks with others present in front of the TV.

3.9.1 DEVICE USAGE AND MMI SCORE

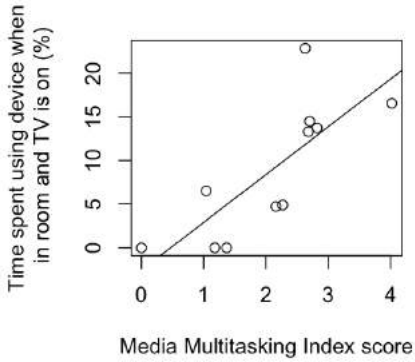
The relationship between MMI score and total device usage (as a percentage of time participants are present while the TV is on), and between MMI score and device uses per hour, was also considered. One participant was removed from these analyses due to misunderstanding the MMI questionnaire, which led to an artificially high value. Across the remaining sample of 11 participants, mean MMI score was 2.08 ($SD = 1.1$). Figure 3.5 contains scatterplots showing the relationship between MMI and time using device and device uses per hour. As can be seen in these figures, participants who had a higher MMI score tended to use their devices for longer periods in total when in front of the TV ($r^2 = .60$), and use their devices more frequently ($r^2 = .48$). Statistical analyses support these observations, showing that MMI score was a significant predictor of time spent using devices in front

(a) Participant A2 organises some financial paperwork, while A1 watches TV.

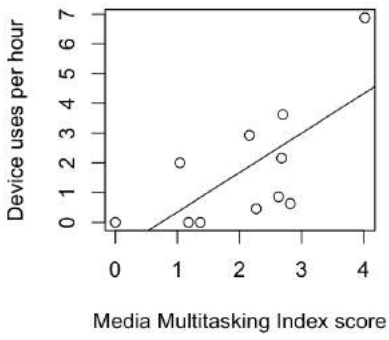
(b) Participant B1, a teacher, marks class work in front of the TV. B2 uses her phone.

(c) One of the children from household D looks over their homework.

Figure 3.4: Participants doing work in front of the TV.



(a) Scatterplot of device usage time against MMI.



(b) Scatterplot of device uses per hour against MMI.

Figure 3.5: MMI correlation scatterplots

of the TV, $F(1, 9) = 13.66, p = .005$, and number of devices uses per hour $F(1, 9) = 8.36, p = .018$. In other words, MMI scores were predictive of people's actual observed media multitasking behaviour at home.

3.9.2 USE OF ON-DEMAND SERVICES

On-demand services, including paid-for subscription services, free catch-up services, and viewer-recorded content, were used by two of the four households. Across all households, on-demand viewing accounted for 14:06:42, 25.9% of all viewing time. Household A used Netflix to watch content on their smart TV for a total of 07:15:00, 37.9% of the time the TV was on. Within this, the content was paused for 00:35:05, or 8.1% of the time. Household D used their Tivo recorder to watch recorded programmes, as well as the BBC iPlayer service. The total time they spent watching on-demand services was 06:51:42, 68% of the time the TV was on. Within this, content was paused for 00:23:37, 5.7% of the time.

Pausing content was a welcome feature of on-demand content, as it allowed participants to leave the room for short periods without losing their place in the programme. This time often seemed to be used to collect food from another room, or to prepare a drink to bring in to the TV area. Table 3.3 shows a small scene in household A, where participant A2 leaves to the room to serve up the evening meal and bring it into the living room, while participant A1 pauses it for her so she doesn't lose her place in the plot. Another scene from a different evening shows A1 using this pause time to use his device — A2 leaves to make a drink while A1 pauses again, and during this time he starts to use his device. When A2 returns with the drinks, A1 puts down his device and resumes the programme (see Table 3.4). Similar device usage by A1 when content is paused was observed on two other occasions.

Household 1 also exhibited what could be considered "binge watching", or consuming a large amount of content in a single sitting Jenner (2015). Dur-

ing two of the three evenings recorded, they watch multiple episodes of the Norwegian crime drama *Lilyhammer*. Over the course of a Saturday evening they watch six episodes, four back-to-back followed by a break, then another two back-to-back. Watching these types of programmes was discussed, and the participants said they are big fans of Scandinavian serial dramas, citing *The Bridge* as an example. Participant A1 said that "They are really interesting. It's subtitled so you really have to watch, erm, what you're doing... You need to watch them constantly".

	18:04:03 - A2 (right) gets up to serve up the evening meal.
	18:04:07 - A2 tells A1 to continue watching, but A1 insists on pausing.
	18:06:53 - A1 waits patiently for a few minutes and retrieves the lap trays.
	18:15:48 - A2 enters with the meals. Once both seated, they discuss the meal briefly.
	18:16:27 - After the brief discussion, A1 resumes the show and they continue to watch while eating.

Table 3.3: Short vignette showing on-demand content being paused while a participant goes to get some food.

	17:39:53 - A2 (left) gets up to collect the cups and make a drink. A1 pauses the content.
	17:40:05 - A2 sits waiting for a while.
	17:40:13 - A2 picks up his tablet and starts to use it.
	17:42:35 - A1 re-enters with the drinks. A1 closes his tablet and puts it back on the chair, then resumes the paused content.

Table 3.4: Short vignette showing on-demand content being paused while a participant goes to get some food.

3.10 DISCUSSION

The results of this study revealed large individual differences in concurrent TV watching and device use habits between participants. Some participants were frequent device users, while others used no devices whatsoever. Looking at the demographic make-up information of the households, the household made of females in their late 20s (household B) recorded the largest

proportion of concurrent mobile device usage while watching TV. This may be in line with expectations that millennials use more technology than older people (Carrier et al., 2009). To further understand this, their domestic circumstances were examined, and found that two of the participants had partners that lived in different cities, whereas all of the other households consisted of couples or children. It is possible that this increased usage could be through messaging their partners. This is supported by the high number of uses recorded, which supports the type of phone checking pattern resulting from asynchronous communication.

The MMI questionnaire asks participants to assess their general multitasking preferences across a range of media. The results show that MMI score was a good predictor of actual media multitasking behaviour. It is interesting to note that the MMI scores of participants in this study were considerably lower than that reported in previous studies that used the MMI — mean MMI score for participants was 2.08, compared to 4.38 in (Ophir et al., 2009), 3.82 in (Lui and Wong, 2012) and 4.07 in (Alzahabi and Becker, 2013). This difference in MMI scores between studies is most likely due to the sample of older participants, compared to the participants in previous studies, which were mainly college students in their early twenties. Both media multitasking and general multitasking has been found to be less common among older generations (Carrier et al., 2009; Duff et al., 2014), which would explain this discrepancy. In general, the results suggest that the rate of media multitasking in the home might vary considerably between households.

Different patterns of device usage were observed, ranging from fewer uses lasting for long periods, to many short uses. This raises interesting questions as to how media multitasking is defined. Multitasking behaviour that sits at

different points on the multitasking continuum (Salvucci et al., 2009) was observed. Frequent, shorter uses could be considered instances of concurrent multitasking, were two tasks are being performed simultaneously (e.g. talking and driving). On the other hand, longer uses with fewer switches could be considered instances of sequential multitasking, where only one task at a time is being actively performed before switching to the other task. This means that when the user is purely concentrating on their device, the TV is likely blurring into the background and they stop following what is happening on the TV. Indeed, in the data collected for this study there were many occurrences of the TV on in the background while the participants were engaged in other activities (e.g. those shown in Figure 3.3 and Figure 3.4). Such nuances may be difficult to convey when using self-reported methods or log analysis to establish how prevalent media multitasking really is, which may call into question the veracity of such methods — simply asking participants if they use their phones and tablets while watching TV may not give a full picture of their behaviour.

The impact of the types of media multitasking observed should also be considered. In general, attempting to perform multiple tasks concurrently can impede performance (Monsell, 2003), and it may be that negative effects also transfer to the TV domain, for instance in terms of reduced engagement (Holmes et al., 2012). It has also been shown that media multitasking specifically can also have detrimental effects, for instance when attempting to work in front of the TV (Brumby et al., 2014; Lottridge et al., 2015), and there is evidence to suggest that those who media multitask the most are often the worst at it (Ophir et al., 2009).

The data showed that of the entire time the television was on across house-

holds, more than one person was present for at least half of the time. In line with prior research, this shows that the television was very much a meeting point for the households in this study (Kubey, 1986; Logan et al., 1995; D'heer et al., 2012). Watching TV was frequently a social activity, and in addition to coming together to watch programmes together, the participants would leave the TV on while doing other tasks seemingly just to be together. This suggests that although the television landscape has changed and become fragmented, people still value the social aspect of sitting together whether or not they are watching TV together.

A large amount of on-demand service usage was observed, such as Netflix, BBC iPlayer, and viewer-recorded programmes, which accounted for 26% of all viewing. However, this also seemed to be tied to personal preference, as only two of the four households engaged in this. The Nielsen Company (2016a) found that 43% of people globally are watching on-demand at least once per day, and so more of this behaviour might be expected in the future.

As a result of the use of on-demand services, an occurrence of binge watching was also observed, where participants watch multiple episodes of the same programme in succession. While there has not been a large amount of academic research on this, it can be regarded as a direct product of having instant access to entire series of content (Conlin et al., 2016). While only a a single occurrence of this was observed, it is a well-known concept in the public consciousness. Netflix describes binge watching as "clearly the new normal"[3] and so may be more common than this data suggests.

[3] `http://media.netflix.com/en/press-releases/netflix-binge-new-binge-scale-reveals-tv-series-we-devour-and-those-we-savor-1` [Last Access 23rd Jan 2017]

3.10.1 LIMITATIONS

This study has described an analysis of video data that gives an interesting snapshot of daily mobile device use in front of the TV. This has allowed these media multitasking moments to be isolated and analysed to give a better understanding of how often and how long they occur in the home. Due to the high level of individual differences observed across participants, and the small sample size, it could be argued that it is difficult to draw strong generalisable conclusions. However, the results do provide good evidence of a strong link between self-reported MMI and observed device usage.

A number of difficulties were had with the technology. Setting up the cameras was not a trivial task, and then when they were installed they still often malfunctioned. Data from an entire household had to be discarded twice due to malfunctioning camera equipment. When the cameras did record, the image quality was lower than desired, which made analysing small movements, such as glances, difficult or impossible. The cameras are intended for basic surveillance purposes and so were not entirely fit for the purposes of this study, but choice was limited by the need to have mains-powered, "set-and-forget" equipment, as well as by cost. Furthermore, the logging software used was at times unpredictable. It only supported Android devices at the time, and even when participants did have Android devices it would not install on some of them. This meant that data could not be recorded for Apple users or those with incompatible Android hardware. Even when it did install, it seemed to only partially log events, resulting in incomplete data which could not be used.

While the video data can provide a rich perspective into events, these types

of data can take significant amounts of time to process and analyse. In this study, a first pass of the video files was used to identify segments without any activity, which were then discarded. The remaining videos then had to be re-encoded and stitched together, which itself is a long process. Once the video file preparation was complete, corresponding files from both camera perspectives had to be manually synchronised due to differences in the camera clocks. The actual video analysis and annotation required a number of passes in order to make sure each type of relevant behaviour and each participant was accounted for. Even though the data was ultimately unused, importing the log data also required preparatory work — SQL queries had to be written to extract the data of interest, and timestamps had to be adjusted to synchronise with the video coding software.

3.11 CONCLUSION

In this study, four households were observed watching TV for three evenings, with cameras observing both the participants and the television. During the 96 hours of observation across each household, participants' televisions were turned on for 54.4 hours (57% of the time), with a mean of 13.6 hours. The results suggest that viewing and device usage habits for individual participants were highly variable. Some participants watched a lot of TV while others watched less. Some participants frequently used their mobile device while other did not use devices at all. MMI was found to be a good predictor of observed media multitasking, taking into account both total device usage and the number of uses, suggesting that people who media multitask with their phone and TV probably do so with other media too. Differing patterns of device use in front of the television were also observed, which could be classified at

different points on the multitasking continuum.

Observations conducted for this study confirmed a common theme in prior studies that the television has a social function in the household. It was observed that for 53% of the total time the TV was turned on, more than one person was present. The TV was used as a household hub, with participants gathering around it even when focusing on other tasks, such as work, using mobile devices, and reading. In addition to these expected behaviours others were also observed, such as the use of on-demand services and binge watching. This suggests that although it is changing, the TV remains a focal gathering point for family life in the home.

The next chapter continues the theme of collecting in the wild data to ascertain the prevalence of on-demand service usage, which made up 26% of viewing time in the study presented in the current chapter. However, this time a diary study is used to establish the level of interaction with these services across all devices and locations, both inside and outside of the home.

CHAPTER 4

A DIARY STUDY OF ON-DEMAND VIDEO VIEWING HABITS

The following publications are based on work featured in this chapter:

Jacob M. Rigby, Duncan P. Brumby, Sandy J.J. Gould, and Anna L. Cox. 2018. "I Can Watch What I Want": A Diary Study of On-Demand and Cross-Device Viewing. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video* (TVX '18). ACM, New York, NY, USA. DOI: `http://doi.org/10.1145/3210825.3210832`

Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J.J. Gould. 2018. Old Habits Die Hard: A Diary Study of On-Demand Video Viewing. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI EA '18). ACM, New York, NY, USA, Paper LBW016. DOI: `http://doi.org/10.1145/3170427.3188665`

4.1 INTRODUCTION

The study in Chapter 3 took at detailed look at people's media and device use practices in their living rooms. One of the findings of this study was that on-demand services accounted for over a quarter of all viewing time recorded. This supports the findings in Ofcom's 2016 communications market report and their "digital day" research, which found that use of on demand video

services is growing. Subscriptions to paid subscription services, such as Netflix and Amazon Video are rising year on year, and total viewing of both free and paid for on demand services, including viewer-recorded content, increased in the UK from 26% in 2015 (Ofcom, 2015b) to 32% in 2016 (Ofcom, 2016a). However, Ofcom's research also shows that mobile devices are becoming increasingly popular as a platform for watching video instead of the traditional TV set, with 21% of the online population watching on a phone, 23% on a tablet and 33% on a computer at least once a month. This rises in the 16-24 and 25-34 age group. Due to the focus on traditional TV viewing, on-demand viewing on mobile devices was not captured in the study in Chapter 3, but nonetheless could account for significant amounts of viewing time outside of the living room setting. This motivated the need to further examine on-demand service usage practices in detail.

With the rise in popularity of on-demand services, what impact is this having on how people consume video content? Large-scale surveys, such as those from Ofcom (2017) and the Nielsen Company (2016a), are useful for giving a general impression of the popularity of on-demand services. However, such surveys can lack the necessary level of granularity to unpick what is driving these viewing practices. For example, Ofcom (2017) suggest that the TV is still the most popular way to view, but that 21% of the online population choose to watch on a phone, 23% on a tablet and 33% on a computer at least once a month. These surveys provide useful data but do not ask about important contextual and situational factors that might be affecting why people choose to watch on one device over another. Does this decision of which device to use depend on where the person is watching? Who they are watching with? What time of day they are watching? What they are watching?

This chapter focuses on what motivates such decisions that people take when viewing.

In this chapter, the results of a diary study conducted to provide a detailed snapshot of everyday viewing behaviours using on-demand services are presented. In particular, it focuses on on unpicking differences in viewing behaviour on handheld mobile devices and non-mobile devices. The results of in-depth interviews that were conducted to better understand these household viewing diaries are also presented. These interviews focused on understanding what motivated different viewing behaviours: why people choose to view on particular devices, watch in different locations, and watch alone or together. An understanding of people's positive and negative perceptions of on-demand services is also developed.

The diary study method was chosen for this study to enable the collection of data in a wide variety of contexts and environments, and over a medium-to-long period of time. While video observation has a number of benefits (see Chapter 3), the necessity for specialist, non-portable equipment means that it is not a feasible method to use in the multitude of environments that on-demand viewing can occur in. Furthermore, free text questions and interviews can allow researchers to gain a more qualitative understand of participant behaviour.

4.2 BACKGROUND

Prior to the advent of on-demand video services, viewers had limited choice in what they watched, and when and where they watched it. Previous research from this era gives us an insight into "traditional" linear TV viewing practices.

For instance, evening viewing after the working day was especially popular, particularly in the living room (Taylor and Harper, 2003); people watched TV regularly, often for multiple hours per day (Logan et al., 1995); and personal viewing schedules were based around broadcast schedules, which in turn influenced other household activities (Gauntlett and Hill, 2002).

Considering the current popularity of on-demand video services and mobile viewing, surprisingly little HCI literature has addressed it. Section 2.2 provides a review of the most relevant literature in this area. In summary, previous research has provided useful perspectives into how people use on-demand video services. A common theme is that people value and take advantage of freedom from the broadcast schedule, allowing them to choose viewing times that suit them. Furthermore, much of the literature reveals a strong social element to watching TV. Be it watching together, selecting content together, or discussing shows with friends and colleagues, social factors appear to affect viewing practices. Prior research also gives us a limited perspective on viewing on mobile devices, specifically regarding motivations for doing so, which are many and varied.

While the phenomena of on-demand viewing and mobile viewing are strongly coupled, they have not been investigated together from an HCI perspective using recent, real-world data, which would allow us to develop deeper behavioural insights. Furthermore, we do not know exactly how people are using these services throughout the day over longer periods of time, across different devices and services, and what motivates particular viewing behaviours. The following sections present the results of a diary study with interviews, conducted over a 14-day period with 20 people from nine households. Participants were asked to record the details of each time someone

viewed on-demand content in the household. These diaries focused on when and where viewing took place, as well as which services and devices were being used. Pre- and post-study interviews were also conducted to further probe and understand these present-day viewing practices.

4.2.1 DIARY STUDIES

Diary studies allow researchers to gather data from participants in situ (Rieman, 1993), which can otherwise be difficult for logistical, ethical, and privacy reasons. Diaries are typically filled in by the participants themselves, over a period which often precludes the presence of a researcher due to resource demands. While the self-reporting nature of these studies gives the potential for misreporting or omission (either intentionally or unintentionally) of important details, by shifting the burden of data collection to the participants, such studies can be scalable to a degree that is not possible using other methods while also reducing observer effects (Carter and Mankoff, 2005). Diary studies have been used successfully by HCI researchers previously (e.g. Czerwinski et al. (2004); Carter and Mankoff (2005); Brown et al. (2000); O'Hara et al. (2007); Cecchinato et al. (2016)), and offer a scaleable and convenient way to gain insights into general behaviour over medium-to-long periods.

4.3 METHOD

4.3.1 PARTICIPANTS

Ten UK households who watched at least five hours of on-demand content a week were recruited through word of mouth and advertisements (for a break-

down of households see Table 4.1). One (household C) withdrew, leaving 20 remaining participants from nine households. Mean age was 29.8 (SD = 13.8). Households were paid £100 (~$137) for 14 days of continuous participation.

4.3.2 MATERIALS

Participants were given the choice of a paper diary or a digital diary. Seven households chose the digital diary and two chose the paper one (households A and B). For the digital diary, data was entered using an online form. Using a shareable link, data could be entered using any device with a web browser, allowing participants to use any device they happen to have to hand. Results were stored in a spreadsheet. For the paper diaries, custom diary booklets were created for each household. Once data collection was complete, they were digitised to make them the same format the digital ones for ease of analysis. Both paper and digital versions were designed to make entering data as easy as possible, e.g. with checkboxes for family members' names, viewing locations, etc.

Participants completed information about each viewing session, defined as a period of viewing with at least 30 minutes of non-viewing activity either side to allow for short-to-medium breaks. Participants were required to fill in basic information about their viewing: who was present, start and finish times, what was watched, how long for, devices and services used, location, and breaks they took. They were also asked to justify and explain their responses, where appropriate.

For the purposes of this study, on-demand content is defined as any content that can be accessed at the convenience of the viewer. This includes

Household	Responses	Location	Notes	Participant	Age	Gender	Nationality
A	27	Birmingham	Cohabiting couple	A1	57	F	British
				A2	68	M	British
B	36	Birmingham	Parents and their children	B1	33	M	British
				B2	38	F	British
				B3	8	M	British
				B4	4	F	British
				B5	2	M	British
C	-	-	Withdrew	-	-	-	-
D	22	London	Cohabiting couple	D1	32	M	Spanish
				D2	29	F	Spanish
E	18	London	Cohabiting couple	E1	31	M	Danish
				E2	29	F	Danish
F	24	London	Cohabiting (others not participating)	F1	27	F	Mexican
G	14	London	Cohabiting couple	G1	32	M	Italian
				G2	32	F	Italian
H	15	Oxford	Cohabiting friends	H1	27	F	British
				H2	30	F	British
I	7	London	Cohabiting couple	I1	27	F	German
				I2	35	M	British
J	15	London	Cohabiting couple	J1	31	M	German
				J2	33	M	British

Table 4.1: Participant household profiles

catch up services such as BBC iPlayer and ITV Hub, subscription services such as Netflix and Now TV, short-form content Accessed on video sharing sites including Youtube and Facebook, and content that the participants have downloaded or recorded themselves on their computers or PVR systems (e.g. TiVo).

4.3.3 PROCEDURE

After recruiting participants, a preliminary interview was conducted to ascertain their general on-demand viewing habits and motivations. They were then briefed on how to enter data in their diaries. Participants were requested to create at least one diary entry per day, but this could simply be to say that no viewing took place. For each household, one participant was nominated to be responsible for the diary and complete it on behalf of others if necessary, though other household members were encouraged to fill in the diary as well. During the study, participants were sent SMS reminders every evening to encourage participation. After the study was over another interview was conducted to ask them about their experiences with using the diary, as well as to explain particular behaviours.

4.4 RESULTS

Participants created 202 diary entries in total. Of these, 24 said that no on-demand service usage occurred that day, leaving 178 remaining entries describing on-demand viewing. Mean entries per household was 20.6 (SD = 9.1). These diaries captured 188:36:00 (HH:MM:SS) of viewing time, with a mean of 20:57:20 per household (SD = 08:10:11). This section provides

both a quantitative and qualitative analysis of diary entries. Interviews were also conducted with participants at the end of the study to learn more about their diary entries and on-demand viewing habits. These interviews were transcribed and were analysed thematically using an inductive coding approach.

This section presents data from both the diary entries and direct quotations from the thematic analysis of our interview data. This is clustered around eight different section headings. First, *viewing screens*, is where the devices people chose for viewing and why are considered. Second, *viewing location*, considers the places both inside and outside of the home where people chose to view. Third, *viewing time of day and duration*, considers how viewing fits into people's daily activities and how long they view for. Fourth, *services used*, considers exactly which on-demand services people used to access content. Fifth, *watching alone and watching together*, considers participants' co-viewing habits. Sixth and seventh, *positive perceptions of on-demand viewing* and *negative perceptions of on-demand viewing*, explores what people like and dislike about these platforms. Finally, *binge watching*, focuses on how on-demand services can facilitate viewing a lot of content in one session, and how people think about and define binge watching.

4.4.1 VIEWING SCREENS

Firstly, focus is given to the kind of screen that participants used to view content. Diary entries fell into five distinct viewing device categories, pre-specified in the participants' diaries. These are shown in Figure 4.1. These were further collapsed into two distinct groups: non-mobile devices (TV, desktop computer, laptop computer) and handheld mobile devices (phone and tablet). Of the 178 entries, 55 (29.9%) contained viewing on a handheld mo-

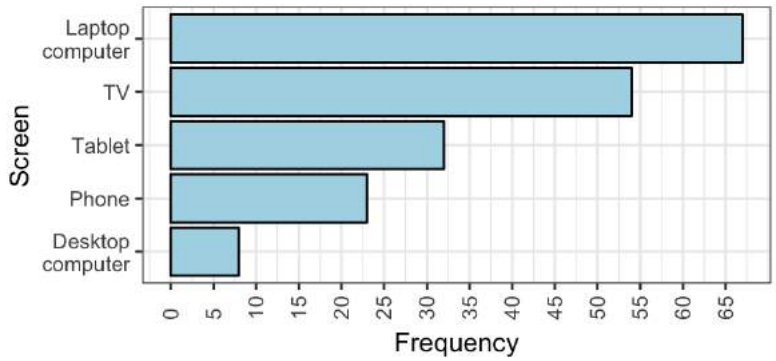


Figure 4.1: Distribution of viewing screens

bile device (i.e., phone or tablet). Households reported more viewing sessions on non-mobile devices (M = 14.0, SD = 8.3) than on handheld mobile devices (M = 5.8, SD = 6.7).

To further understand how people chose a viewing device, their entries and what was said during the end of study interviews were looked at. It became clear that different viewing devices were chosen for different reasons. For example, participant A1 described how she and A2 (her partner) would choose their tablets when they wanted to watch content individually, while still being together in the same room.

> *A1: [We watch] the stuff on the tablet singly — we both watch different things on that — but on the TV we tend to put something on that we both want to watch.*

This was later clarified:

> *A1: I can watch what I want to watch. We both put our earphones on and we can then watch our own watching[...]. The TV, that's*

our bit of relaxation together. But our little bit of YouTube is what we do on our own.

Portability was another factor. Participant B1, an eight-year-old child, said he liked to be able to watch anywhere, instantly:

Interviewer: Why do like to watch it on a tablet?

B3: Because I can take it anywhere. TV, [...] you have to leave it there. And [other devices] take loads of time to set up if you take it somewhere.

Participant J1 said that the device could dictate the content that was viewed, with phone viewing only being for short clips:

J1: Most of the time the phone is usually for only shorter snippets it's like YouTube, or Twitter things... like really short up, to five minutes or so. [...] if I'm taking the time to watching something for longer, I can also take the time to just sit on the couch and relax.

Participant F1 said the phone was her preferred device in many cases, also due to the immediacy of it:

Interviewer: You seem to watch on your phone quite a lot. Is that your preferred device?

F1: Yeah, I mean that's when I'm at home. I think when I'm [at work] I use my laptop.

Interviewer: So what do you like about the phone for watching stuff on?

F1: That it's just more immediate.

However, mobile viewing was consistently seen as being unfavourable and was often avoided if possible. This was typically due to small screen sizes, as stated by household I:

Interviewer: So do you ever watch on tablets or phones?

I2: No.

Interviewer: Never? Absolutely never?

I2: Never.

Interviewer: Okay and why is that?

I1: Screen is too small.

Participant B1 spoke in disbelief that someone could watch for long periods on a small screen:

B1: I was talking to [my friend] about this earlier and he said every night he'll sit and watch a film on his phone. He'll sit there, like, next to [his wife] and she'll sit there watching something he's not interested in and he'll sit there and watch a film or watch videos on YouTube, something to do with work, whatever. And his phone is the same size as mine. I couldn't imagine watching a whole film, just because it's too small.

When asked further about mobile viewing, he clarified:

B1: I don't really get much pleasure holding the tablet to watch something. [...] It doesn't interest me, I'd rather sit and watch it on the telly or not bother. [...] One, you've got to hold it and two, the size of the screen.

However, he did see a benefit to mobile viewing in keeping children occupied:

B1: What I would say about the tablet and the phones, though, is having the kids, when you're out and, say you're going for a meal or something like that, having the phone or tablet with video or like you say, YouTube, is really quite handy because it does keep them occupied.

While the laptop was the most popular viewing screen, participants consistently said that they would prefer to watch on a television. One of the main reasons for this was the bigger screen, but participants also liked the associated comfortable seating. Household A said how watching on the TV was just part of their routine:

Interviewer: Why do you watch, for instance, Better Call Saul on the TV?

A2: Bigger screen.

A1: Bigger screen, yeah. [...] And it's our sort of evening routine, we come in [the living room], we sit down and we watch TV and that's... yeah, it's our routine really.

Participant D1 said watching on a TV is the ideal situation, even though he did not own one himself:

Interviewer: In an ideal world what would you choose to watch on?

D1: A really cool and expensive and nice TV.

Interviewer: And why is that?

D1: Because the quality is quite nice, and if everything is integrated with the streaming service and all that then... lying down on the sofa is the best option.

4.4.2 VIEWING LOCATIONS

The location where participants viewed content is next considered. As shown in Figure 4.2, viewing occurred in 10 distinct locations, with the living room and bedroom being the most popular locations. These locations can again be further collapsed into two distinct groups: watching in the home (living room, bedroom, kitchen, etc.) and watching outside of the home (workplace, public transport, public place). Households reported more viewing sessions inside the home (M = 17.7, SD = 8.9) than outside of the home (M = 2.1, SD = 2.2), and in total 169 viewing sessions (89.9%) took place inside the home. Moreover, four of the nine households never once reported watching content outside of the home. Most viewing sessions were reported to have taken place in a single location; there were just five diary entries (2.8%) in which participants reported moving between two locations, and all of these were entirely inside the home.

Diary entries and interview data suggest that viewing location was often not a concious choice, but a result of situational and contextual factors. Participant F1, living in a shared house in London (where it is common to convert

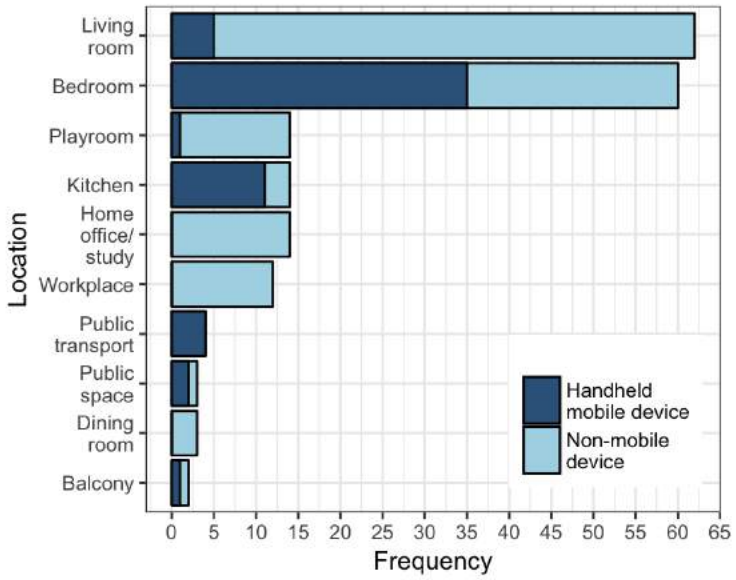


Figure 4.2: Distribution of viewing locations

communal living spaces into extra bedrooms), spoke about how she could not watch in the living room:

> *Interviewer: Why do you prefer to watching the bedroom than in the living room for instance?*
>
> *F1: Because I don't have a living room.*

While small screens on mobile devices were often seen negatively, some participants spoke favourably about being able to view on public transport due to their portability, such as H1:

> *H1: You can use it on a plane.*
>
> *Interviewer: Why is that?*
>
> *H1: Because you can just put it on the little table.*
>
> *Interviewer: Because it's smaller?*
>
> *H1: It's smaller.*

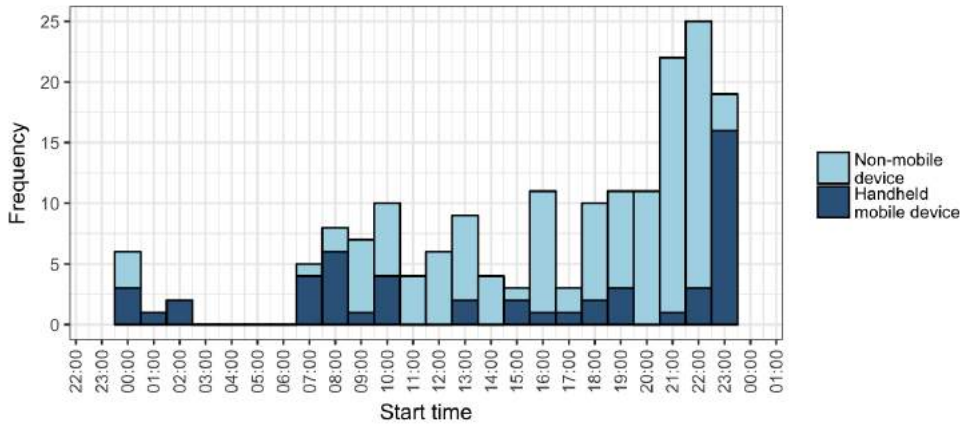


Figure 4.3: Histogram of viewing start times

4.4.3 VIEWING TIME OF DAY AND DURATION

Next, the times at which participants watched during the day, and how long their viewing sessions lasted, are considered. A histogram of viewing start times can be seen in Figure 4.3. It can be seen in the figure that late evening tended to be the most popular time to start viewing, though lower levels of viewing also took place throughout the day, apart from in the very early hours of the morning. In terms of total viewing time, 105:08:00 (55.7%) of viewing took place in the evening period between 18:00 and 00:00. It can also be seen in the figure that viewing on handheld mobile devices was particularly popular in the morning, and during late night and the early hours of the morning. There appears to be a noticeable transition from the pre-bed social ritual of watching on a TV to personal viewing on mobile devices at bedtime.

When considering how long participants viewed for, it was found that mean viewing session duration was 01:03:00 (SD = 00:55:56). A histogram of session durations can be seen in Figure 4.4. Of all the sessions, 122 were one hour or less (69%), and 158 sessions (89%) were two hours or less. Figure 4.5 shows a detailed view of these sessions, where the most common

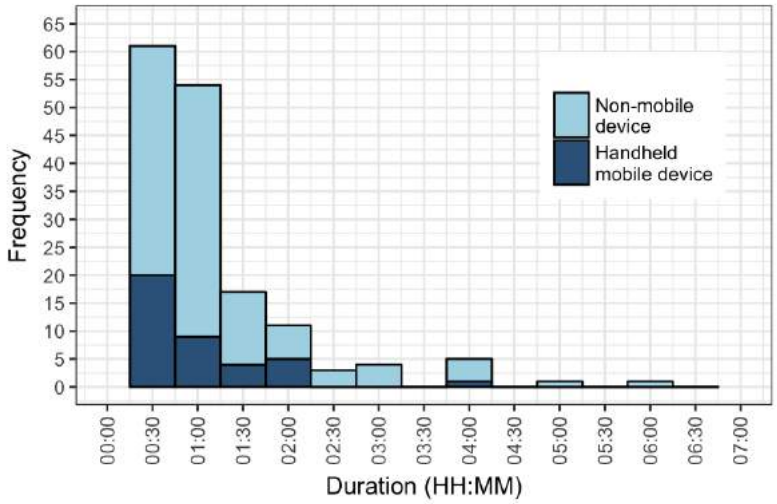


Figure 4.4: Histogram of viewing session durations

durations is 30 minutes (often the length of one episode). Only 22 (12%) viewing sessions were over two hours. The longest session was six hours, and the shortest two minutes. On average, households reported longer viewing sessions on non-mobile devices (M = 01:15:37, SD = 00:34:42) than on handheld mobile devices (M = 00:38:24, SD = 00:18:58).

4.4.4 Amount of Content Viewed

To better understand what was being watched during a session, the amount of content that was watched is also considered. For this analysis, each episode or separate video is considered to be a different item that is watched. Participants reported watching 481 items across 178 sessions; watching 2.7 items per session (SD = 2.7, range: 1–20). The largest number of items viewed in a single session was 20 YouTube videos over 90 minutes. It was found that households tended to watch more items on non-mobile devices (M = 38.1, SD = 33.1) than on handheld mobile devices (M = 15.3, SD = 23.4).

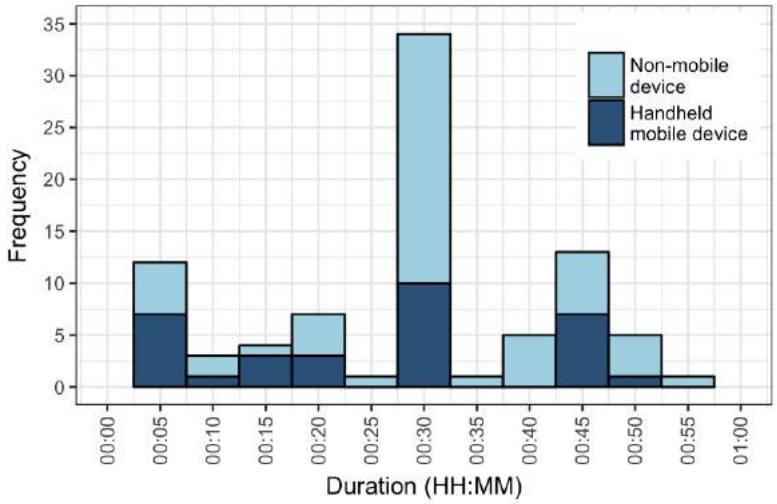


Figure 4.5: Histogram of viewing session durations for sessions with a duration of one hour or less

4.4.5 On-Demand Services Used

Participants were also asked to record which on-demand services they used for viewing. They reported using 13 distinct services. These are shown in Figure 4.6, along with the number of sessions they featured in. These services were divided into two categories: short-form, which consisted of YouTube, Facebook, Lynda iOS app (a training course app), Vimeo, WhatsApp, and The Guardian website (news); and long-form, which consisted of Netflix, Raiplay (Italian on-demand service), BBC iPlayer, unofficial streaming services, home recordings, and Amazon Video. Households reported more sessions featuring only long-form services (M = 10.6 SD = 7.3) than sessions featuring only short-form services (M = 8.6, SD = 6.5).When considering viewing time, households reported longer viewing sessions when sessions featured only long-form services (M = 01:21:49, SD = 01:04:40) than with sessions containing only short-form services (M = 00:38:13, SD = 00:30:40.

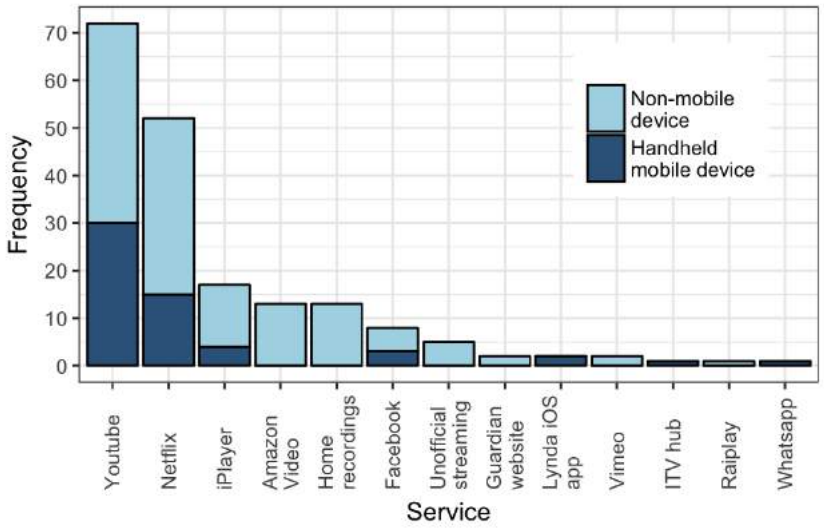


Figure 4.6: Popularity of different services

4.4.6 WATCHING ALONE AND WATCHING TOGETHER

Whether people watched alone or with others (i.e., co-viewing) was also explored. Watching alone was more common than co-viewing. In total, 135 sessions (75.8%) were watched alone, and 43 (24.2%) by multiple people. In this context, co-viewing refers to more than one person actively watching.

Motivations for watching alone were explored during the interviews. A common theme was differing interests. For example, household E (a cohabiting couple with 88.9% of their sessions viewed alone), had very different tastes:

> *Interviewer: So what affects whether you watch together? Is it that you like different things, is it that you're just around at different times?*
>
> *E1: Yeah I think I like watching it more than [E2] does, and different things. I really enjoy watching House of Cards whereas he's more, I guess, logical with what he chooses to watch.*

Household D (53.3% of sessions viewed alone) also cited similar differences in personal tastes:

> *Interviewer: And you said you don't ever watch things together?*
>
> *D1: Not really. [My girlfriend] falls asleep all the time.*
>
> *Interviewer: [...] apart from that is there another reason? Do you like different things?*
>
> *D1: We do really like very different things, and I think the rare occasion we watch something together is... Well actually, we do watch quite often The Big Bang Theory during dinner. But it's fifteen minutes and we watch the same episodes all the time. So it's more as kind of a background thing... We tried watching other TV series that we might enjoy watching together but those haven't existed to this point.*
>
> *Interviewer: Okay, so you said she goes to sleep all the time. Is it because she sleeps early and you go to bed late? Is that a factor?*
>
> *D1: No no, it's more that she only wants to watch what she likes, and if I don't adjust to it she really finds it really boring and just falls asleep.*

Household A had a more even split of watching alone (58.33%) versus together (41.67%), but still expressed different tastes which influenced whether or not they viewed together:

> *Interviewer: So, why don't you want to watch what [A2] wants to watch?*
>
> *A1: Because it's food programmes...*

A2: Health...

A1: (Laughing) I can answer it myself. Yeah, it's generally food programmes, health programmes...

Interviewer: Okay. And [A2], why don't you want to watch what [A1] wants to watch?

A2: I can't watch another camper van conversion [on YouTube]!

Interviewer: (Laughing) Okay.

A2: And [A1]'s tutorials, he watches an awful lot of tutorials, which wouldn't interest me.

However, often watching alone was driven by situational factors rather than conscious choice — sometimes people just happened to be alone when they watched. Participant F1 (90.91% of sessions watched alone), an international student, discussed how she often watches alone when in the UK, but with family when back in her home country:

Interviewer: Do you normally watch alone then, when watching on-demand stuff?

F1: Yeah well, when I'm here, yes. ... If I go for holidays back home then I might do it with my, I don't know, with my sister, or my mom.

Participant G1 discussed how watching at work for a break meant they often watched alone:

G1: ...we like some similar shows and therefore we watch them together, but also because, I mean even for instance, [...] during

lunch break if I'm alone, I watch something and therefore I am alone!

Participant H1 also referenced her living situation, having recently moved into a different household with new housemates:

H1: But maybe that's just because of my living circumstances. Before I used to just watch TV with other people.

Interviewer: So which would you prefer? Or does it depend?

H1: It depends, but I prefer to watch stuff with other people I think.

4.4.7 Positive Perceptions and Motivations for Using On-Demand Video Services

Participants generally had favourable opinions about on-demand services. One of the most obvious themes from the data was the benefits provided by these services in terms of freedom, convenience, and choice for viewers, which was a strong motivator for using them. The results presented above show this clearly — participants watched in a variety of locations, at different times and on different devices. They also spoke about this in interviews, such as the following quote from participant A1:

A1: You can choose when you watch it then can't you? You know you don't have to say "ooh it's on at 8 o'clock tonight, we've got to be there for 8 o'clock". If we watch it on-demand you can think "I'll watch it at 10 o'clock if I want".

A number of households spoke about the catalogue of content available, which can make it easy to find something to watch:

> *E1: I think I watched the first thirty minutes of it but didn't really... And that's the thing about on-demand — if you don't really like it, you can just find something else.*

> *F1: It's been so long since I watched the TV that I don't even remember how it is that you have to wait every week for a new episode or whatever, for the series, right? So now I just watch them whenever I want, whenever I have the time.*

> *J1: I can always find something [more easily] on on-demand, because on broadcast TV I am limited to [...] forty different channels? And most of it is just reruns, and on-demand I have the selection of fifteen-thousand videos or something.*

Participants also remarked on the quality of content available via on-demand services in comparison with broadcast TV, which was generally seen as similar if not better:

> *D1: I think I went into Netflix because of the catalogue they had, and a couple of their own productions like House of Cards, Orange Is the New Black, and the fact [that] they had a lot of stuff [...]. I tried it out and I enjoyed it.*

However, participant J1 did note how broadcast TV was still useful for time-sensitive content:

J1: I think that quality is quite similar, I think. Broadcast TV also has some more things to offer like as news coverage which you don't get on demand.

4.4.8 Negative Perceptions of On-Demand Viewing

While opinions were generally positive, participants noted some negative aspects. Typically these were in relation to watching large volumes of content, or at least the potential to. Some participants spoke about being addicted to particular shows. Household A spoke about AMC's *Better Call Saul*:

A1: Yeah, well we like to watch two or three at a time, don't we?

A2: Yeah.

Interviewer: Why is that?

A1: We can't stop watching them because they're addictive.

A2: The trouble is, when you watch one that's on for almost an hour... you just, you feel as though you've been short-changed, you need to watch some more.

A1: Yeah, you're drawn in aren't you? You just want to watch more.

Participant F1 also spoke about a compulsion to watch:

F1: It becomes a bit addictive now. At least on the TV if you missed an episode you will be like "Oh okay, I'll just watch it next week" and then you will do other stuff. But now, I have this theory about habits. Because I can watch whenever I want, and

it's the sort of thing where I need to be watching now to go to sleep.

Other participants also spoke about becoming hooked on a show, and how certain services made it very easy to watch another episode:

E2: Normally when you watch something you have to say "Should we see one more?". Then we would actually take an active choice to press next button, but Netflix there's like five seconds count-down. So often we [think] "Should we see one more?", "Hmm, I don't really know" and then, the intro screen is on and Netflix started.

E1: It made the choice for us.

E2: Yeah, I think if it didn't start automatically and we actually had to push the button, then I think we would talk. I probably would talk about if we should see one more, because now it was the fourth in a row.

Participants also spoke about trying to control their viewing to ensure they didn't spend more time watching than intended:

E1: You also want to see a lot of these like, Suits, or whatever... We don't want to get dragged into it because I can't get out of it (laughing). [...] I know myself too well that I'll end up spending half a day there.

Participant H1 said something similar:

H1: I didn't want to start a series one time because I knew I would just waste so much of my time watching it.

This type of boundary setting was also mentioned by other participants. Participant F1 thought that excessive viewing might be having a detrimental effect on other areas of her life, and so spoke about creating hardware boundaries to combat it:

F1: No I don't have Netflix on my phone, and I don't want to put Netflix on my phone.

Interviewer: Why is that?

F1: Because at least with the tablet you know I leave it at home, and I know that I won't use it unless I am at home at night.

Interviewer: So that's one way of setting a boundary?

F1: Yes, I mean I always feel to set these boundaries and they work, but the problem is for the last few months I've become an addict to YouTube. I don't think I was like that last year. I was a bit more able to control myself.

Interviewer: How about watching outside of the office and outside of home? So, maybe in a public place or while you were travelling for instance. Can you talk about if you did any of that?

F1: No, because so first of all I don't have enough data to watch videos, and I also deliberately don't pay more [...] so that I can restrict myself from watching videos, because otherwise I would just be watching everything.

Participants also said how watching too much content often meant they wasted time or ended up going to bed too late:

> *A2: I think sometimes we normally stay up a bit late with on-demand.*

> *H1: Um, well, because then I'll watch maybe three episodes in an evening...well, on a bad evening or like I'll watch two and the next evening I'll watch two. If there are lots of episodes in the series then that's a big waste of time.*

4.4.9 BINGE WATCHING

Discussions of consuming too much content often brought the subject specifically to binge watching, which was discussed with all of the participants. Most of the participants were familiar with this behaviour and said they participated in it themselves. It seemed that this phenomenon could be thought of as a particular type of excessive viewing. However, when pushed to define binge watching, few participants had a clear idea of what binge watching was. Some would define it as being based on the number of episodes of a show that was watched, e.g. participant G1 defined it as three or more episodes, but only when watching TV shows:

> *G1: I have always thought about it in terms of TV shows [...]. So, watching many more than just one single episode, in one sitting.*
>
> *Interviewer: So how many episodes is it before you are binge-watching?*
>
> *G1: I would say from three.*

Interviewer: So if I watch three five minute YouTube videos, is that binge-watching?

G1: Not exactly. My understanding was [...] that you are watching episodes of 45 minutes each.

Participant J1 also agreed with this:

J1: I think binge watching should be sort of a TV series episode length. An hour, or 45 minutes, or 42 minutes... and you watch more than two of those in a row.

Others said it was based on the amount of overall time spent, such as participant D1:

Interviewer: How many episodes do you think is binge watching?

D1: Ooh, erm, anything that goes above four or five hours.

Interviewer: Okay, so it's more about the time than the number of episodes for you?

D1: Yeah, because it's not the same to watch a whole TV miniseries that has 10 episodes [that are] an hour and 15 minutes each, [as it is to watch] ten episodes of The Big Bang Theory or Friends.

Participant I2 also agreed with this, specifically noting how the number of episodes was inconsequential. He also seems to think that it is possible binge watch shorter content:

I2: Well, the [time and number of episodes] are synonymous, right? So if the programme was 10 minutes per episode, then I would go through [many] more episodes probably to achieve the same amount of time.

Household B also thought it was based on the amount of time spent, but disagreed about the actual definition.

Interviewer: So how would you define it? Is it the number of episodes or is it the amount of time that you watch?

B2: The amount of time. [...]

Interviewer: So, how many episodes would have to watch and how long would you have to watch for, for it to be binge watching?

B3: I don't know. I guess if you sit there, waste your whole night. [...]

B3: Yeah, I've never thought about it before, so I don't know. Um, four or five hours I guess. [...]

B2: I'd go for three.

Such disagreement as to what constitutes a televisual binge was also present in other households, such as in household A:

A1: Didn't we watch three [episodes] in the last couple of weeks? We watched three [episodes].

A2: Oh three... yeah, but I wouldn't say that constituted binge watching, but maybe it does.

A1: I think three is, yeah. Three is, I would say, yeah.

A2: I don't know... but yeah we did watch three.

Interviewer: So would you say that that's binge watching?

A2: I wouldn't say that's binge watching.

Interviewer: Why not?

A2: I don't think there are enough episodes there.

Interviewer: Okay.

A1: I would say... I think more than two is binge watching.

Unlike some, H1 did not think that episodes necessarily had to be watched back-to-back or even on the same day:

H1: I think it's watching multiple episodes compulsively. [...] it could be one episode but you watch an episode per evening or it could be within a shorter space of time...

Participant J1 also suggested that watching one episode per evening could be binge watching, but was not entirely sure:

J1: Maybe it is... maybe seven episodes in seven days is binge watching [...] it's difficult to say. I think... like in a short period of time, watching something that was made for once a week maybe.

Participant I1 thought binge watching was more related to viewing intentions:

I1: It depends, yes, because if I'm supposed to work and I tell myself, "Okay, one. One video," and then I end up watching six, then it's kind of binge watching too, because I was supposed to take just a 10 minute break.

Interviewer: So does it depend on what you are supposed to be doing, for you?

I1: Yeah, I guess what the intention was. If I really just want a fifteen-minute break and I end up, you know, watching something for thirty minutes, then I kind of escalated there, so in a way that would be binge watching. If it's a lazy Saturday afternoon and it's raining and I end up watching three or four episodes, then yeah I think four or five is turning into binge watching, but otherwise if I have the time and nothing else to do...

4.5 DISCUSSION

The findings of this study show that although on-demand video platforms have the potential to change viewing behaviour, viewers still often conform to traditional viewing habits. For instance, in terms of viewing time, most viewing occurred during the evening "prime time" slot. Furthermore, the most common session duration was 30 minutes, typically the length of one episode of content. The TV was also still a popular viewing screen. However, changes as a result of new technology can also be seen. YouTube was the most common viewing platform, showing how shortform content has become popular. It was also found that a third of viewing happened on a mobile device, and instances of very long viewing sessions.

When considering viewing screens, the laptop was slightly more popular as a viewing device than the television, which may not be possible without the cross-device availability of on-demand services. This could be due to the ease of access to different services via the internet, as well as the balance of screen size and portability that laptops provide. However, for the purposes of viewing they function similarly to a TV — a fairly large screen that can be placed in a comfortable location, with the ability to watch with others. Though most viewing occurred on larger screens, a third of viewing sessions were on handheld mobile devices. This was generally seen as unfavourable, and mostly seemed to be down to necessity — in interviews, participants expressed their dislike for viewing on mobile devices, citing the small screen as a reason. However, participants said that they would watch on a mobile device if no other device were available (e.g. when travelling). Most said they preferred to watch on a TV, due to large screen size and comfortable seating typically found nearby. Individual differences were evident however, with some participants entirely discounting watching content on phones, and others sometimes preferring it.

It was found that viewing device often depended on people's locations. The majority of viewing (89.9%) took place inside the home, and the living room was the most popular location. It is perhaps then not surprising that people tended to watch on larger display TVs and laptops rather than smaller mobile screens when in the living room. Mobile devices tended to be used in the home in locations where there may not be access to a TV, such as the kitchen or bedroom. The participants did report watching on mobile devices when outside of the home, particularly when travelling and commuting to work. In recent years the lower cost of mobile data has made easier to watch

on-demand services on the move. These instances of mobile viewing tended to be during longer journeys, possibly because it allows for an entire episode of content to be watched.

Participants spoke very favourably about on demand services, especially about how they have allowed them more freedom and choice than broadcast TV. However, a study by Vanattenhoven and Geerts (2015) found that some consumers found the amount of choice available to be an annoyance, especially with regard to the number of different services available. Interestingly, this sentiment was not found in any of the data from the present study.

This freedom also allowed participants to select content that matched their personal tastes. Interview data from the present study revealed that these differing tastes among household members could lead to people choosing to view alone, which was reflected in the diary data showing that 75.8% of sessions were watched alone. This shows a stark turnaround of events when compared with an observational study by Saxbe et al. (2011), who found that watching TV with at least one other person happened for 61% of the time, and that the TV provided a platform for togetherness in the household. It also contrasts with the observational study in Chapter 3, where co-viewing was found to be common when watching on a living room TV. While participant D1 said he and his partner generally watched different content in different rooms due to differing tastes, participant A1 described how she and her partner used tablets and earphones to watch different content, but still be in the same room together. This agrees with Ofcom's findings (Ofcom, 2017) who found that people often turn to on-demand services for some "alone time".

Although it was observed that only a quarter of sessions were co-viewed, previous research shows that viewers value the way new broadcasting tech-

nologies can enhance the social aspect of viewing (Barkhuus and Brown, 2009; Irani et al., 2010). While watching alone was more common than co-viewing, it may be that the sessions watched alone were driven by other latent social factors, such as being able to discuss the show with friends. Finally, it could be that co-viewing and other social factors work differently in different household configurations, e.g. it was observed that the households with the highest percentage of co-viewing were household A, an older couple (42.1% of sessions co-viewed) and Household B, a family (41.7% of sessions co-viewed), while the household with the lowest percentage was household H, two young professional cohabiting friends (6.7% of sessions co-viewed). This cannot be speculated upon beyond this with the data from the present study, but it would be an interesting focus of future research.

Participants were often wary of the way instant access to large amounts of content could mean watching for long periods. This led to some participants creating boundaries to prevent this behaviour, either by simply not starting to watch a new show, or by restricting viewing in some other way, e.g. not installing Netflix on their phone. While it may be in the interests of service providers to make it as easy as possible to view large volumes of content for revenue and engagement purposes, this was often troubling to our participants, some of whom commented that Netflix "made the choice for [them]" when deciding whether to watch another episode. As such, the introduction of small "design frictions" to combat automatic behaviours could lead to a better user experience (Cox et al., 2016), either by design or manually by the users themselves.

Discussing consuming large amounts of content typically led to talking about binge watching, which most of the participants said they participated

in. However, when pushed to define what binge watching was, there were widely different responses and definitions often seemed to change depending on the context. This is reflected in other studies, where binge watching is defined differently by different authors. For example, some participants said it was watching two or more episodes in a row (as in Pittman and Sheehan (2015); Ofcom (2017)), and said three or more episodes in a row (as in de Feijter et al. (2016); Walton-Pattison et al. (2016)). Others said it was not so much the number of episodes watched but the total time spent watching, while others said it was a combination of these two features. Others said that it depended on their intentions when they started to watch. In summary, different people seemed to have different ideas of what binge watching is, and this disagreement reflects the diversity of definitions that appear in the literature on this topic. Such varying definitions suggest that it could defined on a scale, and vary with context and type of content, as suggested by Trouleau et al. (2016).

While effort was taken to recruit participants of various ages and living in different parts of the UK, most of the participants were London-based millennials without children. This bias in the sample may have affected our results. For instance, some participants lived in shared housing without a communal living room or TV. In place of this, viewing occurred on laptops and tablets in bedrooms. Considering millennials' typically high level of interaction with technology, more activity that differs from traditional notions of TV viewing might have expected. This may have seen an increase if our sample featured more teenagers and children. Viewing mainly in the evening is perhaps to be expected, as our sample was mostly adults in full-time employment. However, there was a steady amount of daytime viewing, resulting from one household

with children being at home and people viewing during work breaks.

The sample for this study consisted of 20 individuals from nine households. This could be argued to be a small sample size, however it is similar to that of comparable studies (e.g. O'Hara et al. (2007); Barkhuus and Brown (2009); Vanattenhoven and Geerts (2015); McNally and Harrington (2017)). It also reflects the challenges of conducting this type of research, where prolonged studies with involved tasks for participants can deter participation, even when well compensated. Notwithstanding, as the present study is qualitative in nature, it is argued that the sample size is sufficient to illuminate many of the behaviours surrounding on-demand and mobile viewing, especially given the study duration.

A limitation of the diary study method is that some participants may not have recorded everything they watched. During interviews some participants did remark that they sometimes did not record very short viewing sessions (e.g., a short Facebook video) because of the effort involved. However, this was fairly uncommon, with most participants saying they recorded the vast majority of content they watched.

4.6 CONCLUSION

The work presented in this chapter extends our understanding of how on-demand viewing occurs in daily life. The results of a diary study show that this technology leads to new behaviours such as mobile viewing, viewing for long periods, and consuming shortform content. However, the sample in this study still often conformed to traditional viewing habits. Viewing was mostly in the evening on a large screen, though this sometimes happened in new

ways, such as by using a laptop. While mobile viewing did account for a third of all viewing sessions, in general this was seen as less favourable than watching on a large screen. Typically, mobile viewing seemed to occur for contextual reasons, such as being a practical device to use while travelling, or wanting to watch content privately when in the presence of others. It was also found that viewing alone was far more common than viewing with other people. Participants had largely positive opinions about on-demand video services, but generally seemed to be wary about the ability watch for long periods and the impact it could have on other areas of their lives.

The work in this chapter and Chapter 3 have taken an in-depth look at contemporary viewing practices in people's everyday lives. In order to understand some of these specific behaviours in more detail, the following chapter details the development of a questionnaire to measure the effect these behaviours have on viewer experience.

CHAPTER 5

DEVELOPMENT OF THE IMMERSIVE EXPERIENCE QUESTIONNAIRE FOR FILM AND TELEVISION

5.1 INTRODUCTION

Measuring engagement and experience when watching TV in films allows researchers to assess the effect of new habits and technological interventions. The studies in Chapter 3 and Chapter 4 showed how the participants were using technology as part of their viewing, e.g. frequently using their mobile devices, and watching on demand content. These emergent behaviours offer the potential to both benefit and reduce a person's viewing experience, but exactly how viewing experience is affected is not well understood or even necessarily measurable. For instance, while it may seem intuitive to argue that constantly checking one's phone leads to reduced engagement with the content, it is difficult to empirically measure this. Previous measures of engagement and experience have used physiological measurement, which can require specialist equipment such as heart rate monitors and skin conductance measurement devices (e.g. Reeves et al. (1999); Lombard et al. (2000)). Another way is through the use of self reported questionnaires, which often do not measure experiences in standardised ways and often only measure partial elements of media experiences (e.g. presence questionnaires (Witmer and Singer, 1998)).

This chapter details the development of the Film IEQ, a questionnaire designed to measure immersion when watching video media. Through this, we can develop a more detailed understanding about how specific behaviours, practices, and interventions are affecting viewer experiences, and in a standardised way. This is intended to be an easily deployable tool to measure a more holistic definition of experience than other questionnaires, such as presence.

5.2 Related Work

5.2.1 Measuring Viewing Experiences

Presence is a common metric used to measure experience when consuming both digital media (e.g. computer games, virtual reality) and traditional media (e.g. books and film), and refers to the viewer feeling like they are having a non-mediated experience (Lombard and Ditton, 1997). Presence has its roots in virtual environments literature (Sheridan, 1992), and questionnaires have been developed to measure it (Witmer and Singer, 1998; Lessiter et al., 2001).

Presence has also been investigated when viewing video content, using both questionnaires and physiological measures — see Section 2.4 for a review of some relevant literature.

Though presence has been used as a measure of experience, it typically refers to spatial presence (though it can also refer to social presence (Lombard and Ditton, 1997)), or feelings of being physically located in somewhere other than the real world (Schubert et al., 2001). However, the focus on the user feeling physically transported to another place does not offer a holistic view

of experience. Consider someone watching a TV quiz show, where it could be argued that they can have positive viewing experience without feeling that they have been transported to the TV studio. It is due to these limitations that this work looks to other literature where experience measurement has been studied more broadly, and specifically that of computer games.

5.2.2 Experience Measurement in Other Domains

Experience measurement has been widely studied for computer games, operationalised through a number of concepts (see Section 2.4 for a review of some relevant literature). Immersion is seen as a highly desirable quality in computer games, and has itself been widely researched. However, there are differing definitions, and care should be taken to differentiate between these to ensure consistency when attempting to compare effects (Cairns et al., 2014a). While some metrics are quite narrowly defined, immersion has been defined as a generalised concept which measures a number of aspects of experience by Jennett et al. (2008), who developed the Immersive Experience Questionnaire (IEQ) to measure it. This was developed from the related areas of flow, cognitive absorption, and presence, but they specifically highlight to how immersion is distinct (e.g. a player can be immersed in a game of Tetris without feeling like they are physically present in a world of falling blocks). This instrument has been widely used in games research (Sanders and Cairns, 2010; Cox et al., 2012; Thompson et al., 2012; Cairns et al., 2013, 2014b), and also successfully adapted to other domains such as public speaking anxiety (Wortwein et al., 2015) and games without graphics for visually-impaired players (Engström et al., 2015).

Brumby et al. (2014) and du Toit (2013) used a modified version of the IEQ in order to measure video engagement, as part of a lab study to assess the effects of watching video while attempting to do other tasks. They found that watching the video did not make tasks less stressful. Furthermore, in the some cases the tasks distracted users from the television, which lead to lower immersion and engagement. However, this work does not detail exactly how the questionnaire was developed, leaving the reader unable to assess the suitability and validity of the measure.

Due to its broad insight into experience, the robustness of its development, and its wide usage, the IEQ presents a promising way of measuring immersion for film media. However, there are important differences between playing games and watching video media. Firstly, watching video media is typically a "lean back" activity, where the viewer simply observes and does not interact. On the other hand, playing games is a "lean forward" activity, where the player is constantly interacting the with game. Secondly, Jennett et al.'s definition of immersion incorporates flow, which is concerned with the extent to which a user's ability is matched to the task at hand. While there is no real task to be completed when watching TV and film, some researchers argue that states of flow can apply to non-participatory media (see Section 2.4.3).

Even though there are differences between playing games and watching video, the strong theoretical background of the IEQ provides a good basis for usage in non-game domains. It measures experience in mediated environments, and many of the questions contained within are general enough to apply to media outside of games. Taking this into account, the aim of this chapter is to develop a modified version of the IEQ to measure immersion in video media. An exploratory factor analysis was also performed to establish

the underlying factor structure of the questionnaire.

5.3 QUESTIONNAIRE DEVELOPMENT

In order to develop a questionnaire to measure immersion in video media, the original IEQ was used a basis. Firstly, the questions were reviewed to find any wording that was specific to games. In a similar manner to du Toit (2013), these were then reworded to be specific to film and television, providing that the essence of the question remained the same, e.g. “to what extent did the game hold your attention?" became “to what extent did the movie, TV show, or clip hold your attention?".

Some questions intuitively do not apply to the mostly passive experience of watching video content, and were unable to be reworded or modified. These were replaced with questions concerning how well the viewer had followed the content and themes of the video content ("how challenging were the themes?" instead of "how challenging was the game?"), as well as narrative engagement (Busselle and Bilandzic, 2009) in a similar way to Brumby et al. (2014). These were inspired by Busselle and Bilandzic’s concept of narrative engagement which incorporates elements of presence, flow, transportation (Green and Brock, 2000), and cognitive and emotional investment, in a similar way to game immersion but in a non-interactive context.

After completing the process of modifying the original IEQ, the resulting questionnaire consisted of 31 items measured using a 1–7 Likert scales (see Table 5.1). The questionnaire was piloted to ensure the it made sense and the wording was clear. Non-native English speakers were also asked to complete the questionnaire to make sure the language used would be widely understood.

#	Original question	Modified question
1.	To what extent did the game hold your attention?	To what extent did the **movie, TV show, or clip** hold your attention?
2.	To what extent did you feel you were focused on the game?	To what extent did you feel you were focused on the **movie, TV show, or clip**?
3.	How much effort did you put into playing the game?	How much effort did you put into **watching the movie, TV show, or clip**?
4.	Did you feel that you were trying you best?	Did you feel that you were trying you best **to follow the events of the movie, TV show, or clip**?
5.	To what extent did you lose track of time?	*Unchanged*
6.	To what extent did you feel consciously aware of being in the real world whilst playing?	To what extent did you feel consciously aware of being in the real world whilst **watching**?
7.	To what extent did you forget about your everyday concerns?	*Unchanged*
8.	To what extent were you aware of yourself in your surroundings?	*Unchanged*
9.	To what extent did you notice events taking place around you?	*Unchanged*
10.	Did you feel the urge at any point to stop playing and see what was happening around you?	Did you feel the urge at any point to stop **watching** and see what was happening around you?
11.	To what extent did you feel that you were interacting with the game environment?	**To what extent could you picture yourself in the scene of the events shown in the movie, TV show, or clip?***
12.	To what extent did you feel as though you were separated from your real-world environment?	*Unchanged*
13.	To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?	To what extent did you feel that the **movie, TV show, or clip** was something you were experiencing, rather than something you were just **watching**?
14.	To what extent was your sense of being in the game environment stronger than your sense of being in the real world?	To what extent was your sense of being in the environment **shown in the movie, TV show, or clip** stronger than your sense of being in the real world?
15.	At any point did you find yourself become so involved that you were unaware you were even using controls?	**While watching the movie, TV show, or clip, could you easily picture the events in it taking place?***
16.	To what extent did you feel as though you were moving through the game according to you own will?	**To what extent did you find yourself thinking of ways the story could have turned out differently?***
17.	To what extent did you find the game challenging?	To what extent did you find the **concepts and themes of the movie, TV show, or clip** challenging?
18.	Were there any times during the game in which you just wanted to give up?	Were there any times **when you just wanted to give up watching**?
19.	To what extent did you feel motivated while playing?	To what extent did you feel motivated while **watching**?
20.	To what extent did you find the game easy?	To what extent did you find the **concepts and themes of the movie, TV show, or clip easy to understand**?
21.	To what extent did you feel like you were making progress towards the end of the game?	To what extent did you feel like you were making progress towards **understanding what was happening during the movie, TV show, or clip, and what you thought might happen at the end?**

22.	How well do you think you performed in the game?	How well do you think you **understood what happened in the movie, TV show, or clip?**
23.	To what extent did you feel emotionally attached to the game?	To what extent did you feel emotionally attached to the **movie, TV show, or clip**?
24.	To what extent were you interested in seeing how the game's events would progress?	To what extent were you interested in seeing how the **events shown in the movie, TV show, or clip** would progress?
25.	How much did you want to "win" the game?	How much did you want **the events in the movie, TV show, or clip to unfold successfully for the main characters involved?**
26.	Were you in suspense about whether or not you would win or lose the game?	Were you in suspense about **how the events would unfold in the movie, TV show, or clip?**
27.	At any point did you find yourself become so involved that you wanted to speak to the game directly?	At any point did you find yourself become so involved that you wanted to speak to the **movie, TV show, or clip** directly?
28.	To what extent did you enjoy the graphics and the imagery?	*Unchanged*
29.	How much would you say you enjoyed playing the game?	How much would you say you enjoyed **watching the movie, TV show, or clip**?
30.	When interrupted, were you disappointed that the game was over?	When interrupted, were you disappointed that **you had to stop watching**?
31.	Would you like to play the game again?	Would you like to **watch more of this in the future**?

Table 5.1: Modifications made to the original Immersive Experience Questionnaire (IEQ) to create the Film IEQ (changes in bold). Questions marked with * are taken from Green and Brock (2000)

5.4 EXPLORATORY FACTOR ANALYSIS

Following the development of the questionnaire, and in line with Jennett et al. (2008), it was desirable to examine the underlying factor structure to better understand how the concept of immersion is constructed and measured. To accomplish this, an exploratory factor analysis (EFA) was performed, which is used to search for a smaller set of latent factors that represent the variables measured in the questionnaire (Henson and Roberts, 2006).

5.4.1 METHOD

5.4.1.1 PARTICIPANTS

The questionnaire was completed by 415 participants. The first 213 participants were recruited via posts on forums, social media, and through university mailing lists. After exhausting these channels for new responses, the study was also listed on websites where participants were rewarded for participation in order to increase the sample size. The remaining participants received either UCL course credit for participants recruited through the UCL psychology subject pool, or a payment of £0.90 for participants recruited through the crowdsourcing website Prolific.ac. Participants were required to have watched a movie or TV episode in the previous three days.

5.4.1.2 MATERIALS

The Film IEQ was administered using an online form, and consisted of a single page featuring all of the questions. At the very top was a section detailing the study to allow participants to give informed consent.

5.4.1.3 PROCEDURE

The questionnaire was distributed to participants through websites, email, social media, and crowdsourcing platforms, as detailed above. They were invited to help with a scientific research project about how immersed people feel when watching video media. Participants were asked to fill in the questionnaire while thinking about the last movie, show, or episode they watched in the previous three days. To aid recall, participants were asked to provide

some information about what they watched: the title; one of the main actors, characters, presenters or other personnel that featured prominently; a location that featured prominently; and a brief synopsis of what happened. These were only to help participants remember what they watched, and were not used in the analysis.

5.4.2 RESULTS

Prior to analysis the data were checked for missing values, which resulted in one questionnaire response being removed. This left 414 responses on which the EFA was performed. Total immersion scores were computed for each of the participants by first inverting the response of negatively scored items (7 becomes 1, 6 becomes 2, 5 becomes 3, etc.), and then by summing the results of all of the responses to give a value between 31 and 217. Observed immersion scores observed ranged from 48 and 182 ($M = 139.61$, $SD = 16.36$).

As there are a number of subjective decisions to be made when performing an EFA, recommendations and guidance from prior research were broadly followed (Osborne and Costello, 2009; Beavers et al., 2013). Prior to conducting the EFA, the sampling adequacy was tested to ensure the factorability of the variables. A Kaiser-Meyer-Olkin test was performed, which was $.850$ (above the recommended value of $.6$) and Bartlett's test of sphericity was significant ($p < .001$), meaning the the data was suitable for EFA (Beavers et al., 2013).

5.4.2.1 Factor Extraction and Retention

One of the first and most important steps in performing an EFA is to decide how many factors to extract. Multiple methods exist to do this, though two are most commonly used (Osborne and Costello, 2009). First, the eigenvalue-one criterion ("Kaiser's criterion"), which discards factors with an eigenvalue < 1 (Kaiser, 1960). Second, the scree test method (Cattell, 1966), which plots the factors and their eigenvalues on a graph, then retains only those before the point where the line starts to level off horizontally. A less common, though arguably better method (Ledesma and Valero-Mora, 2007), is Horn's Parallel Analysis (Horn, 1965), where random datasets are generated and compared to the current dataset. Due to the inherently subjective nature of deciding on the number of factors, researchers have been advised to assess multiple criteria and use reasoned reflection when deciding on the number of factors (Henson and Roberts, 2006).

For the data in the present study, the eigenvalue one criterion, the scree plot method, and Horn's parallel analysis were all considered. Using the eigenvalue-one criterion suggested five factors, while examining the scree plot suggested three factors. Furthermore, a parallel analysis was also performed, which suggested eight factors. As parallel analysis has be found to be one of the best methods for establishing the number of factors (Ledesma and Valero-Mora, 2007), an eight factor solution was first considered. However, this resulted in a number of crossloaded items (items that load onto more than one factor) which were then removed, which led to some factors containing fewer than three items. As guidance from previous research suggests that factors with fewer than three items are unstable (Yong and Pearce, 2013), these were then further removed. As this resulted in the removal of a large number

of items which compromised the sensitivity of the immersion measure, the number of factors to extract was then repeatedly reduced by one and the analysis repeated until a satisfactory solution was obtained with four factors — i.e., without a large amount of crossloaded items, without factors with fewer than three items, and without a large amount items that did not load onto any factor.

Researchers also have to decide which factor extraction method to use, and which factor rotation method to use in order to make interpreting the data easier and reveal a simple structure (Corner, 2009). Again, there are no absolute guidelines in making these decisions. Multiple factor extraction methods and rotation methods were attempted, until the four-factor solution was arrived at which seemed to best fit the data. This used the maximum likelihood method of extraction, and a direct Oblimin rotation which allows the factors to correlate. A value of .32 was used as a cutoff for factor loadings (Comrey and Lee, 2013), which resulted in questions 5, 7, 10, 15, and 16 being removed. Additionally, crossloaded items 18 and 23 were also removed. The analysis was then repeated and four factors were extracted. This resulted in a 24-item scale (see Table 5.2), giving overall immersion scores between 24 and 168. The resulting factors and their loadings can be seen Table 5.3

5.4.2.2 Factor Identification

After the four factors were retained, the constituent questions were examined and the factors were given appropriate titles to describe them. These were *captivation, real-world dissociation, comprehension,* and *transportation*; they will each be described in turn (Table 5.2 shows an overview of the factors and the questions they contain).

#	Question
1.	To what extent did the movie, TV show, or clip hold your attention?
2.	To what extent did you feel you were focused on the movie, TV show, or clip?
3.	How much effort did you put into watching the movie, TV show, or clip?
4.	Did you feel that you were trying your best to follow the events of the movie, TV show, or clip?
5.*	To what extent did you feel consciously aware of being in the real world whilst watching?
6.*	To what extent were you aware of yourself in your surroundings?
7.*	To what extent did you notice events taking place around you?
8.	To what extent could you picture yourself in the scene of the events shown in the movie, TV show, or clip?
9.	To what extent did you feel like you were separated from your real-world environment?
10.	To what extent did you feel that the movie, TV show, or clip was something you were experiencing, rather than something you were just watching?
11.	To what extent was your sense of being in the environment shown in the movie, TV show, or clip stronger than your sense of being in the real world?
12.	To what extent did you find the concepts and themes of the movie, TV show, or clip challenging?
13.	To what extent did you feel motivated to keep on watching?
14.	To what extent did you find the concepts and themes easy to understand?
15.	To what extent did you feel like you were making progress towards understanding what was happening, and what you thought might happen at the end?
16.	How well do you think you understood what happened?
17.	To what extent were you interested in seeing how the events in the movie, TV show, or clip would progress?
18.	How much did you want the events in the movie, TV show, or clip to unfold successfully for the main characters involved?
19.	Were you in suspense about how the events would unfold?
20.	At any point did you find yourself become so involved that you wanted to speak to the movie, TV show, or clip directly?
21.	To what extent did you enjoy the cinematography, graphics and/or imagery?
22.	How much would you say you enjoyed watching the movie, TV show, or clip?
23.	When it was over, were you disappointed that you had to stop watching?
24.	Would you like to watch more of this, or similar content, in the future?

Table 5.2: Film IEQ questions, coloured by factor. Negatively scored items marked with an asterisk (*).

The first factor was captivation. This consists of 12 items (questions 1–4, 13, 17–19, 21–24) regarding how much the viewer enjoyed watching the video content, how interested they were, and their motivation to watch.

The second was real-world dissociation, and consists of three items (questions 5–7) regarding how much the viewer was aware of their real world surroundings.

	1. Captivation	2. Real-world dissociation	3. Comprehension	4. Transportation
Q.1	.832			
Q.2	.780			
Q.3	.426			
Q.4	.472			
Q.5		-.704		
Q.6		-.803		
Q.7		-.735		
Q.8				.597
Q.9				.652
Q.10				.797
Q.11				.872
Q.12			-.509	
Q.13	.698			
Q.14			.803	
Q.15			.364	
Q.16			.729	
Q.17	.623			
Q.18	.375			
Q.19	.354			
Q.20				.366
Q.21	.509			
Q.22	.806			
Q.23	.441			
Q.24	.662			

Table 5.3: Pattern matrix showing factors and factor loadings (values below 0.32 omitted).

The third, comprehension, consists of four items (questions 12, 14–16) asking about how well the concepts and themes of the video content were understood.

The fourth was transportation, consisting of five items (questions 8–11, 21). This factor describes how much the viewer felt like they were experiencing the events for themselves, and how much they felt they were located in the world portrayed in the video content.

5.5 DISCUSSION

This chapter has examined the concept of immersion, and attempted to transfer a well-grounded definition from computer games research to the domain of video consumption. Even given the similarities due to using much of the same source material and questionnaire items, the experience of immersion when watching video appears to be different from immersion experienced when playing games. When examining the factor structure, the Film IEQ revealed a four-factor structure rather than the five principle components in the IEQ, suggesting that this definition of immersion in video media constitutes fewer latent variables. Such differences may be due to the "lean back" nature of watching film and TV, where the user has little or no interaction with what they are watching and no autonomy in deciding the outcome of events. This is in contrast to the "lean forward" nature of playing games, where the player has direct control over the outcome, which is reflected in the control and challenge factors extracted from the IEQ by Jennett et al. (2008).

There are some similarities between the questionnaires. Both appear to measure a real-world dissociation factor, suggesting that escapism from the real world is a common element. This has been shown to be a motivation for playing games (Yee, 2006) for some players who prefer the exploration and role-playing elements of gaming, and it has been suggested that the psychological detachment that these experiences can afford is beneficial to players, potentially aiding post-work recovery (Collins and Cox, 2014). Similarly, Kubey (1986) notes that television is often chosen as an activity to escape negative feelings caused by work, as well as reality in general.

The captivation factor contains many of the items measuring cognitive

involvement and emotional involvement in the IEQ, suggesting that both cognitive and emotional investment is a common indicator of immersion across both media. This is supported by Busselle and Bilandzic (2009), who argue that following a narrative in non-interactive media requires both cognitive and emotional processes, and can result in a state of flow in some cases.

Due to the Film IEQ having one fewer factor than the IEQ, some items from different factors of the IEQ loaded onto the same factor in the Film IEQ. E.g., the Film IEQ factor comprehension was loaded with items from challenge and cognitive involvement factors of the IEQ. This seems logical — the comprehension factor is concerned with how well the viewer is understanding and following the video content, would would involve cognitive resources and could also present a challenge in some cases.

It is possible that some participants being rewarded for participation may have affected their responses. For this reason, the reward and non-reward sample groups were compared to check for any disparities. Immersion scores were computed, then plotted on a graph and assessed visually. This revealed no obvious differences between the samples. An independent samples t-test was also conducted to compare the samples, and there was no evidence that these groups differed significantly in terms of total immersion scores, $t(413) = .943, p = .346$.

The main contribution of this questionnaire is that allows for the development of a more detailed understanding of the effects of certain interventions on viewer experience — which things negatively affect people's immersion, which things positively affect it, and exactly how it is affected by examining the factor subscales. When compared with other measures of viewing experience, the Film IEQ goes beyond this and attempts to measure a more holistic

definition of experience. Furthermore, it also provides a standardised way of making these comparisons through a tool that is easier to deploy than other methods, such as the observational study in Chapter 3 and the diary study in Chapter 4. It is important to note that a questionnaire cannot simply replace these methods, but can augment them when used as part of a mixed methods approach, leading to a richer understanding of specific behaviours.

5.6 CONCLUSION

This chapter details the development of the Film IEQ. The resulting questionnaire provides a method of better understanding how various interventions affect viewer experience. Exploratory factor analysis revealed a four-factor model consisting of captivation, real-world dissociation, comprehension, and transportation. While this is different from the original IEQ, there were similarities with the extracted factors of both questionnaires. This suggests that although immersion experienced while watching film and TV is different from that of computer games, they share common elements.

In the following two chapters, this questionnaires utilised to develop a deeper understanding of two common behaviours observed in the studies in Chapter 3 and Chapter 4, watching on devices of different sizes, and being interrupted by device notifications while viewing. A lab study method will be used for both of these, which will also serve to validate the questionnaire through its use.

CHAPTER 6

INVESTIGATING THE EFFECT OF SCREEN SIZE ON IMMERSION

The following publication is based on work featured in this chapter:

Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J.J. Gould. 2016. Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (MobileHCI '16). ACM, New York, NY, USA, 714-721. DOI: `http://doi.org/10.1145/2957265.2961843`

6.1 INTRODUCTION

Through the collection of situated data, the previous studies in Chapters 3 and 4 have given rich qualitative insights into how on-demand services are a popular way of accessing and consuming video content, often across a range of devices. These multi-device ecosystems allow users to complete computing tasks at times and places that are convenient for them — e.g. it is no longer necessary to bring a heavy laptop on the train to check emails or to watch videos, as this can all be achieved on a smartphone or tablet (Levin, 2014). Devices can be seamlessly synchronised with personal accounts, with user interfaces sensitive to the benefits and constraints of each particular device.

While cross-device applications and services offer greater flexibility and convenience for users, they also presents a lack of control for developers and content producers over the experience provided to the user — consider a mobile game that may be easy to control when using a tablet touchscreen, but could be more frustrating on a smaller phone screen and therefore provide a worse experience to the player. An example of this is the popular game Angry Birds, where cartoon birds are catapulted across the screen. Some stages in this game require a high amount of precision, which is easier to achieve on a larger screen where the user can be more accurate with their finger movements when aiming. A study by Thompson et al. (2012) investigated the effect of touch screen size on game immersion by comparing a small iPod screen to a larger iPad screen, and found that a higher level of immersion was experienced when playing on the larger screen.

The variety of screen sizes present in everyday personal computing is also present in the domain of TV and film. Large screens are often thought of as providing a better experience — moviegoers often pay a premium for large IMAX screens, whose website describes themselves as "the world's most immersive movie experience"[1], and people often purchase large TVs for an enjoyable home viewing experience. Furthermore, a number of prominent directors have expressed their belief that watching movies on phones results in a lesser experience, such as George Lucas[2], Spike Lee[3], and David Lynch[4].

Despite the apparent preference for larger screens, content is being in-

[1] http://www.imax.com/ [Accessed 17th July 2018]

[2] http://www.techinsider.io/george-lucas-discusses-watching-movies-on-phones-2015-4 [Accessed 17th July 2018]

[3] http://www.theguardian.com/film/2015/may/01/spike-lee-watching-movies-digitally-is-heartbreaking [Accessed 17th July 2018]

[4] http://www.digitalspy.com/movies/news/a491726/david-lynch-watching-movies-on-a-smartphone-is-pathetic/ [Accessed 17th July 2018]

creasingly consumed on mobile devices. A recent UK communications market report (Ofcom, 2015b) showed that subscriptions to on-demand video services are increasing, and that subscribers are using them more and more. On-demand content accounted for 15% of all viewing for adults, with 33% of the online population watching on a computer, 21% on a smartphone, and 23% on a tablet at least once a month. Viewing on a variety of devices was also recorded in the diary study in Chapter 4, where 29% of viewing sessions featured viewing on a phone or tablet. Given this wide variety of viewing devices, it is possible that they may provide the user with differing experiences in the same way as has been shown for games (Thompson et al., 2012). Does a viewer watching a movie on a smartphone have a comparatively worse experience than if they watch it on a larger screen?

In this study, participants watched video content on three different devices to see if screen size correlated with self-reported immersion scores. A lab experiment was chosen for this, as it allows researchers to examine specific phenomena in detail. This can be difficult when using other methods, e.g. in the wild studies (such as the video observation study in Chapter 3), where the behaviour of interest may not manifest itself for the duration of the study. Conversely, a lab study allows certain phenomena to be recreated artificially, which can then be accurately measured. Furthermore, a greater degree of control of confounding factors allows researchers to isolate phenomena and their effects.

6.2 Related Work

Past research has investigated how people respond to different sized screens when watching film and television, but has not focused on measuring a well-defined concept of immersion. Section 2.4.1 reviews relevant literature on this subject, showing that various experiential measures are sensitive to screen size. However, these measures are quite narrowly defined and do not give a holistic view of the viewer's experience. In this chapter, the Film IEQ detailed in Chapter 5 is used to measure viewer immersion to give a broader insight into viewer experience. Furthermore, this study also intends to validate the Film IEQ through its use.

6.3 Method

6.3.1 Aims and Hypotheses

The aim of this study was to investigate the effect that screen size had on the level of immersion felt by participants when watching film content, which was self-reported by participants using the Film IEQ described in Chapter 5. In the diary study in Chapter 4, participants expressed a general dislike of mobile viewing, and previous work has shown more intense responses to film content shown on larger screens (Lombard et al., 1997; Reeves et al., 1999; Lombard et al., 2000; IJsselsteijn et al., 2001), and higher levels of immersion when playing games on larger screens in a study similar to the present study (Thompson et al., 2012). Given this, the hypothesis is that larger screens will lead to higher immersion scores in general.

6.3.2 PARTICIPANTS

A total of 19 participants (12 female, 7 male) were recruited through the UCL Psychology subject pool. They were granted course credit for 50 minutes of their time.

6.3.3 DESIGN

The study used a within subjects design. The independent variable was the screen size of the device they were watching the footage on, and there were three levels: a 4.5-inch phone, a 13-inch laptop and a 30-inch monitor. The dependent variable was the level of immersion the participants reported using the Film IEQ.

6.3.4 MATERIALS

The experiment took place in a lab with a desk present for participants to sit at using a fixed chair. Three devices were used to play the clips using the Netflix online streaming service: a Motorola Moto G smart phone with a 4.5-inch screen (held in the participants' hands with their arms on the desk); a Dell laptop with a 13-inch screen (placed on the desk approx. 50 cm away), and a 30-inch monitor (also placed 50 cm away). Participants used the laptop to select a movie from the Netflix website, which was required to be one they wanted to watch but had not yet seen. The first 30 minutes of this was split into three 10-minute clips. Audio was played through Sony over-ear headphones in order to control for sound level. Before the experiment, participants completed a questionnaire to collect demographic information, and after watching each clip they completed the Film IEQ detailed in the previous chapter.

6.3.5 PROCEDURE

Participants were greeted, and then asked to read the information sheet and sign a consent form. They were then given the opportunity to ask any questions they had. To begin the study, participants were seated at the desk and were told how the study would proceed and what they should do, then they filled in a questionnaire to collect demographic data. They were then asked to choose a movie to watch from the Netflix catalogue, which they had not seen before but would like to see. They watched the first 10 minutes of their chosen movie on their first assigned device and filled out the Film IEQ. The next 10 minutes were then watched on the second device followed by filling out another Film IEQ, then finally the remaining 10 minutes were watched on the remaining device followed by the final Film IEQ. The order of the devices was counterbalanced to control for order effects. Finally, participants were given another opportunity to ask questions before leaving.

6.4 RESULTS

Immersion scores were calculated by summing all questions in the Film IEQ. Questions 5, 6, and 7 were scored negatively.

Mean immersion scores were lower in the phone condition ($M = 106.05$, $SD = 15.53$) than in the laptop ($M = 114.47$, $SD = 12.42$) or monitor conditions ($M = 116.89$, $SD = 13.55$). A one-way repeated measures ANOVA was used to analyse this data, and showed a significant main effect of screen size on immersion score, $F(2, 36) = 5.09$, $p = .011$. Post-hoc t-tests were performed to examine pairwise differences between each condition, using Bonferroni corrections. Results found a significant difference in immersion

score between the phone condition and the laptop condition, $t(18) = 2.65$, $p = .048$, and the phone condition and the monitor condition, $t(18) = 2.69$, $p = .045$. There was no significant difference in immersion score between the laptop condition and the monitor condition, $t(18) = .48$, $p > .99$. These results suggest that watching content on a very small screen results in lower immersion than when watching content on a much larger screen.

While the above analyses give a detailed understanding of how immersion varies between different screen sizes, we were interested in the various subscales of the constituent factors of the Film IEQ (see Chapter 5). To examine this we performed a series of one-way repeated measures ANOVAs on responses to each subscale of the questionnaire. As can be seen in Table 6.1, there was a significant main effect of screen type on comprehension, $F(2, 36) = 5.48$, $p = .008$. However, there was no effect of screen size on captivation, Real-world Dissociation, or transportation, all p values > 0.05. Post-hoc tests were again performed on the significant subscale. Paired sample t-tests revealed significant differences in the phone-monitor conditions of the comprehension subscale (see Table 6.2).

Immersion factor	***F* (2, 36)**	***p***
Captivation	2.85	.108
Real-World Dissociation	2.01	.149
Comprehension	5.48	.008*
Transportation	2.49	.098

Table 6.1: Repeated measures ANOVA results for each subscale of the Film IEQ. p values $< .05$ marked with an asterisk.

Immersion factor	Conditions	*t*(18)	*p*
Comprehension	Phone - Laptop	2.09	.153
	Phone - Monitor	3.48	.008*
	Laptop - Monitor	1.09	.867

Table 6.2: Post-hoc paired sample t-test results for the statistically significant subscale. p values $< .05$ marked with an asterisk.

6.5 DISCUSSION

The hypothesis that larger screen sizes would result in greater immersion was supported. The significant main effect of screen size across conditions suggests that it is more difficult to experience high levels of immersion when viewing very small screens, but after a certain size there is less of a disparity. This fits with the results of the study by Thompson et al. (2012), where immersion scores reported using the IEQ when playing a simple game were significantly lower on a 3.5-inch screen than on a 9.7-inch screen. It appears that this result is consistent with the findings concerning mobile viewing from the diary study in Chapter 4, where participants often spoke of their dislike of small screens and would prefer to watch on large screen if possible. Nonetheless, mobile viewing was present in 29% of all of the sessions recorded in the diary study, showing that although it was not preferable, people still find it acceptable in some situations. Furthermore, this result also serves to further validate the Film IEQ as a immersion measurement tool.

The reason why a larger screen seems to provide a better experience is not well understood when looking at previous literature observing responses to screen size. Hatada et al. (1980) found that viewers perceived a greater sense of realism when viewing on larger screens. Furthermore, some research argues that increased responses could be due to the way humans perceive

objects on screen as larger (Troscianko et al., 2012), and that images in a video are not just representational, but are objects in their own right (Detenber and Reeves, 1996). Therefore bigger images are bigger objects, which can cause a reaction at a primitive level. Consider an image of spider — for some people, larger images may illicit stronger responses than smaller ones, because they are closer to their real-world counterparts. Such responses are involuntary and cannot be controlled, and can generate different emotions and actions.

When examining the subscales of the questionnaire, only the comprehension factor was significantly affected by screen size (questions 12, 14, 15, and 16 of the Film IEQ). These questions asked about how well the viewer understood the concepts and themes of the video, as well as what was happening in general. This result suggests that the smaller screen size leads to a lower level of understanding for viewers — this could perhaps be due to missing small details that may be critical to the plot, or because the small screen fills less of their visual field, potentially allowing more distractions from outside of the video (though an impact on the real-world dissociation factor might be expected in this case). Some research has also suggested that larger images can improve memory for the content, which may aid comprehension when the viewer has to remember details of what they have previously seen Lombard and Ditton (1997). Additionally, having to physically hold the device may have introduced some discomfort, which could also have been distracting.

Examining the remaining subscales of captivation, real-world dissociation, and transportation did not produce any statistically significant results. This suggest that screen size does not affect the viewers' enjoyment and interest in the video (captivation), their awareness of their real-world surroundings (real-world dissociation), or how much they feel like they are experiencing

events for themselves (transportation). This would go some way to explaining why watching content on smaller screens is fairly common — while some elements of immersion are affected (comprehension), it is still possible to have a enjoyable and immersive experience.

One confounding factor in this study was the freedom of choice of content that the participants were given when selecting content to view, as it is likely that some movies, or certain sections of them, could be considered more immersive than others. When designing the experiment, we considered giving every participant the same video to watch, but as the IEQ is partially based on factors determined by personal preference (e.g. Q21, "To what extent did you enjoy the cinematography, graphics and/or imagery?", and Q22, "How much would you say you enjoyed watching the movie, TV show, or clip?") it was decided that participants should have the freedom to choose content that would give the best experience. However, the issue still remains that some participants may have watched a more immersive clip than others. Interestingly, a number of participants did pick the same movie to watch, but sample sizes were too small to perform any statistical analyses — four participants chose *The Wolf of Wall Street*, two chose *The Dallas Buyer's Club*, two chose *She's Funny That Way*, and two chose *The Hunger Games: Mocking Jay Part 1*. It is quite possible that this was because of the way Netflix places certain movies on the front page, making them more likely to be chosen.

6.6 CONCLUSION

With the popularity of on-demand film and TV content being rising year on year, and with content increasingly being consumed on devices other that the

traditional television, it is important to examine how viewer experience may be affected by these different screen sizes. In this chapter, a lab-based experiment was conducted to examine the relationship between screen size and self-reported immersion scores while watching content on screens of three different sizes. It was found that watching content on a the small 4.5-inch phone screen recorded the lowest immersion scores, and that there was a significant main effect of screen size on immersion scores when compared to both the laptop and monitor screens. There was no significant effect when comparing the laptop and monitor screens. This suggests that watching content on a phone screen results in a worse experience than watching on a medium or large screen, and viewers wanting a more immersive experience should reserve content for larger screens.

Furthermore, this chapter also provides a level of validation for the Film IEQ instrument, as the results agree with the results of similar studies. The following chapter details another lab study that investigates the phenomenon of being interrupted by mobile devices, further investigating behaviours observed previously in this thesis while also further validating the Film IEQ.

CHAPTER 7

INVESTIGATING THE EFFECT OF MOBILE NOTIFICATION INTERRUPTIONS ON IMMERSION

The following publication is based on work featured in this chapter:

Jacob M. Rigby, Duncan P. Brumby, Sandy J. J. Gould, and Anna L. Cox. 2017. Film, interrupted: investigating how mobile device notifications affect immersion during movies. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '17). ACM, New York, NY, USA, Article 93, 8 pages. DOI: `http://doi.org/10.1145/3098279.3122136`

7.1 OVERVIEW AND MOTIVATION

In Chapter 6, it was shown that watching content on a small screen resulted in a less immersive experience that watching on a larger screen. This study was performed under lab conditions without any other distractions. However, in reality television viewing is often accompanied by many interruptions or secondary activities. A prominent example is the external interruptions present in the form of notifications from mobile devices, which prompt the user to stop what they are doing and attend to the device (Iqbal and Horvitz, 2010). Instant messaging interruptions, often driven by notifications, have be shown to be detrimental to performance in the workplace environment (Czerwinski

et al., 2000a,b). Following this, it seems reasonable that similar interruptions may have a detrimental effect on the viewer's experience when watching a movie or TV. Therefore, this study aims to answer the question of whether interruptions in the form of mobile device notifications affect a viewer's immersion when watching movies. This is important to both viewers, who may wish to have the best watching experience possible when watching their favourite shows and movies, and also content producers, who need to ensure they are not negatively affecting a viewer's experience if interrupting the viewer to provide second-screen content.

The relationship between immersion and interruptions was investigated across two emergent viewing paradigms: watching content on a traditional living room TV with mobile devices present, and watching content on mobile devices. Participants were split into two groups — one group watched content on the TV while having the phone present, the other group watched content on the phone itself. Both groups received notifications on the phone. This study aims to see if interruptions have negative effects in both of these contexts in terms of the immersion experienced by the participants. A lab study was chosen for this for the same reasons as the previous chapter (See Section 6.1) — greater experimental control, and the ability to study specific phenomena in detail.

7.2 Related Work

Much of the previous work investigating interruptions shows that being interrupted can be detrimental to performance in a workplace environment, and in safety critical environments such as driving or piloting a plane (see Sec-

tion 2.3 for a review of relevant literature). This is also true when looking specifically at interruptions in the form of email and instant messaging notifications (Czerwinski et al., 2000a,b), where it has been shown that some workers work longer and more consistently when notifications are turned off (Iqbal and Horvitz, 2010; Mark et al., 2012).

There is also work examining pervasive interruptions in the form of device notifications from a day-to-day perspective (Sahami Shirazi et al., 2014; Pielot et al., 2014), which reveal that users have to deal with high volumes of notifications every day. Notifications from mobile devices differ from traditional desktop notifications in a number of ways (Sahami Shirazi et al., 2014). Firstly, they are presented in a unified way. Notifications are often delivered to the user with broadly similar sounds, vibrations and graphical presentation, regardless of the nature of the type of notifications (e.g. text message, email, application update, etc.). Secondly, they are used to provide information to the user about a wide variety of applications and services, so much that it can lead to an overload of notifications (Church and de Oliveira, 2013). Fischer et al. (2010) showed that users assign differing levels of importance to notifications from different sources, depending on the context and time. When presented in a unified manner, users cannot differentiate between these levels of importance without attending to the notification. Finally, they are nearly always with the user due to the pervasive nature of mobile devices, which can lead to the formation of checking habits (Oulasvirta et al., 2012). Regardless of the interruption and frustration sometimes caused by notifications, some previous work has shown that users to attach value to notifications (Iqbal and Horvitz, 2010), and hence completely disabling them is often not welcomed.

Little research has examined the impact of interruptions during leisure

time, specifically when watching films and TV. On the basis of the related work showing the negative effects of interruptions, the hypothesis for this study is that interruptions in the form of notifications will negatively affect immersion across both devices. Furthermore, immersion will be further reduced in the phone condition due to the smaller screen. This is also informed by the previous study (see Chapter 6), which showed that the Film IEQ measure was sensitive to the effect of interventions that have been shown to affect viewer experience in prior work; i.e. that watching content on smaller screens results in lower immersion (Thompson et al., 2012), consistent with findings by Lombard and Ditton (1997); Reeves et al. (1999); IJsselsteijn et al. (2001); Thompson et al. (2012).

7.3 METHOD

7.3.1 PARTICIPANTS

A total of 29 participants (14 female, 15 male) were recruited through the UCL Psychology subject pool. The mean age was 25.1 years ($SD = 7.5$), and ages ranged between 18 and 52 years. They were paid £7.50 for 40 minutes of their time.

7.3.2 DESIGN

The study used a 2 $\times$ 2 (device $\times$ interruption) mixed factorial design. Device was manipulated by having participants watch content on either a phone or computer monitor, which was a between subjects manipulation. Interruption was manipulated by regularly interrupting the participants with phone

notifications in one condition and removing interruptions in the other. This was a within subjects manipulation. The dependent variables were the participants' self-reported level of immersion, measured using the Film IEQ (see Chapter 6), and the time to taken to respond to notifications.

7.3.3 MATERIALS

The experiment took place in a dedicated usability lab, with a desk present for participants to sit at using a fixed chair. A Motorola Moto G smart phone with a 4.5-inch screen (held in the participants' hands with their arms on the desk) and a 30-inch monitor (placed on the desk approx 50 cm away) plugged into a laptop were used to play the clips. Both were loaded with Netflix to allow the participants to play the content, which was selected using the 13-inch laptop screen. Three participants chose *Kung Fu Panda 2*, two chose *Ride Along*, two chose *Maleficent*, and the remaining participants chose something unique. Participants were required to select an unseen movie that they wanted to watch, which was freely selected by the participant from the entire Netflix catalogue. The first 20 minutes of this were split into two 10-minute clips. Audio was played through a pair of desktop speakers to control for sound level and quality.

Participants were also sent notifications on the smart phone through a basic messaging app written for this experiment. Participants were alerted to the notification through a notification sound and a vibration, and the notification was positioned in the centre of the screen. Figure 7.1 Shows the a screenshot of one of the notifications participants were presented with. The app asked the participant five simple questions, the order of which was shuffled:

- "What did you eat for breakfast today?";

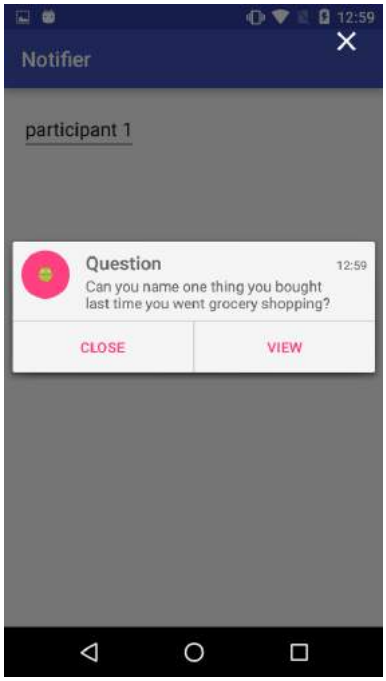


Figure 7.1: A screenshot of a notification shown to participants.

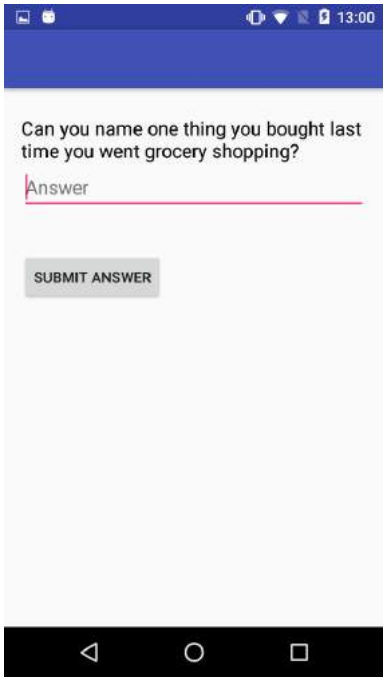


Figure 7.2: A screenshot of the answer submission screen.

- “Can you name one thing you bought last time you went grocery shopping?";
- “What are your plans for the weekend?";
- “What was the last restaurant you went to?";
- “If you were to get a new pet, what would it be?";

After viewing a notification, they were taken to another screen where they could fill in their answer using the on screen keyboard and submit it, as shown in Figure 7.2. The app recorded the time it took for participants to act upon the notifications they received as well as their responses. Before the experiment, participants filled out a short questionnaire to collect demographic information, and after watching each clip they filled out the Film IEQ to measure their level of immersion.

7.3.4 PROCEDURE

Participants were greeted, and then asked to read the information sheet and sign a consent form. They were then given the opportunity to ask any questions they had. To begin the study, participants were seated at the desk and were told how the study would proceed and what they should do, then they filled in a questionnaire to collect demographic data. Every participant was given a smart phone (even those not watching content on the phone) and told they may receive a notification containing a message which they should respond to, which they were shown how to do. They were then asked to choose an unseen movie that they would like to watch by using the Netflix streaming service catalogue, and depending on their assigned device they watched the first 10 minutes on either the phone or monitor then filled out the Film IEQ, then they watched the next 10 minutes on the same device followed by filling out another Film IEQ. For one of the clips, the participants received notifications on the smart phone which asked them to answer simple questions.

In the notification condition, participants watched uninterrupted for five minutes, then were asked a question every minute for five minutes. This left an uninterrupted minute of viewing before the clip finished. In addition to the notification appearing as a pop-up in the centre of the screen, it was accompanied by a notification sound and a vibration. Participants responded to the question by using the phone's on-screen keyboard and pressing a submit button. Both the order of whether participants received notifications and the order of the questions were counterbalanced to control for order effects. After participants had watched two clips, participants were given another opportunity to ask questions and then left.

7.4 RESULTS

Immersion scores were calculated by summing all questions in the Film IEQ. Questions 5, 6, and 7 were scored negatively.

A factorial ANOVA was conducted to compare the main effects of the presence of interruptions and device and the interaction effect between interruptions and device on the level of self-reported immersion. Mean immersion scores were lower in the notification condition ($M = 110.24, SD = 25.29$) than in the uninterrupted condition ($M = 118.38, SD = 14.43$). Statistical analysis found that there was a significant main effect of interruptions on immersion score, $F(1, 27) = 4.54$, $p = .042$. Mean immersion scores were slightly lower in the phone condition ($M = 113.21, SD = 16.8$) than in the monitor condition ($M = 115, SD = 15.08$). There was no significant main effect of screen size on immersion scores, $F(1, 27) = .82, p = .373$ (see Figure 7.3 for a bar plot of condition means). There was no significant interaction between the screen size and the presence of interruptions on immersion scores, $F(1, 27) = 1.06, p = .311$.

The various subscales of the Film IEQ were also analysed, and the results are shown in Table 7.1. When examining within subjects (interruption) effects, the results show there was a significant main effect of interruptions on real-world dissociation, $F(1, 27) = 8.627, p = .007$. There was no effect of interruptions on captivation, comprehension, or transportation, all p values > 0.05. When examining between subjects effects, there was no significant main effect of device on captivation, real-world dissociation, comprehension, or transportation, all p values > 0.05.

Finally, the response times between the participants receiving the notifi-

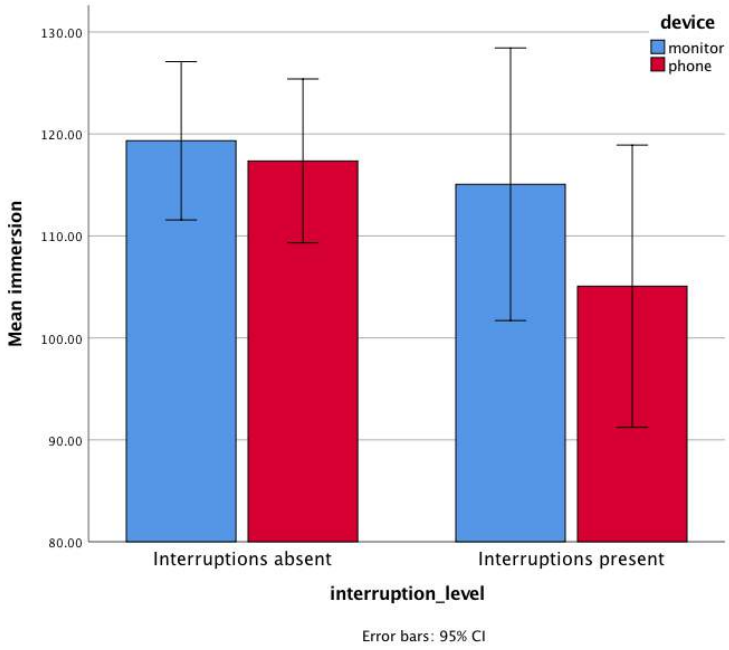


Figure 7.3: 2x2 bar plot showing condition means

	Interruption effects		Device effects	
Immersion factor	***F* (1, 27)**	***p***	***F* (1, 27)**	***p***
Captivation	.205	.654	.1	.754
Real-World Dissociation	8.627	.007*	.284	.598
Comprehension	2.453	.129	.973	.333
Transportation	1.791	.192	.069	.795

Table 7.1: Factorial ANOVA results for each subscale of the Film IEQ, showing both within subjects (interruption) effects and between subjects (device) effects. p values $< .05$ marked with an asterisk.

cation and submitting their answer were analysed. The mean response time across the five answered questions was computed for each participant, and these were analysed using an independent samples t-test to compare the phone and monitor groups. The results showed that the monitor group had significantly lower response times (14.46 seconds, $SD = 4.62$) than the phone group (19.12, $SD = 6.38$), $t(26) = 0.36$, $p = .036$.

7.5 DISCUSSION

The main finding of the study was that interruptions lead to lower immersion scores regardless of the device used for viewing, which supports the hypothesis. This finding agrees with previous work showing that interruptions are detrimental a number of environments, including the workplace safety-critical settings. This has implications for the viewer, because it shows that constant interruptions from mobile devices may be leading them to have a worse experience than they would otherwise. Interruptions have been shown to affect emotional state and feelings of annoyance (Zijlstra et al., 1999; Bailey and Konstan, 2006), and so this may lead to feelings of frustration. This also has implications for producers and distributors of content, because they risk reduced engagement, or even losing their audience altogether, if viewers become disinterested as a result of frustration caused frequent interruptions.

One solution is to simply disable notifications, but we know that users value the awareness they provide (Iqbal and Horvitz, 2010), and may even add to the viewing experience (e.g. discussing shows with friends). Furthermore, social pressures often make people feel obliged to attend to notifications (Pielot and Rello, 2015). With this in mind, these findings reinforce the case for greater management of notifications. This could be performed manually at operating system level, or through use of intelligent notification management systems such as that suggested by Iqbal and Bailey (2008), which suggest delivering notifications at breakpoints in a task. For office-based tasks break points are frequent and identifiable (e.g., switching between documents); this is not the case for film and television, where it is not clear where breaks points lie. Consider a three-hour movie: should notifications be disabled for the

entirety of the movie, or collected and batch-released at intervals? More intelligent approaches could also be taken, e.g. a smart watch might detect user movement to deliver interruptions when they are fidgety, which has been linked to lower immersion (Bianchi-Berthouze et al., 2006), and boredom could be inferred by pressure-sensitive seats to detect restlessness (D'Mello et al., 2007) or through machine learning learning applied to phone usage data (Pielot et al., 2015). Electronic programme guides could be used to establish start and finish times of TV programmes, or a notification manager could detect cinematic cuts which can provide cognitive separation of events (Schwan et al., 2000). While some previous work has examined managing attention for multi-screen TV settings (Neate et al., 2015), this has focused of specially designed companion experiences and do not account for spontaneous interruptions.

Somewhat unexpectedly, immersion scores were not found to be significantly lower when watching on the phone. This is inconsistent with the results of the previous study in Chapter 6 investigating how screen size affects immersion, where smaller screens were found to lead to reduced immersion, and the study by Thompson et al. (2012) which that smaller screens led to lower levels of immersion when playing games. It is also inconsistent with other studies looking at the effects of screen size, where larger screens elicited greater presence (Lombard and Ditton, 1997; IJsselsteijn et al., 2001) and stronger physiological responses (Reeves et al., 1999). A possible explanation for this is that the previous study used a within subjects design where all participants watched on all devices, and could therefore directly compare the screen sizes of the devices. While there was no significant affect of device, the mean immersion scores were slightly lower for the phone group than the

monitor group across both the interruption and no interruption conditions.

Analysis of the subscales of the questionnaire showed that there was a significant main effect of interruption on Real World Dissociation. This may be because immersion is something that develops progressively (Brown and Cairns, 2004), and being interrupted constantly resets this feeling by bringing the participants out of the movie world and back into the real world.

Analysis of the response times revealed that participants watching on the monitor had significantly lower response time than those watching on the phone. This may be because the content was automatically paused when the notification appeared when watching on the phone, but was not when watching on the monitor. This perhaps meant participants felt less pressured to respond as they were not missing the movie.

A limitation of this study was the type of task that was chosen as in reality notifications come from various sources (Church and de Oliveira, 2013). Furthermore, the questions asked (see Section 7.3.3) were fairly generic and did not require much thought. Participants were also likely to be aware that they were answering a computer, so may not have put as much thought into the answers compared with if they were answering a friend.

Another limitation is the lab setting, which could be seen as unsuitable for studying living room behaviour. This feeds into a wider discussion about experimental control versus ecological validity when considering research methods in this domain. A number of methods have been used when looking at device usage while watching television, each with pros and cons. Researchers have often relied on self reporting from participants such as surveys (Foehr, 2006; Rideout et al., 2010), diary studies (Foehr, 2006; Rideout et al., 2010; Vanattenhoven and Geerts, 2012), and interviews (Vanattenhoven and

Geerts, 2012). These can be effective in ascertaining general practices, but can be unsuitable for developing a fine-grained understanding of specific behaviours. One solution is to conduct situated studies (e.g Holz et al. (2015)), but the living room environment presents a number of challenges (Brown et al., 2014), such as difficulties with data collection, experimental setup in a non-controlled environment, and monitoring attention over multiple screens. On the other hand, controlled experiments can be less resource-intensive to conduct and allow for a detailed understanding of very specific behaviours. It appears that there is no panacea for this problem, and that incorporating a number of complimentary methods can allow us to develop a balanced understanding when conducting research in this domain.

Finally, as in the previous study in Chapter 6, participants selecting stimuli could present a lack of control. While it is possible to expose participants to the same stimuli, the immersion measure is partially based on personal interest, and for this reason it was hoped that participant choice would maximise potential immersion.

7.6 CONCLUSION

This study investigated how external interruptions in the form of device notifications affected participants' experience when watching a movie, measured using self reported immersion scores. It was found that interruptions resulted in significantly lower immersion no matter which device was used. This suggests that viewers who wish to experience the maximum amount of immersion when watching content should attempt to minimise interruptions, perhaps by disabling some of the notifications on their mobile devices during these times.

From the perspective of designers and producers of film and TV who wish to incorporate second screen content, the results suggest that they should at least be mindful of this effect, and be careful not to interrupt the viewer during times when they wish for them be fully immersed.

The study found that there was no main effect of device on the level of immersion reported by participants, which disagrees with related work in the area, an well as the previous study in this document. Possible reasons for this may be the device being a between subjects variable in this study vs. a within subjects variable in the previous study, or due to the sampling or number of participants.

This chapter concludes the empirical work in this thesis. The following chapter provides a general discussion of the findings of this thesis, as well as their implications.

CHAPTER 8

GENERAL DISCUSSION

This chapter begins by briefly summarising the research undertaken in this thesis. A more detailed look at the research is then taken with reference to the research questions in order to conclusively answer them. The contributions that this research makes are then clearly presented, along with the limitations and implications of this work. Finally, areas of future research are suggested.

8.1 SUMMARY OF RESEARCH AND METHODOLOGY

This thesis features five chapters of empirical work, using a variety of methodologies in order to utilise a mixed methods approach. The first two were intended to explore the new ways and contexts in which people are consuming video media, and how it fits into their everyday lives. The first study, presented in Chapter 3, established the prevalence of media multitasking with mobile devices and TV. An in the wild approach was used by placing cameras in participants' homes in order to observe their behaviour. Following this, Chapter 4 details a 14-day diary study undertaken by nine households to investigate how people are interacting with on-demand video services. This chapter also explored participants' perceptions of on-demand services via interviews.

The studies in Chapter 3 and Chapter 4 were effective in establishing people's everyday viewing behaviours, and the contexts and locations in which

people viewed. Following this, specific phenomena that were observed in these two studies were focused on in detail, using a questionnaire developed for this research and lab experiments. To do this, three empirical chapters explore how viewer immersion can be used as a measure to assess the impact of these specific emergent behaviours. Chapter 5 describes the development of a questionnaire (Film IEQ) to measure viewer immersion, through adaptation of a computer games questionnaire and an Exploratory Factor Analysis. This was then used for two lab experiments. In Chapter 6, a lab study investigated how watching on small screens affects viewer immersion, a practice that has become common with on-demand services. The affect of interruptions on immersion while viewing was investigated in a lab study in Chapter 7, both when viewing on a TV screen and on a phone.

As described above, this research utilised a mixed methods approach in order to answer the overarching research questions, combining both qualitative and quantitative approaches, and both situated studies and lab experiments. This meant the strengths of some of the methods used could make up for the weaknesses of others (Johnson and Onwuegbuzie, 2004), and is a widely-used approach in HCI (van Turnhout et al., 2014). For instance, the in the wild study in Chapter 3 showed that participants frequently used their mobile devices while watching TV, though it was not possible to know exactly how this affected their viewing experience. The particular phenomenon of being interrupted by notifications while viewing was then isolated and investigated in detail in a lab study, which allowed for greater experimental control and systematic manipulation of variables in order to understand the problem and minimise any confounding factors. Using such a mixed methods approach allowed for the problems to be focused on from different perspectives, leading

to a richer understanding of the outcomes (Creswell and Creswell, 2017). In the case above, it was found in Chapter 7 that the interruptions were detrimental to viewer immersion, which would have been very difficult to isolate and measure outside of the lab environment. Similarly, viewing on a variety of screen sizes was recorded in the diary study in Chapter 4, which was then isolated and studied in the lab in Chapter 6 to reveal that small screens led to lower immersion. This would have been difficult to measure as part of the diary study.

8.2 Revisiting the Research Questions

This thesis has investigated new ways in which people are consuming video media, namely media multitasking with mobile devices and TV, and using on demand services as a means of accessing video, and sought to answer the following research questions:

RQ 1: How prevalent are technology-driven emergent viewing practices?

RQ 2: What are the motivations for participating in these viewing practices?

RQ 3: How do these viewing practices affect viewer experience?

The following sections will show how this research has answered these questions.

8.2.1 Concurrent TV and Mobile Device Usage

8.2.1.1 RQ 1: Prevalence

In Chapter 3, a situated video observation study was conducted to establish real-world levels and patterns of mobile device usage while watching television over a period of three evenings. Such a study was not found to have been performed before in the literature. It was found that device usage was highly individual, ranging from 0-23% of TV time, and from 0-28 individual uses recorded over the duration of the study. Different patterns of usage were observed that could sit at different points on the multitasking continuum, from frequent, short uses to much longer, less frequent uses.

8.2.1.2 RQ 2: Motivation

Ophir et al. (2009) suggest that media multitasking propensity is a stable behavioural trait rather that being influenced by situational factors, and Loh and Kanai (2014) found physiological differences between high and low media multitaskers. This may help to explain motivations for engaging in this behaviour. While very specific "in the moment" motivations can be difficult to measure accurately, the evidence of such a trait could explain general motivations. This was explored by comparing participants' MMI scores (a measure of media multitasking propensity) to the amount of media multitasking observed in Chapter 3. This showed MMI score to be a significant predictor of real-world media multitasking, helping to explain some of the motivations behind this — some may just be more likely to media multitask than others.

8.2.1.3 RQ 3: EFFECTS

A number of studies in the past have evaluated the effect of the use of specific companion apps (e.g. (Murray et al., 2012; Silva et al., 2015; Neate et al., 2015)). However, outside of these highly designed and curated experiences little work has examined the effects of ad-hoc device usage. A controlled lab study was conducted to assess the effects of media multitasking in a common living room scenario: watching a video while being interrupted by chat notifications. Interruptions from notifications were found to be detrimental to immersion, mirroring the findings of other research in other areas — interruptions are not only bad in the specialist domains of knowledge work (González and Mark, 2004), piloting a plane (Latorella, 1998) and healthcare (Li et al., 2011), but also to the average person trying to unwind and enjoy watching a movie.

8.2.2 USE OF ON-DEMAND SERVICES

8.2.2.1 RQ 1: PREVALENCE

The real-world extent of on-demand viewing was first uncovered in the video observation study in Chapter 3, where is was found that viewing on-demand content made up 26% of all viewing time. Interaction with these services was investigated in a more holistic manner via the diary study in Chapter 4, and confirmed a high level of usage. Households recorded a mean average of 20.6 viewing sessions each (SD = 9.1), detailing a mean of 20:57:20 of viewing (SD = 08:10:11). Viewing in the living room was most popular, closely followed by the bedroom. 56% of viewing took place in the evening (between 6pm and midnight), and 69% of sessions were one hour or less in duration.

Furthermore, 89% of sessions were two hours or less. Moreover, this study also established that there were various contextual factors that affected various aspects of viewing, such as which devices were used, when they watched, and how long for. For instance, while the living room was the most popular location for viewing, this depended on the device used — when considering only mobile viewing, the bedroom was the most popular location, followed by the kitchen. Similarly, viewing after 11pm was dominated by mobile devices.

8.2.2.2 RQ 2: MOTIVATION

Motivations for using on-demand viewing platforms were explored through the interviews conducted for the diary study in Chapter 4. Overwhelmingly, participants referred to the freedom and choice that these platforms provide, which allowed them to watch content they wanted to watch according to their own schedule and circumstances. Participants also referenced the wide range of content available, meaning there was generally something that they wanted to watch.

Participants also spoke about the ability to watch multiple episodes of a series, and/or to watch for long periods of time, including the phenomenon of binge watching. This was a complex issue; many participants said that it was something they were wary of because it was "addictive". Due to this, some participants said how they tried to avoid watching too much, even going so far as to create their own boundaries such as not installing on-demand applications on certain devices. Therefore, this may have discouraged use in some cases. However, when considering total usage time, the ability to easily watch for long periods was likely a motivator for use, encouraging people to watch more often and for longer periods due to the availability of content.

Other features, such as automatically playing the next episode, were also cited as features contributing to this.

8.2.2.3 RQ 3: EFFECTS

One of the by-products of the freedom provided by on-demand viewing platforms is the consumption of video on different sized devices, as recorded in the diary study in Chapter 4. The effects of viewing on a small screen versus a large one were investigated in a controlled experiment in Chapter 6, and found that watching on a very small screen led to a less immersive experience than watching on larger screens, confirming the general belief that watching on larger screens is typically a "better" experience. In line with this, the diary study in Chapter 4 found that participants were reluctant to view on mobile devices, typically only doing so when larger screens were not available or were inconvenient to use. The smaller screen size was generally cited as the reason for this, with participants consistently saying they prefer to watch on a TV.

Another effect on viewing experiences was watching large volumes of content in a single sitting, which is a significant departure from the weekly episode "appointment viewing" that was common in the past. This lead to potential instances of binge watching, and long viewing sessions of up to six hours in duration. While this was often seen as something driving engagement with on-demand services, when considering wider effects of these behaviours outside of purely viewing experiences participants spoke about how this could easily get out of control, and affect other areas of their lives due to the amount of time spent watching.

8.3 CONTRIBUTIONS

This research makes a number of contributions to the current body of research in the area of new video experiences, which are described in turn in this section.

8.3.1 THE NATURE OF MEDIA MULTITASKING

Studies looking at concurrent TV and mobile device usage have typically relied on self-report measures such as questionnaires and interviews, with little direct observation conducted. The study in Chapter 3 did exactly this through the use of cameras placed in participants' homes. As a result, this observational data provides a necessary real-world insight into modern TV viewing behaviour. By analysing video data, a number of accurate conclusions could be drawn.

Firstly, individual media multitasking behaviour is subject to large individual differences. Some participants were observed to be heavy media multitaskers, spending a large proportion of overall TV time interacting with their mobile devices, or exhibiting a large number of individual device interactions. On the other hand, some participants did not use their mobile devices at all. Related to this is the correlation observed between total device usage and Media Multitasking Index score (MMI), and device uses per hour and MMI score. For both of these measures, MMI was found to be an accurate predictor of actual media multitasking behaviour. Though the MMI measure has attracted some criticism in other research (Baumgartner et al., 2017; Wiradhany and Nieuwenstein, 2017), these findings suggest its effectiveness as a correlate for media multitasking behaviour.

Secondly, different patterns of media multitasking behaviour were observed, such as long infrequent uses and shorter, more frequent uses. The multitasking continuum can be used to understand these better (Salvucci et al., 2009); allowing occurrences of media multitasking to be categorised according to the time between task switches. This means specific instances of multitasking could be concurrent multitasking (switching between tasks in quick succession, such as talking and driving), sequential multitasking (switching between tasks at a less frequent pace, such as cooking and reading a book), or somewhere in between. This allows us to better understand these individual occurrences, and means they can be viewed in the context of the wider literature of multitasking outside of the specific domain focused on here. This can in turn be used to advance the development of interventions that improve the experience of multitasking for users, such as through notification management systems.

8.3.2 Extent and Context of On-Demand Service Usage

Much like the research surrounding media multitasking, most studies about the prevalence of on-demand service usage have focused on general surveys and qualitative interview studies. Furthermore, those that do take a more fine-grained look at people's behaviours are generally out of date, being conducted before on-demand video services were commonplace. Therefore, it is difficult to establish exactly how people are interacting with these services in their everyday lives. This was addressed by conducting the diary study in Chapter 4, which sought to ascertain exactly how much people were using these services, as well as the context surrounding this use.

This thesis not only established how much people are using these services,

but also goes further in establishing the context around this usage. This contributes to our understanding of the decisions people make when watching and why they do so, such as which device to use and in which locations to watch. This study, along with the study in Chapter 3 contributes to the wider body of literature that examines TV and media use in everyday life (e.g. (Lull, 1980; Kubey, 1990; D'heer et al., 2012; Holz et al., 2015)) as well as other emerging technologies (e.g. (Taylor and Harper, 2003; Brown et al., 2013; Pizza et al., 2016)).

8.3.3 Immersion as an Experiential Metric for Watching Video

Immersion has been a well-studied concept in interactive systems such as computer games (Jennett et al., 2008) and virtual reality (Witmer and Singer, 1998). However, experiential measures in non-participatory media such as when reading or watching video typically look to the concept of (spatial) presence, which describes only the sense of being physically present in a mediated location. The research in this thesis adapted the relevant concepts and themes of Jennett et al. (2008)'s well-defined concept of immersion to the area of video consumption, as well as incorporating some questions relating to narrative engagement (Busselle and Bilandzic, 2009) allowing for a multifaceted experiential measure that still incorporates elements of presence. This resulted in the Film IEQ, which can be used as a standardised way to assess immersion when watching video. Comparing the factor structure of the Film IEQ with the original IEQ reveals that there are both similarities and differences with regard to how immersion can be defined and measured in both games and video media. A more detailed look at these similarities and dif-

ferences could be an interesting avenue for future research in order to further understand the nature of immersive experiences.

The Film IEQ provides a metric for activities that are not task-based so that the impact of various interventions can be quantified outside of these environments, which is important in an environment where error rates and task completion cannot readily be measured. This allows us to better understand the effects of interventions, and the underlying factor structure of the questionnaire can also be examined in order to establish exactly where the effects are felt. Furthermore, the Film IEQ provides a convenient and less resource-intensive method of measurement when compared with other techniques, such as physiological measurement or direct observation.

8.3.4 Small Screens Negatively Affect Viewer Immersion

Previous work has shown that screen size can affect viewer experience, for example in terms of presence (Lombard et al., 2000; IJsselsteijn et al., 2001) and physiological responses (Reeves et al., 1999). The study in Chapter 6 showed that watching on smaller screens can also lead to a reduction in viewer immersion. This may be expected as the definition of immersion used is related to presence, but it also serves to show that presence is not the only element of experience to be affected by small screen sizes. Indeed, it was found that the comprehension subscale of immersion was significantly reduced when watching on a phone. These overall findings agree with a study by Thompson et al. (2012), who found that video game immersion was similarly reduced when using smaller screens.

8.3.5 Interruptions While Watching Video Have Detrimental Effects

The effect of interruptions and task switching has been a strong focus of research in a number of areas such as the workplace (e.g. González and Mark (2004)), and safety-critical settings such as driving (e.g. Brumby et al. (2009)) and medical care (e.g Li et al. (2011)). In these areas, detrimental effects from interruptions can have serious consequences, at worst resulting in loss of life. However, less attention has been given to the potential negative effects of interruptions in leisure time. While such effects may not be immediately salient as those mentioned above, leisure time is nonetheless important to people's psychological and physical wellbeing (Pressman et al., 2009), allowing for relaxation, post-work recovery, and time to spend with friends and family. As such, the importance of quality leisure time should not be trivialised, and care should be taken to safeguard it.

This thesis contributes to the body of research on interruptions by specifically showing that interruptions while watching TV can also have detrimental effects in terms of reduced immersion. Being able to quantify this draws attention to effects of interruptions during leisure time, allowing us to understand the implications of this and look toward developing ways of mitigating these problems.

8.4 Limitations

A mixed methods approach was used in order to mitigate the shortcomings of using any single method. However, some limitations are still present in this research.

8.4.1 Ecological Validity and Experimental Control

Lab studies were conducted as part of this research in order to study very specific behaviours and phenomena: the effect of screen size on viewer immersion (see Chapter 6), which has become a concern due to the rise in mobile viewing; and the effect of interruptions from mobile notifications on viewer immersion (see Chapter 7), which have become ubiquitous in everyday life.

Studying behaviours which occur during leisure time and in the home can present difficulties when translating these to a lab setting — participants may behave differently under lab conditions, without their personal devices and outside of their own routines. This could call into question the ecological validity of using such methods. Careful consideration was given to each of the studies presented in this thesis to ensure the most appropriate study methodology was chosen for each one. A lab study was selected to be the best choice for the two above studies for a number of reasons. Firstly, the lab setting allows for a level of experimental control which can otherwise be difficult to achieve, for instance ensuring that all participants use the same devices in the same environment, which would be problematic in the wild. Secondly, the specific behaviours that were studied are often ephemeral, and for some participants may occur only infrequently (or not at all) outside of the lab setting. These phenomena can be reliably recreated in the lab, and therefore in these instances lab studies were selected as the best way of examining them.

On the other hand, it could also be argued that the studies in this thesis that were conducted in the wild (see Chapter 3 and Chapter 4) suffer due to a lack of control. Again, different study designs were considered when choosing the best way to conduct these, and as they are mostly concerned with reporting

real-world behaviour, observational and diary methods were selected as the best fit. As these studies took place over multiple days (three and 14 days respectively), lab studies would have been inappropriate. Such long durations were favourable in these instances, as it gave enough time for the participants to exhibit a range of behaviours of interest.

Overall, there does exist a tension between ecological validity and experimental control when conducting research in this domain. Furthermore, previous research has long acknowledged the numerous difficulties that can present themselves in the living room setting (Lull, 1980; Bernhaupt et al., 2008; Obrist et al., 2009; Brown et al., 2014). Unfortunately, when dealing with activities such as video consumption for leisure, which is often conducted in periods of unstructured time, it appears there is no silver bullet to allay all of these difficulties. For this reason, careful consideration of the most suitable research methodologies to use, and employing a mixed methods approach offer a way to mitigate some of these concerns. As such, that was the approach taken in this thesis.

8.4.2 Immersion as a Measure

Another point to consider is the type of content with which the Film IEQ should be used. In order to be used with a variety of content, general language was used where possible, e.g. "movie, TV show, or clip". It is intuitively applicable to videos portraying story-driven content, and indeed the studies in which the questionnaire was used both used movies as stimuli; it may be that immersion is affected differently with different types of stimuli (e.g sport or documentaries). This could provide an avenue for further research in order to assess the sensitivity to different types of content.

The broader question of "what is immersion?" can still be asked, and is unlikely to be conclusively answered by researchers in the future. The use of a subjective term certainly leaves room for disagreement For instance, in the field of virtual reality (VR), Slater (2003) describes how immersion is a concept that can be objectively assessed, and is independent of the human response to it (for which the author uses the concept of presence), disagreeing with fellow VR researchers Witmer and Singer (1998), who describe it as a psychological state. Murray (2017) also describes immersion as a subjective experience — "the sensation of being surrounded by a completely other reality" — and seems to be a common view outside of VR circles. The work in this thesis subscribes to this notion, and used the concept of immersion to apply to people's subjective experience based on a popular definition of immersion computer games (Jennett et al., 2008). However, as this thesis uses a modified version of Jennett et al. (2008)'s measure, it is still fair to question if immersion is still being measured. In response to this, it is argued that the immersion measure used in this thesis is still measuring a person's response to media experiences, which is grounded in similar concepts, and which is sensitive to various interventions. Though this definition does not align with the definition Slater (2003) uses, it does agree that it may be an issue of terminology — sometimes, people are talking about the same things but using different words to describe them.

8.4.3 Sample Sizes

It is possible that some of the samples sizes for the qualitative studies in Chapter 3 (four households, 12 participants) and Chapter 4 (nine households, 20 participants) could be considered low. While in general larger sample sizes are

better, it can be argued that the sample sizes chosen are suitable for the type of research being conducted. Furthermore, they are in line with similar studies in the same area, especially when considering there are multiple participants in each household. E.g. Barkhuus and Brown (2009), 21 participants; Shokrpour and Darnell (2017), 10 participants; McNally and Harrington (2017), 24 participants; Vanattenhoven and Geerts (2015), seven households. Perhaps contributing to this was a difficulty in recruiting participants for these studies. This may be due to the effort involved in participating, the duration of the studies, and also potential invasions of privacy. In order to compensate for these potential inconveniences participants were well paid for participation, though interest was still lower than had been hoped for. However, some argue that generalisability is not the primary aim of such qualitative research, but rather to shine a light on some of the common and interesting practices and phenomena that occur (Labuschagne, 2003). Furthermore, a common school of thought is that qualitative research can be generalisable given thoughtful study design (Blandford et al., 2016).

8.5 IMPLICATIONS

Aside from academic interest, this research has implications for a number of parties. Firstly, viewers themselves should consider their own behaviours and be mindful of the potential for negative effects. In terms of media multitasking, viewers should be aware of the potential for distraction and interruption that mobile devices can cause, and the effect this can have on their immersion. The negative effects of interruptions and distractions are taken seriously in other areas (Janssen et al., 2015), and interruptions while watching TV could have effects on physical and psychological wellbeing as they interrupt

leisure time (Pressman et al., 2009), and work-life balance if work and personal devices are combined (Stawarz et al., 2013). In light of the findings in this thesis, viewers might consider strategies to mitigate this, perhaps through operating system level interventions (such as "do not disturb mode"). Viewers could also simply turn off their devices, or leave them in another room, though due to the present-day ubiquity of mobile devices this could seem less appealing than it might have five or ten years ago – previous research has shown that people value the awareness that notifications provide (Iqbal and Horvitz, 2010) and often feel a social obligation to be available and attend to notifications (Pielot and Rello, 2015). In terms of on-demand video services, entire libraries of content can lead to viewers consuming large amounts of video, which can turn into binge watching. The research in this thesis found that often this was seen as a negative point by participants, who disliked the ease with which their viewing could get out of control. This sometimes lead to participants setting their own boundaries, such as not installing the service on their devices, and even completely disengaging with the service by not starting to view a new series over fears they would watch too much.

While this research has highlights issues that should be considered in order to maintaining positive viewing experiences and actions that can be taken by viewers, it seems unfair that the the responsibility should fall entirely to the viewers themselves. Firstly, the average viewer may not be aware of their own behaviour and how it might be leading to a worse experience, and therefore may not be motivated to take any action. Secondly, they may not have the knowledge or skills to know how they can improve the situation. It is therefore encouraging to see that technology companies like Google and Apple have started to take issues such as excessive screen time more seriously,

such as through Apple's recently announced enhanced screen time and notification management features for iOS [1], and Google's "take a break reminder" for YouTube [2] which pauses videos after a set amount of time. However, the efficacy of these particular features remains to be seen.

Content producers should also be aware of the new ways in which people are consuming their content. Watching on mobile device is now a popular way to consume video media, but offers a very different experience when compared with watching on a large screen in the cinema. As such, excessive small details and various cinematographic techniques may be lost on mobile viewers simply due to lack of screen real estate and lower resolution. We might envisage there being slightly different versions of the same TV show or movie, adapted for a viewer's specific viewing device. Consideration would need to be given to increased production time and cost, though technological solutions such as intelligent use of image cropping may be possible. Furthermore, the move towards Object-Based Broadcasting (Armstrong et al., 2014), where individual assets and metadata describing their relationships are broadcast separately and assembled on-the-fly, could help facilitate this kind of solution.

From the point of view of networks and advertisers, both on-demand content and mobile device usage during shows can offer a direct link to consumers which can be used to deliver adverts as well as extra content and features to keep viewers engaged. However, care should be taken when attempting to drive engagement at any cost. For instance, the "post play" feature on the

[1] `http://www.apple.com/newsroom/2018/06/ios-12-introduces-new-features-to-reduce-interruptions-and-manage-screen-time/` [Accessed 20th June 2018]

[2] `http://support.google.com/youtube/answer/9012523` [Accessed 20th June 2018]

Netflix platform automatically plays the next episode of a serious after a short period of time (this varies, but can be as little as five seconds). Objectively, this appears to be effective in driving engagement with their service by keeping viewers watching, but in the interviews conducted in this thesis, participants often said that excessive viewing was a concern for them, and some even specifically referenced their dislike at how the post play feature removes their agency. This could result in people using the service less in the long term due to the negative behaviours it could facilitate, which has been observed in other systems (Baumer et al., 2013). A recent study by Hiniker et al. (2018) examined the use of a tablet-based video platform designed for children, where they specifically examined the use of a post play feature. Their findings showed that the presence of such a feature resulted in significantly longer tablet sessions, and that parents were more likely to intervene, often causing conflict. Therefore, adding small design frictions to discourage automatic behaviours could lead to a better user experience in these cases (Cox et al., 2016).

8.6 Future Research

This thesis gives a firm grounding for future research to build upon the work presented here. One possible direction is to extend the in the wild observational study in Chapter 3 to include device logging (this was attempted, but ultimately failed). This would allow us to understand exactly what people are doing on their mobile devices at particular times. This could be coupled with identifying what is being watched on the TV at that particular time to give important contextual insights into behaviour. It may be that people are more likely to perform some tasks than others when watching certain content, e.g. using social media during a sporting event to share their opinions, or check-

ing emails when commercials start. However, this does open up ethical and privacy issues — participants were reluctant to have their personal devices monitored, which seemed to affect the ease with which participants could be recruited. Furthermore, video recording people in their homes can be seen as an invasion of privacy. Any future research on this topic should carefully consider what data is necessary to be collected in order to minimise the effects on participants and their recruitment.

Another worthwhile extension of this work would be in the area of interruption management systems, which can delay notifications until more opportune times. Interruptions can cause errors and reduce task performance, and so technological systems to manage these have been explored and a means of mitigating these issues, e.g. in the workplace (Iqbal and Bailey, 2008). Furthermore, notification management in day-to-day life has also been researched, typically looking at sensing a user's context to decide when to serve notifications, for instance using sensors (Kern and Schiele, 2003) or existing mobile interactions (Fischer et al., 2011). Other similar systems attempt to infer if a user is bored and therefore more susceptible to interruptions (Pielot et al., 2015), which can improve engagement with a system (Dingler et al., 2017). However, these systems have not been implemented specifically for leisure situations such as when watching video. Such a system could also incorporate knowledge of the content being watched to more intelligently manage notifications by holding them back until a less disruptive time, such as the time between programmes, cinematic cuts, or even parts of the video where there is less action. As mentioned above in this chapter there are some solutions that can help users manage their own notifications, but currently the onus is very much on the user and they are very much blunt instruments; more

sophisticated, automatic approaches could be explored.

8.7 CONCLUSION

This thesis has examined emergent video viewing practices motivated by the uptake of fast internet connections and mobile devices. Two emergent practices were focused on: media multitasking with TV and mobile devices; and usage of on-demand video services to access content. A mixed-methods approach was used, incorporating both qualitative and quantitative methods.

The prevalence of these practices was established through situated studies: an in the wild video observation study examined the extent media multitasking; and a diary study examined the extent on-demand service use. These studies found these practices to be common, and uncovered other behaviours surrounding them. Following this, specific phenomena resulting from these studies were then the focus of lab studies. A questionnaire to measure viewer experience in terms of the immersion they feel was developed, and was used in two experiments. The first experiment investigated the effect of viewing on different devices with varying screen sizes on immersion, and the second looked at the effect of interruptions from mobile notification on immersion. In both cases, it was found that both of these practices had the potential to negatively affect viewer immersion.

Overall findings of this thesis showed that the emergent viewing practices examined offered the potential for new and improved viewing experiences, but could also result in detrimental effects that could reduce viewer experience in some cases. All stakeholders in this domain should be aware of the positives and negatives of these practices in order to provide enjoyable view-

ing experiences. The work presented in this thesis makes a number of significant contributions to this research area, which can help us to better understand how people are consuming video media in the modern age. Furthermore, it provides a foundation for future work that might try and tackle some of these problems, such as through notification management systems.

APPENDIX A

THE IMMERSIVE EXPERIENCE QUESTIONNAIRE FOR FILM AND TELEVISION (FILM IEQ)

A.1 INTRODUCTION

This appendix contains the full Immersive Experience Questionnaire for Film and Television (Film IEQ) as well as instructions on how to use it and analyse the results. For further details of the development of the questionnaire, see Chapter 5.

A.2 USING THE QUESTIONNAIRE

To use the Film IEQ, participants should answer all questions in Table A.1 using 1–7 Likert scales. This typically takes less than five minutes. Calculate overall immersion scores by first inverting responses to items 5, 6, and 7 (1 becomes 7, 2 becomes 6, 3 becomes 5, etc.) then summing all responses. This results in an overall immersion score between 24 and 168. Individual factor subscale scores (for captivation, real-world dissociation, comprehension, and transportation) are calculated in the same way, first inverting responses where necessary then summing the responses to all questions within that subscale.

#	Question	Factor
1.	To what extent did the movie, TV show, or clip hold your attention?	1
2.	To what extent did you feel you were focused on the movie, TV show, or clip?	1
3.	How much effort did you put into watching the movie, TV show, or clip?	1
4.	Did you feel that you were trying your best to follow the events of the movie, TV show, or clip?	1
5.*	To what extent did you feel consciously aware of being in the real world whilst watching?	2
6.*	To what extent were you aware of yourself in your surroundings?	2
7.*	To what extent did you notice events taking place around you?	2
8.	To what extent could you picture yourself in the scene of the events shown in the movie, TV show, or clip?	4
9.	To what extent did you feel like you were separated from your real-world environment?	4
10.	To what extent did you feel that the movie, TV show, or clip was something you were experiencing, rather than something you were just watching?	4
11.	To what extent was your sense of being in the environment shown in the movie, TV show, or clip stronger than your sense of being in the real world?	4
12.	To what extent did you find the concepts and themes of the movie, TV show, or clip challenging?	3
13.	To what extent did you feel motivated to keep on watching?	1
14.	To what extent did you find the concepts and themes easy to understand?	3
15.	To what extent did you feel like you were making progress towards understanding what was happening, and what you thought might happen at the end?	3
16.	How well do you think you understood what happened?	3
17.	To what extent were you interested in seeing how the events in the movie, TV show, or clip would progress?	1
18.	How much did you want the events in the movie, TV show, or clip to unfold successfully for the main characters involved?	1
19.	Were you in suspense about how the events would unfold?	1
20.	At any point did you find yourself become so involved that you wanted to speak to the movie, TV show, or clip directly?	4
21.	To what extent did you enjoy the cinematography, graphics and/or imagery?	1
22.	How much would you say you enjoyed watching the movie, TV show, or clip?	1
23.	When it was over, were you disappointed that you had to stop watching?	1
24.	Would you like to watch more of this, or similar content, in the future?	1

Table A.1: Film IEQ questions, coloured and numbered by factor (1: captivation, 2: real-world dissociation, 3: comprehension, 4: transportation). Negatively scored items marked with an asterisk (*).

BIBLIOGRAPHY

Adamczyk, P. D. and Bailey, B. P. (2004). If not now, when?: The effects of interruption at different moments within task execution. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04, pages 271–278, New York, NY, USA. ACM.

Altmann, E. M. and Trafton, J. G. (2004). Task interruption: Resumption lag and the role of cues. In Proceedings of the 26th annual conference of the Cognitive Science Society, Cogsci '04.

Altmann, E. M., Trafton, J. G., and Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. Journal of Experimental Psychology: General, 143(1):215.

Alzahabi, R. and Becker, M. W. (2013). The association between media multitasking, task-switching, and dual-task performance. Journal of Experimental Psychology: Human Perception and Performance, 39(5):1485.

Angell, R., Gorton, M., Sauer, J., Bottomley, P., and White, J. (2016). Don't distract me when i'm media multitasking: Toward a theory for raising advertising recall and recognition. Journal of Advertising, 45(2):198–210.

Armstrong, M., Brooks, M., Churnside, A., Evans, M., Melchior, F., and Shotton, M. (2014). Object-based broadcasting - curation, responsiveness and user experience. IET Conference Proceedings, pages 12.2–12.2(1).

Arroyo, E. and Selker, T. (2011). Attention and intention goals can mediate disruption in human-computer interaction. In IFIP Conference on Human-Computer Interaction, pages 454–470. Springer.

Bailey, B. P. and Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. Computers in Human Behavior, 22(4):685–708.

BARB (2017). The viewing report 2017. `http://www.barb.co.uk/download/?file=/wp-content/uploads/2017/04/Barb-Viewing-Report-2017.pdf`.

Barkhuus, L. and Brown, B. (2009). Unpacking the television: User practices

around a changing technology. ACM Transactions on Computer-Human Interaction (TOCHI), 16(3):15:1–15:22.

Baumer, E. P., Adams, P., Khovanskaya, V. D., Liao, T. C., Smith, M. E., Schwanda Sosik, V., and Williams, K. (2013). Limiting, leaving, and (re)lapsing: An exploration of facebook non-use practices and experiences. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13, pages 3257–3266, New York, NY, USA. ACM.

Baumgartner, S. E., Lemmens, J. S., Weeda, W. D., and Huizinga, M. (2017). Measuring media multitasking. Journal of Media Psychology, 29(4):188–197.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., and Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. Practical Assessment, Research & Evaluation, 18(6):1–13.

Becker, M. W., Alzahabi, R., and Hopwood, C. J. (2013). Media multitasking is associated with symptoms of depression and social anxiety. Cyberpsychology, Behavior, and Social Networking, 16(2):132–135.

Bernhaupt, R., Obrist, M., Weiss, A., Beck, E., and Tscheligi, M. (2008). Trends in the living room and beyond: Results from ethnographic studies using creative and playful probing. Computers in Entertainment, 6(1):5:1–5:23.

Bianchi-Berthouze, N., Cairns, P., Cox, A., Jennett, C., and Kim, W. W. (2006). On posture as a modality for expressing and recognizing emotions. In BCS-HCI workshop: Emotion and HCI.

Biocca, F. (2003). Can we resolve the book, the physical reality, and the dream state problems? from the two-pole to a three-pole model of shifts in presence. In EU Future and Emerging Technologies, Presence Initiative Meeting.

Bittman, M. and Wajcman, J. (2000). The rush hour: The character of leisure time and gender equity. Social Forces, 79(1):pp. 165–189.

Blandford, A., Furniss, D., and Makri, S. (2016). Qualitative HCI re-

search: going behind the scenes. Synthesis Lectures on Human-Centered Informatics, 9(1):1–115.

Bracken, C. C. and Pettey, G. (2007). It is really a smaller (and smaller) world: Presence and small screens. In Proceedings of the 10th International Workshop on Presence, pages 283–290.

Brasel, S. A. and Gips, J. (2011). Media multitasking behavior: Concurrent television and computer usage. Cyberpsychology, Behavior, and Social Networking, 14(9):527–534.

Brasel, S. A. and Gips, J. (2017). Media multitasking: How visual cues affect switching behavior. Computers in Human Behavior, 77:258–265.

Brown, A., Evans, M., Jay, C., Glancy, M., Jones, R., and Harper, S. (2014). HCI over multiple screens. In CHI '14 Extended Abstracts on Human Factors in Computing Systems, CHI EA '14, pages 665–674, New York, NY, USA. ACM.

Brown, B., McGregor, M., and Laurier, E. (2013). Iphone in vivo: Video analysis of mobile device use. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI '13, pages 1031–1040. ACM.

Brown, B. A. T., Sellen, A. J., and O'Hara, K. P. (2000). A diary study of information capture in working life. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '00, pages 438–445, New York, NY, USA. ACM.

Brown, E. and Cairns, P. (2004). A grounded investigation of game immersion. In CHI '04 Extended Abstracts on Human Factors in Computing Systems, CHI EA '04, pages 1297–1300, New York, NY, USA. ACM.

Brumby, D. P., Du Toit, H., Griffin, H. J., Tajadura-Jiménez, A., and Cox, A. L. (2014). Working with the television on: An investigation into media multitasking. In CHI '14 Extended Abstracts on Human Factors in Computing Systems, CHI EA '14, pages 1807–1812, New York, NY, USA. ACM.

Brumby, D. P., Salvucci, D. D., and Howes, A. (2009). Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while

driving. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09, pages 1629–1638. ACM.

Burgess, P. W., Veitch, E., de Lacy Costello, A., and Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. Neuropsychologia, 38(6):848 – 863.

Bury, R. and Li, J. (2015). Is it live or is it timeshifted, streamed or downloaded? watching television in the era of multiple screens. New Media & Society, 17(4):592–610.

Busselle, R. and Bilandzic, H. (2009). Measuring narrative engagement. Media Psychology, 12(4):321–347.

Caird, J. K., Willness, C. R., Steel, P., and Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. Accident Analysis & Prevention, 40(4):1282 – 1293.

Cairns, P., Cox, A. L., Day, M., Martin, H., and Perryman, T. (2013). Who but not where: The effect of social play on immersion in digital games. International Journal of Human-Computer Studies, 71(11):1069–1077.

Cairns, P., Cox, A. L., and Nordin, A. I. (2014a). Immersion in digital games: A review of gaming experience research. Handbook of digital games, MC Angelides and H. Agius, Eds. Wiley-Blackwell, pages 339–361.

Cairns, P., Li, J., Wang, W., and Nordin, A. I. (2014b). The influence of controllers on immersion in mobile games. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, pages 371–380, New York, NY, USA. ACM.

Calvillo-Gámez, E. H. and Cairns, P. (2008). Pulling the strings: A theory of puppetry for the gaming experience. In Conference Proceedings of the Philosophy of Computer Games 2008, pages 308–323.

Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., and Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of americans. Computers in Human Behavior, 25(2):483–489.

Carter, S. and Mankoff, J. (2005). When participants do the capturing: The

role of media in diary studies. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05, pages 899–908, New York, NY, USA. ACM.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1(2):245–276.

Cecchinato, M. E., Sellen, A., Shokouhi, M., and Smyth, G. (2016). Finding email in a multi-account, multi-device world. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '16, pages 1200–1210, New York, NY, USA. ACM.

Cesar, P., Bulterman, D. C., and Jansen, A. J. (2008). Usages of the secondary screen in an interactive television environment: Control, enrich, share, and transfer television content. In Proceedings of the 6th European Conference on Changing Television Environments, EUROITV '08, pages 168–177, Berlin, Heidelberg. Springer-Verlag.

Church, K. and de Oliveira, R. (2013). What's up with whatsapp?: Comparing mobile instant messaging behaviors with traditional sms. In Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '13, pages 352–361, New York, NY, USA. ACM.

Collins, E. and Cox, A. L. (2014). Switch on to games: Can digital games aid post-work recovery? International Journal of Human-Computer Studies, 72(8-9):654–662.

Comrey, A. L. and Lee, H. B. (2013). A First Course in Factor Analysis. Psychology Press.

Conlin, L., Billings, A. C., and Averset, L. (2016). Time-shifting vs. appointment viewing: the role of fear of missing out within tv consumption behaviors. Communication & Society, 29(4).

Corner, S. (2009). Choosing the right type of rotation in PCA and EFA. JALT Testing & Evaluation SIG Newsletter, 13(3):20–25.

Courtois, C. and D'heer, E. (2012). Second screen applications and tablet users: Constellation, awareness, experience, and interest. In Proceedings

of the 10th European Conference on Interactive TV and Video, EuroITV '12, pages 153–156, New York, NY, USA. ACM.

Cox, A., Cairns, P., Shah, P., and Carroll, M. (2012). Not doing but thinking: The role of challenge in the gaming experience. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, pages 79–88, New York, NY, USA. ACM.

Cox, A. L., Gould, S. J., Cecchinato, M. E., Iacovides, I., and Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. In CHI '16 Extended Abstracts on Human Factors in Computing Systemss, CHI EA '16, pages 1389–1397, New York, NY, USA. ACM.

Creswell, J. W. and Creswell, J. D. (2017). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications.

Csikszentmihalyi, M. (1996). Creativity: Flow and the Psychology of Discovery and Invention. Modern classics. HarperCollins Publishers.

Czerwinski, M., Cutrell, E., and Horvitz, E. (2000a). Instant messaging and interruption: Influence of task type on performance. In OZCHI 2000 Conference Proceedings, volume 356, pages 361–367.

Czerwinski, M., Cutrell, E., and Horvitz, E. (2000b). Instant messaging: Effects of relevance and timing. In People and Computers XIV: Proceedings of HCI, volume 2, pages 71–76. British Computer Society.

Czerwinski, M., Horvitz, E., and Wilhite, S. (2004). A diary study of task switching and interruptions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04, pages 175–182, New York, NY, USA. ACM.

de Feijter, D., Khan, V.-J., and van Gisbergen, M. (2016). Confessions of a 'guilty' couch potato understanding and using context to optimize binge-watching behavior. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, TVX '16, pages 59–67, New York, NY, USA. ACM.

Deloitte, LLP (2017). Digital democracy survey. https://www.ngi-summit.org/wp-content/materials/Deloitte/us-tmt-deloitte-digital-democracy-executive-summary.pdf.

Detenber, B. H. and Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. Journal of Communication, 46(3):66–84.

Dey, A. K., Wac, K., Ferreira, D., Tassini, K., Hong, J.-H., and Ramos, J. (2011). Getting closer: An empirical investigation of the proximity of user to their smart phones. In Proceedings of the 13th International Conference on Ubiquitous Computing, UbiComp '11, pages 163–172, New York, NY, USA. ACM.

D'heer, E., Courtois, C., and Paulussen, S. (2012). Everyday life in (front of) the screen: The consumption of multiple screen technologies in the living room context. In Proceedings of the 10th European Conference on Interactive TV and Video, EuroiTV '12, pages 195–198, New York, NY, USA. ACM.

Dias, P. (2016). Motivations for multi-screening: An exploratory study on motivations and gratifications. European Journal of Communication, 31(6):678–693.

Dingler, T., Weber, D., Pielot, M., Cooper, J., Chang, C.-C., and Henze, N. (2017). Language learning on-the-go: Opportune moments and design of mobile microlearning sessions. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '17, pages 28:1–28:12, New York, NY, USA. ACM.

Dismukes, R. K., Loukopoulos, L. D., and Jobe, K. K. (2001). The challenges of managing concurrent and deferred tasks. In Proceedings of the 11th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University. Citeseer.

D'Mello, S. S., Chipman, P., and Graesser, A. (2007). Posture as a predictor of learner's affective engagement. In Proceedings of the Cognitive Science Society, volume 29.

du Toit, H. (2013). Working while watching TV, is it really work?: The impact of media multitasking on stress and performance. Master's thesis, University College London.

Duff, B. R.-L., Yoon, G., Wang, Z., and Anghelcev, G. (2014). Doing it

all: An exploratory study of predictors of media multitasking. Journal of Interactive Advertising, 14(1):11–23.

Dunning, D., Heath, C., and Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. Psychological Science in the Public Interest, 5(3):69–106.

Engström, H., Brusk, J., and Östblad, P. A. (2015). Including visually impaired players in a graphical adventure game: A study of immersion. IADIS International Journal on Computer Science and Information System, 10(2):95–112.

Fischer, J. E., Greenhalgh, C., and Benford, S. (2011). Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. In Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '11, pages 181–190, New York, NY, USA. ACM.

Fischer, J. E., Yee, N., Bellotti, V., Good, N., Benford, S., and Greenhalgh, C. (2010). Effects of content and time of delivery on receptivity to mobile interruptions. In Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '10, pages 103–112, New York, NY, USA. ACM.

Foehr, U. G. (2006). Media multitasking among american youth: Prevalence, predictors and pairings. Henry J. Kaiser Family Foundation.

Gauntlett, D. and Hill, A. (2002). TV Living: Television, Culture and Everyday Life. Routledge.

Gil de Zúñiga, H., Garcia-Perdomo, V., and McGregor, S. C. (2015). What is second screening? exploring motivations of second screen use and its effect on online political participation. Journal of Communication, 65(5):793–815.

González, V. M. and Mark, G. (2004). "constant, constant, multi-tasking craziness": Managing multiple working spheres. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04, pages 113–120, New York, NY, USA. ACM.

Gould, A. J. (2014). What Makes an Interruption Disruptive? Understanding

the Effects of Interruption Relevance and Timing on Performance. PhD thesis, UCL (University College London).

Green, M. C. and Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. Journal of Personality and Social Psychology, 79(5):701.

Gürkök, H. (2012). Mind the Sheep! User Experience Evaluation & Brain-computer Interface Games. PhD thesis, Enschede, the Netherlands.

Gysbers, A., Klimmt, C., Hartmann, T., Nosper, A., and Vorderer, P. (2004). Exploring the book problem: Text design, mental representations of space, and spatial presence in readers.

Hatada, T., Sakata, H., and Kusaka, H. (1980). Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. SMPTE Journal, 89(8):560–569.

Hembrooke, H. and Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. Journal of Computing in Higher Education, 15(1):46–64.

Henson, R. K. and Roberts, J. K. (2006). Use of exploratory factor analysis in published research common errors and some comment on improved practice. Educational and Psychological Measurement, 66(3):393–416.

Hiniker, A., Heung, S. S., Hong, S. R., and Kientz, J. A. (2018). Coco's videos: An empirical investigation of video-player design features and children's media use. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '18, pages 254:1–254:13, New York, NY, USA. ACM.

Holmes, M. E., Josephson, S., and Carney, R. E. (2012). Visual attention to television programs with a second-screen application. In Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA '12, pages 397–400, New York, NY, USA. ACM.

Holz, C., Bentley, F., Church, K., and Patel, M. (2015). "I'm just on my phone and they're watching TV": Quantifying mobile device use while watching television. In Proceedings of the ACM International Conference

on Interactive Experiences for TV and Online Video, TVX '15, pages 93–102, New York, NY, USA. ACM.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30(2):179–185.

Ichikawa, F., Chipchase, J., and Grignani, R. (2005). Where's the phone? a study of mobile phone location in public spaces. In 2nd Asia Pacific Conference on Mobile Technology, Applications and Systems, pages 1–8.

IJsselsteijn, W., Ridder, H. d., Freeman, J., Avons, S. E., and Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. Presence: Teleoperators and Virtual Environments, 10(3):298–311.

IJsselsteijn, W. A., de Ridder, H., Freeman, J., and Avons, S. E. (2000). Presence: Concept, determinants, and measurement. In Electronic Imaging, pages 520–529. International Society for Optics and Photonics.

Iqbal, S. T. and Bailey, B. P. (2005). Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In CHI '05 Extended Abstracts on Human Factors in Computing Systems, CHI EA '05, pages 1489–1492, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2007). Understanding and developing models for detecting and differentiating breakpoints during interactive tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07, pages 697–706, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2008). Effects of intelligent notification management on users and their tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08, pages 93–102. ACM.

Iqbal, S. T. and Horvitz, E. (2010). Notifications and awareness: A field study of alert usage and preferences. In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, CSCW '10, pages 27–30, New York, NY, USA. ACM.

Irani, L., Jeffries, R., and Knight, A. (2010). Rhythms and plasticity: television temporality at home. Personal and Ubiquitous Computing, 14(7):621–632.

Janssen, C. P., Gould, S. J., Li, S. Y., Brumby, D. P., and Cox, A. L. (2015). Integrating knowledge of multitasking and interruptions across different perspectives and research methods. International Journal of Human-Computer Studies, 79(C):1–5.

Jenner, M. (2015). Binge-watching: Video-on-demand, quality TV and mainstreaming fandom. International Journal of Cultural Studies, page 1367877915606485.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. (2008). Measuring and defining the experience of immersion in games. International Journal of Human-Computer Studies, 66(9):641 – 661.

Johnson, R. B. and Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. Educational Researcher, 33(7):14–26.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement.

Kern, N. and Schiele, B. (2003). Context-aware notification for wearable computing. In Proceedings of the seventh IEEE International Symposium on Wearable Computers, pages 223–230.

Knoche, H. and McCarthy, J. D. (2005). Design requirements for mobile TV. In Proceedings of the 7th International Conference on Human-computer Interaction with Mobile Devices and Services, MobileHCI '05, pages 69–76. ACM.

Kubey, R. (1990). Television and the quality of family life. Communication Quarterly, 38(4):312–324.

Kubey, R. W. (1986). Television use in everyday life: Coping with unstructured time. Journal of Communication, 36(3):108–123.

Kushlev, K., Proulx, J., and Dunn, E. W. (2016). "Silence your phones": Smartphone notifications increase inattention and hyperactivity symptoms. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '16, pages 1011–1020, New York, NY, USA. ACM.

Labuschagne, A. (2003). Qualitative research-airy fairy or fundamental? The Qualitative Report, 8(1):100–103.

Latorella, K. A. (1998). Effects of modality on interrupted flight deck performance: Implications for data link. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 42, pages 87–91. SAGE Publications.

Ledesma, R. D. and Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. Practical Assessment, Research & Evaluation, 12(2):1–11.

Lee, J., Lin, L., and Robertson, T. (2012). The impact of media multitasking on learning. Learning, Media and Technology, 37(1):94–104.

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. Presence, 10(3):282–297.

Levin, M. (2014). Designing Multi-device Experiences: An Ecosystem Approach to User Experiences across Devices. O'Reilly Media, Inc.

Li, S. Y., Magrabi, F., and Coiera, E. (2011). A systematic review of the psychological literature on interruption and its patient safety implications. Journal of the American Medical Informatics Association, 19(1):6–12.

Logan, R. J., Augaitis, S., Miller, R. H., and Wehmeyer, K. (1995). Living room culture - an anthropological study of television usage behaviors. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 39, pages 326–330. SAGE Publications Sage CA: Los Angeles, CA.

Loh, K. K. and Kanai, R. (2014). Higher media multi-tasking activity is associated with smaller gray-matter density in the anterior cingulate cortex. PLOS One, 9(9):e106698.

Lombard, M. and Ditton, T. B. (1997). At the heart of it all: The concept of presence. Journal of Computer-Mediated Communication, 3(2):0–0.

Lombard, M., Ditton, T. B., Grabe, M. E., and Reich, R. D. (1997). The

role of screen size in viewer responses to television fare. Communication Reports, 10(1):95–106.

Lombard, M., Ditton, T. B., and Weinstein, L. (2009). Measuring presence: The temple presence inventory. In Proceedings of the 12th Annual International Workshop on Presence, pages 1–15.

Lombard, M., Reich, R. D., Grabe, M. E., Bracken, C. C., and Ditton, T. B. (2000). Presence and television. the role of screen size. Human Communication Research, 26(1):75–98.

Lottridge, D. M., Rosakranse, C., Oh, C. S., Westwood, S. J., Baldoni, K. A., Mann, A. S., and Nass, C. I. (2015). The effects of chronic multitasking on analytical writing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '15, pages 2967–2970, New York, NY, USA. ACM.

Loukopoulos, L. D., Dismukes, R. K., and Barshi, I. (2003). Concurrent task demands in the cockpit: Challenges and vulnerabilities in routine flight operations. In Proceedings of the 12th International Symposium on Aviation Psychology, pages 737–742. Wright State University Press Dayton, OH.

Lui, K. F. and Wong, A. C.-N. (2012). Does media multitasking always hurt? a positive correlation between multitasking and multisensory integration. Psychonomic Bulletin & Review, 19(4):647–653.

Lull, J. (1980). The social uses of television. Human Communication Research, 6(3):197–209.

Lyons, G. and Urry, J. (2005). Travel time use in the information age. Transportation Research Part A: Policy and Practice, 39(2):257–276.

Mark, G., Gonzalez, V. M., and Harris, J. (2005). No task left behind?: Examining the nature of fragmented work. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 321–330. ACM.

Mark, G., Iqbal, S. T., Czerwinski, M., Johns, P., and Sano, A. (2016). Neurotics can't focus: An in situ study of online multitasking in the workplace. In Proceedings of the 2016 CHI Conference on Human Factors in

Computing Systems, CHI '16, pages 1739–1744, New York, NY, USA. ACM.

Mark, G., Voida, S., and Cardello, A. (2012). "A pace not dictated by electrons": An empirical study of work without email. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, pages 555–564, New York, NY, USA. ACM.

Matrix, S. (2014). The Netflix effect: Teens, binge watching, and on-demand digital media trends. Jeunesse: Young People, Texts, Cultures, 6(1):119–138.

McNally, J. and Harrington, B. (2017). 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, pages 31–39, New York, NY, USA. ACM.

Meyer, D. E. and Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. basic mechanisms. Psychological Review, 104(1):3.

Monk, C. A., Boehm-Davis, D. A., and Trafton, J. G. (2002). The attentional costs of interrupting task performance at various stages. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 46, pages 1824–1828. SAGE Publications.

Monsell, S. (2003). Task switching. Trends in Cognitive Sciences, 7(3):134–140.

Murray, J., Goldenberg, S., Agarwal, K., Chakravorty, T., Cutrell, J., Doris-Down, A., and Kothandaraman, H. (2012). Story-map: Ipad companion for long form TV narratives. In Proceedings of the 10th European Conference on Interactive TV and Video, EuroITV '12, pages 223–226, New York, NY, USA. ACM.

Murray, J. H. (2017). Hamlet on the Holodeck: The Future of Narrative in Cyberspace. The Free Press.

Nandakumar, A. and Murray, J. (2014). Companion apps for long arc TV series: Supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations. In Proceedings of the 2014 ACM

International Conference on Interactive Experiences for TV and Online Video, TVX '14, pages 3–10, New York, NY, USA. ACM.

Neate, T., Jones, M., and Evans, M. (2015). Mediating attention for second screen companion content. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '15, pages 3103–3106, New York, NY, USA. ACM.

Neate, T., Jones, M., and Evans, M. (2017). Cross-device media: A review of second screening and multi-device television. Personal and Ubiquitous Computing, 21(2):391–405.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. Journal of Personality and Social Psychology, 28(1):28.

Nielsen Holdings, N. (2014). From live to 24/7: Extending twitter TV engagement beyond the first airing. http://www.nielsen.com/us/en/insights/news/2014/from-live-to-24-7-extending-twitter-tv-engagement-beyond-the-live-airing.html.

Nogueira, J., Guardalben, L., Cardoso, B., and Sargento, S. (2017). Catch-up TV analytics: Statistical characterization and consumption patterns identification on a production service. Multimedia Systems, 23(5):563–581.

Obrist, M., Moser, C., Alliez, D., Holocher, T., and Tscheligi, M. (2009). Connecting tv & pc: An in-situ field evaluation of an unified electronic program guide concept. In Proceedings of the 7th European Conference on Interactive TV and Video, EuroITV '09, pages 91–100, New York, NY, USA. ACM.

O'Conaill, B. and Frohlich, D. (1995). Timespace in the workplace: Dealing with interruptions. In Conference Companion on Human Factors in Computing Systems, CHI '95, pages 262–263, New York, NY, USA. ACM.

Ofcom (2013). The communications market report 2013. http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr13/2013_UK_CMR.pdf.

Ofcom (2014). The communications market report 2014. http:

//stakeholders.ofcom.org.uk/binaries/research/cmr/cmr14/2014_UK_CMR.pdf.

Ofcom (2015a). Beyond broadcast consumer research 2015. http://www.ofcom.org.uk/__data/assets/pdf_file/0031/76990/beyond_broadcast_uk_datatables.pdf.

Ofcom (2015b). The communications market report 2015. http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr15/CMR_UK_2015.pdf.

Ofcom (2016a). The communications market report 2016. http://www.ofcom.org.uk/__data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf.

Ofcom (2016b). Linear vs. non-linear viewing: A qualitative investigation exploring viewers' behaviour and attitudes towards using different TV platforms and services providers. http://www.ofcom.org.uk/research-and-data/tv-radio-and-on-demand/tv-research/linear-vs-non-linear-viewing.

Ofcom (2017). The communications market report 2017. http://www.ofcom.org.uk/__data/assets/pdf_file/0017/105074/cmr-2017-uk.pdf.

Office for National Statistics (2017). Internet access - households and individuals: 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2017.

O'Hara, K., Mitchell, A. S., and Vorbau, A. (2007). Consuming video on mobile devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07, pages 857–866, New York, NY, USA. ACM.

Ophir, E., Nass, C., and Wagner, A. D. (2009). Cognitive control in media multitaskers. Proceedings of the National Academy of Sciences, 106(37):15583–15587.

Osborne, J. W. and Costello, A. B. (2009). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Pan-Pacific Management Review, 12(2):131–146.

Oulasvirta, A., Rattenbury, T., Ma, L., and Raita, E. (2012). Habits make smartphone use more pervasive. Personal and Ubiquitous Computing, 16(1):105–114.

Pea, R., Nass, C., Meheula, L., Rance, M., Kumar, A., Bamford, H., Nass, M., Simha, A., Stillerman, B., Yang, S., et al. (2012). Media use, face-to-face communication, media multitasking, and social well-being among 8-to 12-year-old girls. Developmental Psychology, 48(2):327.

Pielot, M., Church, K., and de Oliveira, R. (2014). An in-situ study of mobile phone notifications. In Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services, MobileHCI '14, pages 233–242, New York, NY, USA. ACM.

Pielot, M., Dingler, T., Pedro, J. S., and Oliver, N. (2015). When attention is not scarce - detecting boredom from mobile phone usage. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, pages 825–836, New York, NY, USA. ACM.

Pielot, M. and Rello, L. (2015). The do not disturb challenge: A day without notifications. In CHI '15 Extended Abstracts on Human Factors in Computing Systems, CHI EA '15, pages 1761–1766, New York, NY, USA. ACM.

Pittman, M. and Sheehan, K. (2015). Sprinting a media marathon: Uses and gratifications of binge-watching television through Netflix. First Monday, 20(10).

Pizza, S., Brown, B., McMillan, D., and Lampinen, A. (2016). Smartwatch in vivo. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '16, pages 5456–5469, New York, NY, USA. ACM.

Pressman, S. D., Matthews, K. A., Cohen, S., Martire, L. M., Scheier, M., Baum, A., and Schulz, R. (2009). Association of enjoyable leisure activi-

ties with psychological and physical well-being. Psychosomatic Medicine, 71(7):725.

Reeves, B., Lang, A., Kim, E. Y., and Tatar, D. (1999). The effects of screen size and message content on attention and arousal. Media Psychology, 1(1):49–67.

Rideout, V. J., Foehr, U. G., and Roberts, D. F. (2010). Generation m ^{2}: Media in the lives of 8-to 18-year-olds. Henry J. Kaiser Family Foundation.

Rieman, J. (1993). The diary study: A workplace-oriented research tool to guide laboratory efforts. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI '93, pages 321–326, New York, NY, USA. ACM.

Roberts, D. F. and Foehr, U. G. (2008). Trends in media use. The Future of Children, 18(1):11–37.

Rogers, R. D. and Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124(2):207.

Rooksby, J., Rost, M., Morrison, A., Bell, M., and Chalmers, M. (2014). Practices of parallel media: Using mobile devices when watching television. Designing with Users for Domestic Environments: Methods, Challenges and Lessons Learned. Workshop at CSCW '14.

Rooksby, J., Smith, T. E., Morrison, A., Rost, M., and Chalmers, M. (2015). Configuring attention in the multiscreen living room. In Proceedings of the 14th European Conference on Computer Supported Cooperative Work, ECSCW '15, pages 243–261. Springer.

Rubin, A. M. (1981). An examination of television viewing motivations. Communication Research, 8(2):141–165.

Rubin, A. M. (1983). Television uses and gratifications: The interactions of viewing patterns and motivations. Journal of Broadcasting, 27(1):37–51.

Sahami Shirazi, A., Henze, N., Dingler, T., Pielot, M., Weber, D., and Schmidt, A. (2014). Large-scale assessment of mobile notifications. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, pages 3055–3064, New York, NY, USA. ACM.

Salvucci, D. D. and Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. Psychological Review, 115(1):101.

Salvucci, D. D., Taatgen, N. A., and Borst, J. P. (2009). Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09, pages 1819–1828, New York, NY, USA. ACM.

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., and Watson, J. M. (2013). Who multi-tasks and why? multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. PlOS One, 8(1):e54402.

Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. Nature Reviews Neuroscience, 6(4):332–339.

Sanders, T. and Cairns, P. (2010). Time perception, immersion and music in videogames. In Proceedings of the 24th BCS Interaction Specialist Group Conference, BCS '10, pages 160–167, Swinton, UK. British Computer Society.

Saxbe, D., Graesch, A., and Alvik, M. (2011). Television as a social or solo activity: Understanding families' everyday television viewing patterns. Communication Research Reports, 28(2):180–189.

Schirra, S., Sun, H., and Bentley, F. (2014). Together alone: Motivations for live-tweeting a television series. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, pages 2441–2450, New York, NY, USA. ACM.

Schubert, T. and Crusius, J. (2002). Five theses on the book problem: Presence in books, film and VR. In Proceedings of the Fifth International Workshop on Presence, pages 53–59. Universidad Fernando Pessoa. Porto,, Portugal.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments, 10(3):266–281.

Schwan, S., Garsoffky, B., and Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activitiy sequences? Memory & Cognition, 28(2):214–223.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. Presence: Teleoperators & Virtual Environments, 1(1):120–126.

Sherry, J. L. (2004). Flow and media enjoyment. Communication theory, 14(4):328–347.

Shokrpour, A. and Darnell, M. J. (2017). How people multitask while watching tv. In Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video, TVX '17, pages 11–19, New York, NY, USA. ACM.

Silva, P., Amer, Y., Tsikerdanos, W., Shedd, J., Restrepo, I., and Murray, J. (2015). A Game of Thrones companion: Orienting viewers to complex storyworlds via synchronized visualizations. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, TVX '15, pages 167–172, New York, NY, USA. ACM.

Slater, M. (2003). A note on presence terminology. Presence Connect, 3(3):1–5.

Stawarz, K., Cox, A. L., Bird, J., and Benedyk, R. (2013). "I'd sit at home and do work emails": How tablets affect the work-life balance of office workers. In CHI '13 Extended Abstracts on Human Factors in Computing Systems, CHI EA '13, pages 1383–1388, New York, NY, USA. ACM.

Sweetser, P. and Wyeth, P. (2005). Gameflow: A model for evaluating player enjoyment in games. Computers in Entertainment, 3(3):3–3.

Tamborini, R. and Skalski, P. (2006). The role of presence in the experience of electronic games. Playing Video Games: Motives, Responses, and Consequences, pages 225–240.

Taylor, A. and Harper, R. (2003). Switching off to switch on. In Harper, R., editor, Inside the Smart Home, chapter 7, pages 115–126. Springer London, London.

The Nielsen Company (2016a). On-demand demographics: VOD viewing

across generations. http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2016-reports/global-video-on-demand-report-mar-2016.pdf.

The Nielsen Company (2016b). Video on demand. http://www.nielsen.com/uk/en/insights/news/2016/on-demand-demographics-vod-viewing-across-generations.html.

Thompson, M., Nordin, A. I., and Cairns, P. (2012). Effect of touch-screen size on game immersion. In Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers, BCS-HCI '12, pages 280–285, Swinton, UK. British Computer Society.

Troscianko, T., Meese, T. S., and Hinde, S. (2012). Perception while watching movies: Effects of physical screen size and scene type. i-Perception, 3(7):414–425.

Trouleau, W., Ashkan, A., Ding, W., and Eriksson, B. (2016). Just one more: Modeling binge watching behavior. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, pages 1215–1224, New York, NY, USA. ACM.

van Turnhout, K., Bennis, A., Craenmehr, S., Holwerda, R., Jacobs, M., Niels, R., Zaad, L., Hoppenbrouwers, S., Lenior, D., and Bakker, R. (2014). Design patterns for mixed-method research in HCI. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, NordiCHI '14, pages 361–370, New York, NY, USA. ACM.

Vanattenhoven, J. and Geerts, D. (2012). Second-screen use in the home: An ethnographic study. In Proceedings 3rd International Workshop on Future Television, EuroITV, EuroITV '12, page 12.

Vanattenhoven, J. and Geerts, D. (2015). Broadcast, video-on-demand, and other ways to watch television content: A household perspective. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, TVX '15, pages 73–82, New York, NY, USA. ACM.

Voorveld, H. A. M. and van der Goot, M. (2013). Age differences in media

multitasking: A diary study. Journal of Broadcasting & Electronic Media, 57(3):392–408.

Voorveld, H. A. M. and Viswanathan, V. (2014). An observational study on how situational factors influence media multitasking with TV: The role of genres, dayparts, and social viewing. Media Psychology, 0(0):1–28.

Walton-Pattison, E., Dombrowski, S. U., and Presseau, J. (2016). 'Just one more episode': Frequency and theoretical correlates of television binge watching. Journal of Health Psychology, 23(1):17–24.

Wang, Z. and Tchernev, J. M. (2012). The "myth" of media multitasking: Reciprocal dynamics of media multitasking, personal needs, and gratifications. Journal of Communication, 62(3):493–513.

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., and Groner, R. (2008). Playing online games against computer-vs. human-controlled opponents: Effects on presence, flow, and enjoyment. Computers in Human Behavior, 24(5):2274–2291.

Wiese, J., Saponas, T. S., and Brush, A. (2013). Phoneprioception: Enabling mobile phones to infer where they are kept. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 2157–2166. ACM.

Wiradhany, W. and Nieuwenstein, M. R. (2017). Cognitive control in media multitaskers: Two replication studies and a meta-analysis. Attention, Perception, & Psychophysics, 79(8):2620–2641.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and virtual environments, 7(3):225–240.

Wortwein, T., Morency, L.-P., and Scherer, S. (2015). Automatic assessment and analysis of public speaking anxiety: A virtual audience case study. In Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, pages 187–193. IEEE.

Yee, N. (2006). Motivations for play in online games. CyberPsychology & Behavior, 9(6):772–775.

Yong, A. G. and Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in Quantitative Methods for Psychology, 9(2):79–94.

Zacks, J. M., Tversky, B., and Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. Journal of Experimental Psychology: General, 130(1):29.

Zijlstra, F. R., Roe, R. A., Leonora, A. B., and Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. Journal of Occupational and Organizational Psychology, 72(2):163–185.