A Review of Collocated Multi-User TV

Examining the changing role of the TV in the multi-viewer, multi-screen home

Mark McGill · John H. Williamson · Stephen A. Brewster

Abstract The TV plays a central role in our homes, bringing friends and family together by providing a common reference point for live, on-demand and interactive media. However the capabilities of the modern day TV poorly reflect this social, multi-user context. The TV is in the process of being supplemented, and in some cases replaced, by other devices and displays (e.g. phones, tablets) in what is termed "multi-screening". This allows users to engage in separate and private activities alongside TV usage. The result of this is that users are "together alone" inhabiting private digital bubbles, their activity and interactions cut off from those around them. This paper outlines why facilitating multi-user use of the TV is important within the context of existing multi-screen usage, and reviews how we might go about designing TVs that better support the social contexts in which they reside. Firstly we discuss how we interact with and share use of the TV, before examining how our TV usage has been augmented by multi-screening. We then review designs for multi-user TV: from shared use TV interfaces, to using the TV as a medium by which multi-screen activity can be shared, to TVs capable of providing multiple independent physical views, allowing for both private and collaborative activity on a single shared display. Finally we discuss new technologies that might have a part to play in determining the future role of the TV. This paper demonstrates that, contrary to existing usage, the TV can be of increasing relevance in the multi-user, multi-screen era.

Mark McGill †
E-mail: m.mcgill.1@research.gla.ac.uk

John H. Williamson ‡
E-mail: jhw@dcs.gla.ac.uk

Stephen A. Brewster †
E-mail: stephen.brewster@glasgow.ac.uk
† Glasgow Interactive Systems Group
‡ Inference, Dynamics, and Interaction
University of Glasgow, Scotland, UK

1 Introduction

The TV has been the dominant means of audio-visual media consumption in the home for decades, supporting shared experiences and attracting the gaze and attention of those nearby. However, in recent years, this dominance has been eroded by the advent of "multi-screening", whereby viewers utilize two or more screens or devices at the same time. For example, in Australia 74% of the population that have internet connectivity have dual-screened (meaning they used two screens simultaneously, e.g. using a TV and a phone together), whilst 26% had triple-screened (meaning they typically utilized a combination of TV, phone and tablet / laptop) [20] (hereafter n denotes sample size; $n = 4980$). This transition toward multi-screen usage occurred because the technology and interface of the TV could not keep pace with the demands of users.

Prior historic interactive TV research, conducted at the dawn of the smartphone era [6], organized interactive TV into three areas: editing, sharing, and controlling content. However new technologies (e.g. smartphones and tablets) and behaviours (e.g. multi-screening) have changed our usage of the TV, its role in our homes, and blurred the lines between these distinctions. Whereas previously users typically consumed the same media content together through the shared reference point of the TV, the advent of digital media has brought about the means for independent on-demand consumption and activity through phones and tablets. These devices have supplemented, and in some contexts replaced, the TV viewing experience for some users. The humble TV has failed to keep pace with these changes, having faltered both in terms of facilitating shared use, and private / inde-

pendent activity. In contrast, semi-private personal devices offer users the capability to operate entirely independently of those around them, and escape the restriction of having to rely on a single medium for consumption. Furthermore, multi-screen devices enabled new multi-tasking behaviours e.g. interacting with or communicating about TV content ("media meshing"), or engaging with content unrelated to the TV content ("media stacking") [42].

Multi-screen usage has had significant consequences for users. These devices are often inferior to the TV in some important respects e.g. physical display area, casual accessibility to others, and socialization (with users in their own private "digital bubble" [23], together but alone). Moreover, not every user in the room may have a secondary device, or wish to use one instead of the TV; this leads to what one survey termed "digital divorce" [46] ($n = 1423$), whereby 24% of polled couples resorted to going into different rooms in order to watch TV separately.

Despite this, the TV remains a dominant feature of our homes. It is typically the largest display in the room, and the most accessible in terms of gaze angle (based on its position in a room). Furthermore, it is regarded as the most immersive consumer display [17], although VR headsets may in time supplant the TV in this respect. As such, in this paper we examine the potential for facilitating multi-user use of the TV. We demonstrate that through utilizing the shared reference point of the TV, we can support socialization, collaboration, and awareness of others' activity. We do so by reviewing literature from CSCW (where previous work in large projection displays and awareness has become increasingly relevant as the TV has converged upon a similar size, fidelity, and accessibility in the connected home), HCI (where a breadth of work has been conducted regarding how we use and interact with TV) and consumer market research (where we examine the current role the TV has in our multi-screen homes).

We suggest the TV has the capability to negate the isolating effects of multi-screen usage, through both making this multi-screen activity more accessible to others, and diminishing the necessity to resort to it. We discuss the social context of the TV and examine existing issues regarding multi-user use in standard smart TVs (Section 2), before examining multi-screen usage, the role these devices play in the TV ecosystem, and the issues they contribute to this social context (Section 3). We then describe the design space around sharing TV displays, reviewing systems and technology which address the issues discussed, expanding the capabilities of TVs to facilitate increased shareability, independent and collaborative use, and awareness of multi-screen activity (Section 4). Finally, we draw conclusions from this review regarding both our capability to design TVs that support multi-user use, what impact this might have on the social experience of future living-rooms, and which fu-

ture technologies have the potential to achieve a similar impact in terms of changing our relationship with the TV (Section 5 and Section 6).

2 TV Displays In The Home

2.1 The Importance Of TV: Who do we watch TV with, and why?

The TV is a central component of home life: in the UK alone there are ~52.2 million [42] TVs, equating to 2.34s TVs per home on average [58]. The TV offers a large, high-resolution, gaze-accessible and immersive view of media content, and is often found in both shared social spaces (e.g. the living room) and private spaces (e.g. the bedroom). Of interest to us is the former: TVs which inhabit shared social spaces, often used or attended to by more than one person, as it is these displays that have the ability to significantly affect users' capabilities to interact with, be aware of, and collaborate with each other.

The social groups that utilize these displays tend to be intimacy groups, meaning family and friends. A study by Thinkbox [53] (based on BARB data[1] for 5100 homes, $n \approx 11500$) found that "52% of our live viewing (including single households) is shared, and time-shifted viewing is even higher at 56%", with "most shared viewing [conducted] with one other person", whilst a report by Ofcom [42] suggested that "people are still coming together to watch TV in the living room - 91% of UK adults view TV on the main set each week, up from 88% in 2002" (also based on BARB data). Indeed, this report emphasized the importance of the living-room TV by stating that people were "increasingly reverting to having just one TV in their household - 41% of households in 2012 compared to 35% in 2002", with only 52% 5-15 year olds having a TV in their bedroom, compared to 69% in 2007.

The living-room TV is an important part of our homes, with 63% of users polled by Thinkbox [54] ($n = 802$) stating that the television is central to their relationship with the living-room. Why are people drawn toward using the shared living room TV? Whilst there are likely a number of contributing factors to this (e.g. availability of set-top box content), two stand out in our review of consumer TV market research: the quality and size of the display and its social context. With respect to the quality of the display, a 2012 Ofcom report [41] found that:

> "In the past 10 years we have seen the development of widescreen television, HD television, screens getting flatter and very importantly screens are getting

---

bigger... What this is doing is actually bringing people back into the living room and television is taking on a new role as a family experience whereas 10 years ago, in the early 2000s, we were seeing kids, different members of the family watching different television shows in different rooms using different sets." [39].

With respect to the social context, this report found that 52% of the "individualistic" 16-24 year olds watched TV with the purpose of experiencing it with family or friends, whilst 31% watched for "a bit of company". Similarly, a study by Deloitte [9] ($n = 4006$) found that for younger age groups, watching TV together was more enjoyable than watching on their own, with approximately 60% of 16-18 year olds and just over half of 19-24 year olds holding this view. Approximately half of 16-44 year olds agreed that "watching TV is a good way of bringing the family together". This is of note because it is often readily assumed that younger generations are freely abandoning TV usage for alternate displays and devices, however this view is not entirely accurate; the TV remains an important social binding agent in households.

2.2 "Pass me the remote": Problems with sharing use of the TV

Given this social context of usage, we might reasonably expect that TV designs have been refined to support the multi-user settings they inhabit. However, this is not readily the case, with problems regarding multi-user interaction and the changing role of the TV with respect to both collaborative and personal / private usage.

2.2.1 Many Inputs, One Display

Traditionally, our management of the TV was based on social conventions developed over decades of use. Interaction with the TV has iterated upon a single device that is now considered a *de facto* standard: the remote control. It is a device of ubiquity and a universally accepted means of controlling what is displayed. With this ubiquity comes a host of associated management behaviours through which social use can be accommodated: it can be passed, taken, shared, relinquished, hidden, denied.

However these social conventions are in a process of change, with new interaction techniques relying on alternate input mechanisms and modalities allowing anyone in the room to exert control. For example it has become commonplace for modern SmartTVs to bundle multiple remotes (e.g. a standard button remote and a touchpad or gestural remote). Many Smart TVs can also be controlled by apps from any mobile device in the room, whilst consumer TVs (e.g.

Samsung Smart TVs[2]), set-top boxes (e.g. the Xbox One[3], and dongles (e.g. Amazon Fire[4]) often provide voice and / or gesture controls. Users have been shown to be receptive to adopting these newly utilized input modalities, depending on situational and application-specific factors [35]. In the case of voice user interfaces (VUIs) consumer adoption of voice activated systems such as Amazon Echo, Apple Siri, and Google Voice search is likely to lead to an increase in both demand and acceptability of these types of interactions in home settings. In the case of gestural interfaces, a study examining low-energy free-hand gestures for TV tasks [60] found that in some cases gesturing was preferable to remote control usage. It is therefore likely that in some cases these modalities may become the preferred input modality for the TV, with implications for how multi-user use of TV interfaces is facilitated, given the potential for input channels that are open to use by everyone in the room.

In terms of interface design, interaction has typically remained discrete and event based, with some capacity for switching to continuous, pointer-based controls. For example, Samsung Smart TVs feature interfaces designed to support both discrete navigation and pointer-based navigation, whilst Android TV offers a similar capability[5]. In both cases, facilitating multi-user use is problematic: in discrete systems there is the issue of cursor sharing versus relying on multiple potentially visually distracting cursors. Managing events (e.g. remote control shortcuts, voice commands, gesture commands) also becomes problematic e.g. if a user initiates a transition to another view whilst another user is interacting with some element currently being displayed, what is the appropriate action to prioritise? Whilst some modalities have associated social cues that might help prevent problems of concurrent usage (e.g. voice usage and the acceptability of talking over another user), other modalities and inputs (e.g. gestural controls, remote control, apps) do not naturally have blocking mechanisms. In pointer-based systems, there is additional bandwidth of input, but also significant visual distraction due to the necessity for multiple pointers, whilst continuous input is likely to increase the effort, mental demand, and physical demand required (depending on the input modality and sensing technology in use).

2.2.2 Many Activities, One Display

There is also the question as to whether concurrent and shared-use interfaces are sufficient for the variety of activities that multiple users might engage in, and the effect that these activities might have on other users of the TV display. If two

---

[2] www.samsung.com/us/2013-smart-tv/
[3] www.xbox.com/en-US/xbox-one/entertainment
[4] www.amazon.co.uk/Amazon-CL1130-Fire-TV/
[5] play.google.com/store/apps/details?id=com.google.android.tv.remote

or more activities are to be conducted on the TV, e.g. one person viewing live TV whilst the other interacts with an Electronic Programme Guide (EPG), this necessitates dividing the TV display so that each activity has a given region of the display. This screen division is arbitrary and can be designed to suit the content being accommodated. For example, picture-in-picture suits fixed aspect ratio content as there is no unused screen area; similarly a 4-way split of the display, as seen in Samsung multi-link[6] (see Figure 1), allows for content designed for the aspect ratio of the display to be scaled down whilst maintaining this aspect ratio.

Fig. 1 Samsung Multi-link: here the screen can be divided into 4 views, allowing for multiple concurrent and independent activities to be performed, at the expense of sacrificing screen area and increasing visual distraction.

If the aspect ratios of the two activities differ, the screen can be divided in any number of arbitrary ways. For example in the case of the XBOX One interface the aspect-ratio of video content is maintained and an overlay for interactive applications appears on a vertical slice of the display. In this way interactivity is provided, at the expense of a portion of the media being viewed (see Figure 2).
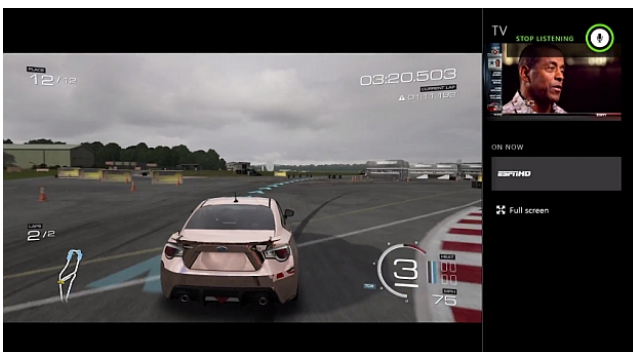
Fig. 2 XBox One Snap UI: here applications can be snapped to various parts of the display, with the primary content aspect ratio being maintained but shrunk to use a diminished area of the TV.

---

[6] www.samsung.com/global/microsite/tv/uhdtv/mobile /multi_link_screen.html

Split-screen and picture-in-picture approaches are inherently sub-optimal, compromising use of the display to accommodate multiple activities and users, through either obscuring part of one view to provide another view of poor legibility / size, wasting screen area, or compromising the aspect ratio of the content being consumed. More fundamentally, they offer no privacy considerations; checking email or using a social media application on a TV, whilst feasible with such screen division approaches, is often socially unacceptable either to the user conducting the activity who wants privacy, and to the users forced to give up part of their TV view for this potentially irrelevant activity. Thus, whilst the display can facilitate collaborative activity to an extent, independent activity is problematic and likely to be a distracting addition to the display for other users, whilst private activity is impossible in a multi-user context.

3 The Role Of Additional Screens And Devices: Autonomy And Privacy In A Shared Social Space

Personal devices circumvent many of the problems TV displays have in multi-user contexts; they guarantee the user full use of a display that remains private through social conventions but physically shareable if they so choose, a display whose interface they alone control and customize as they see fit. Because these devices are personal, they are invariably connected to personal social media and messaging accounts and offer a semi-private space for conducting activity. However their usage introduces new problems regarding shareability and social impact, as we will discuss.

3.1 Adoption And Usage: Multi-Screening Behaviours

The adoption of these personal devices, and their usage in TV-viewing contexts, is highly indicative of the importance of being able to operate independently and privately. In the UK smartphone adoption reached 61% in 2014 (up 10% since 2013), whilst tablet adoption almost doubled (to 44%) in the past year [43]. This is a global phenomenon e.g. in Australia tablet adoption was 42% in 2014, up 10% from 2013 [38] (based on OzTAM, 3500 homes[7]). These are devices that are widely available and have had a significant impact on the TV-viewing experience, through their use alongside the TV in what is known as multi-screening. This refers to usage of a mobile internet-connected device at the same time as television viewing [52], and it is in this way that users attempt to get the best of both worlds in the home: utilizing the TV for immersive, shared entertainment experiences, whilst utilizing smaller displays for personal and private experiences. A report by Google [13] ($n = 1611$) found that:

---

[7] http://www.oztam.com.au/AboutOzTAM.aspx

"TV no longer commands our undivided attention, with 77% of viewers watching TV with another device in hand. In many cases people search on their devices, inspired by what they see on TV."

This multi-screen usage has typically been categorised into sequential and simultaneous use [13, 34] (see Figure 3). Sequential multi-screening (also known as shifting [34] or "quantum" [18] referring to leaps in both time and space) refers to one task or activity being transferred between devices as and when required, e.g. performing a search for an item on a laptop then continuing that search on a tablet later in the day. Of particular note in the context of TV usage, however, is simultaneous usage - that is, usage of more than one screen at the same time.
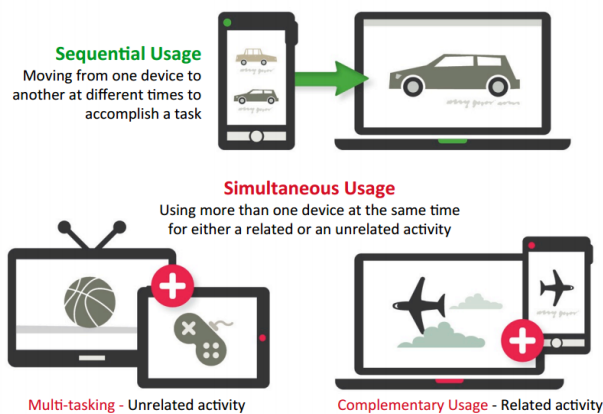
Fig. 3 A common categorisation of multi-screening behaviour [13].

Estimates vary regarding the extent of simultaneous usage. A report by Millward Brown [34] ($n > 12000$) suggested this constitutes 35% of the time, whilst an Ericsson Consumerlab study [12] ($n = 15000$) stated that 75% of users polled had at some point engaged in multi-screen multi-tasking. This is clearly a highly prevalent behaviour in the home. For example, a Nielsen study of Australian multi-screen usage [20] ($n = 4980$) suggested that 74% of online Australians dual-screened, whilst 26% had triple-screened, most often using laptops / netbooks (63%) smartphones (50%) and tablets (36%). This simultaneous behaviour typically relied on a combination of smartphone and TV displays, with 81% of users polled using this combination every day, whilst 66% used laptops in conjunction with TV. One user remarked that:

"My phone... I consider it my personal device, my go-to device. It's close to me, if I need that quick, precise feedback" [13]

This simultaneous usage has been broadly categorised in various publications; Microsoft defined different pathways

for multi-screen usage such as content grazing, investigative spider-webbing and social spider-webbing [18]. Millward Brown discussed it in terms of meshing (simultaneous use for related content, which was employed 38% of the time that users were engaging in simultaneous usage) and stacking (simultaneous use for unrelated content, employed 61% of the time). Of note here is that this usage is not always driven by the need to engage with interactive media or applications:

"The online research ... shows nearly one in five men (18 per cent) and over one in ten women (11 per cent) polled have watched two live TV programmes simultaneously within the past year whilst three per cent of respondents said they had watched three programmes at the same time" [12]

Resorting to personal devices for video media consumption instead of relying on the (in some ways) superior shared TV is a common theme in recent multi-screening surveys, and just as prevalent as relying on multi-screening for interactive applications and internet usage. For example, a UK survey by Thinkbox [54] ($n = 802$) found that 56% of the sample had watched TV on screens other than the TV set whilst in the living room, whilst an Australian survey by Nielsen [38] (based on OzTAM, 3500 homes) found that 74% of Australians aged 16 and upwards had watched TV and used the Internet simultaneously - up 14% since 2011.

3.2 Impact On TV Usage

The question then is: given the prevalence of multi-screen usage, the variety of multi-screen combinations in use and the vast breadth of user behaviours and activities observed, what effect does this have on the usage of the TV and the ability to socialize and interact in the living room? For a start, multi-screening reinforces use of the TV, with "viewers more likely to stay in front of the TV for longer (64% of multi-screeners viewed for over 15 minutes per time compared to 47% of non multi-screeners)" [52] ($n = 1000$). There is now less of a need to resort to "digital divorce", where couples resorted to going into different rooms in order to watch TV separately [46]; any private and personal media activity can be undertaken in this shared public space, at the expense of the ability to potentially use the best display in this space. Indeed one study [55] suggested that "people are planning their evenings around the TV schedule more - the TV is an important social point both within the home and beyond", meaning that multi-screen usage was driving viewers back to shared TV experiences.

Device usage is not necessarily a distracting presence; viewers browse the internet as much as they talk to other people in the room [9] ($n = 4006$)(70% frequently or occasionally browsing versus 80% frequently or occasionally

talking to others in the room); distracted viewing is the norm. Moreover second-screening activities such as the use of companion applications or social media (such as Twitter hashtags) accompanying TV content have become a regular occurrence. However, device usage can impact enjoyment and focus of attention: in a UK survey two-thirds of respondents agreed with the statement "If I am really enjoying a programme, I don't really want to use another device at the same time" [9].

3.3 Interactions between TV and Additional Devices: Social Sharing And Transitions

Whilst device usage has played a significant part in enabling independent and private activity in the living room, there have also been a number of consumer technologies that have attempted to open the TV up to these devices, allowing for an element of shareability through the TV. Screen-mirroring (also screen sharing, casting, annexing) technologies such as Apple Airplay[8] or Miracast[9] are available in most new mobile devices, allowing the mirroring of screen content via a dongle in the TV, as well as driving entirely separate presentations, expanding the capability of users to share both presentational and interactive content.

Fig. 4 Example of social sharing using the Google Chromecast dongle and associated smart device apps[10]. Here a media experience is controlled by both a phone and a tablet, with the Chromecast-enabled TV having the ability to stream media directly from the internet, or from a given device.

Similarly, playlisting technologies have been incorporated into TV displays, most notable of which is the "casting" capability integrated into Google Chromecast TV dongles (as seen in Figure 4). These devices, along with appropriate software integration on multi-screen devices, allow

for shared presentational use of the TV through the playlisting of video streams, as well as limited mirroring capability much like Miracast. At a higher level, these shareable actions can be considered mechanisms by which content or activity transitions can occur; this sequential or shifted use of different displays allows for users to migrate the content being consumed or interacted with onto the most appropriate display available, for example moving a mobile gaming experience onto the TV to continue playing with the best display in the room.

Indeed there is significant scope for inferring this appropriation (or "cyber-foraging" [8]) of available displays. However the acceptability of interactions such as these is likely tied to the social makeup of the group using the TV e.g. is it acceptable for a relative outsider to the household to appropriate the TV for sharing content? Devices like Google Chromecast offer guest modes specifically designed to allow for such use cases, whereby proximate users not connected to the household network can still interact with the TV (e.g. sharing video, photos etc.) with no restrictions. As yet we do not know the extent to which variations in the familiarity / social makeup of a group might effect willingness to appropriate the TV, nor whether the TV should be made as accessible as this e.g. is proximity / geo-fencing sufficient to allow multi-screen devices access to the TV[11]?

3.4 Problems With Multi-Screening

Given the adoption and usage of multi-screen devices, it is reasonable to assume that multi-screening fulfils the majority of user needs: the living room TV remains a shared social focal point with one common media experience, whilst independent and collaborative activity is offloaded to individual personal devices of varying capabilities and sizes. However, supposing that everyone in the room has access to alternate displays to the TV (which will not always be the case), this usage presents two problems: firstly that users are together, but alone, ensconced in their own private media experiences, and secondly that users are having to resort to smaller, less immersive displays whose content is not readily accessible to others. Whilst there is a capacity for explicitly sharing content using devices, shared-use interactive content and casual awareness of non-private activity are greatly impeded. The result of this is that users are potentially cut off from a significant portion of the experiences and activities of others around them.

---

[8] www.apple.com/airplay/
[9] www.wi-fi.org/wi-fi-certified-miracast
[10] www.google.co.uk/chrome/devices/chromecast

[11] Victims of the Chromecast "RickMote" controller might suggest not www.wired.com/2014/07/rickroll-innocent-televisions-with-this-google-chromecast-hack

3.4.1 Casual Awareness & Shareability: Together Alone

The private "digital bubble" [23] of device usage has long been discussed as raising a problematic barrier to socialization and interaction, with mobile phone use in particular having significant anti-social connotations [57]. Efforts have been made to penetrate this bubble, for example Lucero *et al* [27] proposed mobile collocated interactions, whereby users would "take an offline break together", pooling their device resources toward "shared-multiuser experiences". They aimed to facilitate joint attention, whilst enforcing a break from online socialization, appropriating mobile device displays in order to pass photos around a table. This emphasis on shareability and joint attention is important as it underlines how collocated interactions are made to be more effective, through the ability to share awareness, and take part in shared activities.

However, mobile devices are not the most shareable displays in the room. McGill *et al.* [31] demonstrated that physically sharing device views was inferior to utilizing the TV in terms of sharing activity with others and thus collaborating effectively, whilst Terrenghi *et al* [51] discussed scale of displays relative to users' visual angle and distance, noting that the scale of the display must match the social interaction space. In using multi-screen devices, we erect barriers to socialization and our ability to be casually aware of, and perhaps join in with, the activity of others is impeded.

These barriers have been reduced by technologies such as Miracast and Chromecast, however these approaches are sub-optimal. The adoption and usage of of Miracast and other screen mirroring technologies is low. An NPD survey [40] ($n = 2600$) of smartphone users found a 40% awareness of the existence of screen-mirroring capabilities, with only 7% having ever used such features. Of these individuals, 75% had used this capability for mirroring videos, whilst approximately 50% had mirrored photos. The study stated that:

> "Bringing sharing experiences to a larger consumer base will require simplifying hardware requirements [and] amplifying the value of being able to share content across screens"

Although screen sharing is a low-cost way of sharing content between multiple users, it also has some notable limitations. In mirroring screen content, elements of the device interface that are not relevant, or not being attended to, may also be shared. Additionally, screen-mirroring restricts the ability for multiple users to concurrently interact, as it is essentially multi-screen single-interface groupware. It is feasible that these reasons have contributed to the lack of adoption. In the case of Chromecast, as with Miracast, there is little facility for multi-user interaction, or collaborative activity; the TV is treated as a dumb terminal, presenting playlisted media aggregated across the group. In both cases, there remains user activity that is isolated on devices with no technological facilities for casual awareness. Sharing must be explicitly managed regardless of if the activity needs to be private or not, and no more than one device or piece of media can be shared at a time.

3.4.2 Size & Immersion

When we watch a TV programme or movie on a TV, we do so because the TV offers the most immersive and shareable experience: the TV is the largest display in the room, often has additional capabilities for improving immersion (e.g. 3D rendering), is accessible from a variety of gaze-angles, and presents a shared audio experience often employing positional audio. Indeed larger displays have been shown to increase immersion, with a study by Hou *et al.* [17] finding that large displays resulted in a greater sense of self-presence than smaller displays. A user choosing to instead watch live TV on an alternate device such as a tablet or phone is inherently sacrificing many of the benefits of larger displays in order to be able to indulge in a personal media experience other than the one currently presented on the TV. This is a trade-off that is inherently less than ideal; if a TV and its associated audio system had the capability to allow for multiple independent viewers, would the 56% of users that were found to have watched TV on screens other than the TV set whilst in the living room still resort to this behaviour? While there would be justifications for this (for example if content were inappropriate or private), it is reasonable to assume that shared utilization of the TV would be preferable if possible.

4 Designing A Multi-User TV

There are a number of ways in which TV media-system user interfaces can be designed to accommodate multi-user use e.g. multi-pointer / cursor interfaces, split-screen / screen division interfaces, or offloading interaction onto other devices or screens. Additionally there is a need to design for the realities of modern homes where multi-screen usage has become the norm. Although we may be able to expand the capabilities of the TV display significantly, there is a culture of personal device usage which is growing year by year. Each of these approaches has particular problems and trade-offs, as we have discussed thus far; in designing a multi-user TV we must rectify specific problems regarding:

– **Interaction:** The integration of new sensing technologies opens up the capability for new input modalities (such as gesture and voice) and mechanisms (such as using smartphones, wearables such as smart watches etc.).

How do we design TVs that support interaction from every corner of the room? And to what extent do we strive to retain existing behaviours for managing use of the TV that have evolved around the existing, dominant input mechanism of the remote control? For example parents might once have taken the remote control away from a child; how do we incorporate such behaviours in a multi-user display where there is no physical token of control to manage? And do we support concurrent multi-user interaction, or do we mediate between users?

– **Facilitating collaborative and independent activity:** The reliance on multi-screening behaviours is fueled in part due to limitations in TV display technology: currently, the TV cannot support private independent use. Although screen-division approaches allow for a degree of independence, they increase visual distraction and sacrifice display area and immersion. Similarly, multi-screen approaches sacrifice immersion through the use of smaller displays, and erect isolating barriers between users, with awareness of the activities and experiences of others greatly impeded. Given that multi-screen usage is now well established, this raises two questions: can we decrease the isolating effects of multi-screening and provide some form of shared awareness using the shared focal point of the TV display, and can we tackle one of the root causes of multi-screening, namely that the TV cannot support private independent activity?

With respect to these problems we examine relevant papers proposing ways in which the capabilities of TVs can be significantly expanded in order to diminish and potentially solve these issues. Firstly, we discuss mediated and concurrent interaction techniques as potential solutions to the "Many Inputs, One Display" interaction problem. Secondly, we examine an extension of mediated interaction which provides mediated use of screen-mirroring, enhancing awareness of multi-screening activity through use of the TV, both demonstrating and decreasing the isolating effects of device usage. Thirdly, we examine TVs that can support semi-private independent activity, personal immersive experiences, and user controlled awareness of other activities in the room through usage of TV displays capable of providing multiple independent views to those present. We additionally review shared audio technologies that can complement such displays, preventing audio conflicts, helping users to manage shared audio spaces, and allowing for personal and private audio spaces. These papers provide competing snapshots of how the TV can facilitate multi-user use; from shared use of the TV interface, to pushing interaction to the devices in the room and using the TV to provide awareness of this activity, to TVs that can accommodate independent multi-user use and reduce the necessity of device usage. In doing so we ask the question: what should the role and scope of the TV be in the future? Should it serve basic media interactions as it has done for decades? Should it augment and support other devices? Or can it be the default display for interaction, superseding devices?

4.1 Many Inputs, One Display: Mediated And Concurrent Interaction

Our capability to provide input to the TV has expanded rapidly in recent years; gesture, voice, and device inputs can be considered relatively common in COTS TVs providing the potential to enable anybody in the room to interact with a shared TV interface. How can we design the TV interface to support multi-user interaction? Of note are two approaches here: mediated interaction, whereby a group self-manages their use of the TV display, and concurrent interaction, where the interface is designed to support and facilitate concurrent inputs and activity from one or more members of the group.

4.1.1 Concurrent Interaction

Concurrent use interfaces are either managed (e.g. You *et al* [62] used computer vision techniques to detect users and partition and rearrange personal space on a shared display), self-managed (e.g. Tse *et al* [56] demonstrated how users were found to self-partition shared workspaces in order to achieve optimal collaboration), or achieved through the combination of the two. This is a common feature of tabletop interaction for example; LunchTable [37] integrated a multi-touch display table with a large, vertical display for rich information, allowing the sharing of content among a whole group, whilst control of the display was managed concurrently via the multi-user table.

Outwith tabletops, Single-Display Groupware [49] multi-pointer systems are perhaps the most relevant example of concurrency [4], to the extent that strategies for multi-pointer management are becoming increasingly relevant [47]. However, the multi-pointer approach is not without its flaws, requiring greater dexterity / continual adjustment when manipulating said pointer compared to typically utilized discrete interfaces, increased visual noise, and potentially decreased performance [25]. It is likely because of these reasons that modern smart TV interfaces are still designed to primarily accommodate discrete controls, typically enacted through traditional button-based remote controls and still prevalent / preferred in existing TV interfaces. For example Android TV still retains discrete cursor controls, whilst Samsung SmartTVs support both discrete and pointer-based interactions. Whilst concurrent interaction of discrete interfaces can be facilitated, many of the same problems arise e.g. increased visual noise. More fundamentally, there are common issues regardless of whether concurrent interaction

is enacted through discrete or pointer-based controls, for example how do we manage navigational events (e.g. switching to a different application on a smart TV) when more than one user might be currently interacting with the displayed interface? One argument is to resort to previously mentioned screen-division approaches, but this lacks scalability and compromises use of the full display.

4.1.2 Mediated Interaction

The alternative is to retain the TV interface as a single-user application, much like it has traditionally been except that instead of this being a limitation of the input technology (a single shared remote control), it is now an enforced limitation. In 1990, Greenberg et al. [14] demonstrated a means of sharing single-user applications through view-sharing and turn-taking, and this concept has been frequently extended and re-implemented since. For example, two decades later, Abe et al. [1] examined tolerant sharing of single-user applications amongst multiple users. The idea of adapting single-user systems represents a pragmatic approach, one that is often deployed due to some constraint preventing the redesign of the underlying single-user system. Often it is a wish to retain the mental model and learned behaviours users have developed, or an acknowledgement that systems are often targeted at the single-user model, even though there will be use cases where multi-user use is likely to occur.

There have been a number of papers proposing shared-use media systems. For example, Ballendat et al. [3] developed a system whereby a large vertical display enabled media related tasks (browsing, viewing), adapting the presentation based on the angle and proximity of the user, and pausing when the user was no longer engaged with the system. In this scheme, the user closest to the system was considered most engaged with it, thus essentially sharing the system through a hierarchy of proximity.

Pohl et al. [45] proposed that interaction could be defined by the extent to which the user was engaged in a task. They suggested that there was a set of scenarios where casual interaction might be better suited for a given task, and that determining this level of engagement (and thus which form of interaction, casual or focused / engaged) be up to the user. The system would then adapt depending on how much attention and effort the user chose to invest. They too discussed proximity, for example pointing to the fact that the bandwidth of user interactions decreases proportional to distance to the device with which the user is interacting, thus mapping engagement to proximity. However, these approaches may not be appropriate for collocated groups in shared spaces interacting with media systems. For example, the proxemic approach does not take into account the fact that proximity to a media system is dictated not by engagement, but by seating arrangement: it might be just as likely

to be fully engaged in the system, without being the closest person to said system, as being entirely disengaged from the system at close proximity, given the variety of seating arrangements in living spaces.

In contrast, approaches have been undertaken to design "seamless" interaction techniques such that, regardless of proximity, the same mechanics for interaction would be retained. Clark et al. [7] proposed a proximity-based interface that allowed users to interact with a media system both within range of touch, and at a distance, transitioning to pointing or device input when far away. Of note here was the fact that in the evaluation of this system, the proximity-based interaction was not frequently used; additionally, having the interface change depending on distance via zooming was found to be counter-intuitive.

This raises some important discussions regarding whether an interface should be adaptive within the domain of the living-room: is there enough space typically available such that the interface becomes unusable at a distance and thus needs to adapt? And how is shared use facilitated? If a group of users is currently attending to the display, with one user browsing through available media, to whom should the display be targeted? Group interaction with media systems overlaps with these techniques, but is fundamentally different in many ways. Proximity is in all likelihood rendered irrelevant in static seated contexts, whilst attentional interfaces are muddled by the fact that many users may be attending to the display, and all may intend to interact with it at some point. Additionally, attempting to adapt to attention is fraught with difficulty: if a user looks away from the screen, perhaps to talk to someone, that does not give sufficient justification that they might want their media paused: in providing interaction techniques that are low effort and seamless, both casual and engaged interaction are potentially adequately facilitated. Finally, there are also social and cultural issues: any given interaction technique may contradict societal norms (e.g. undermining the control of the head of the household) or cultural norms (e.g. a particular gesture set being inappropriate).
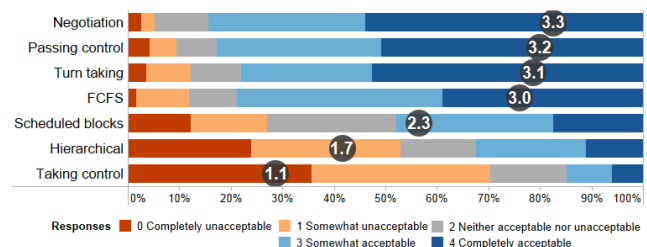
Fig. 5 Responses to the question "How acceptable do you find the following ways of controlling media systems?" from [30]. Responses were Likert-type five point scale, ranging from completely unacceptable to completely acceptable, and converted into 0-4 scale for mean acceptability (labeled in grey circles, higher is better) for relative comparisons.

McGill *et al.* [30] attempted to side-step many of these issues by developing schemes for mediating between inputs in a social group based upon the existing management behaviours of the single physically shared remote control, in what was termed "mediation of control". In a survey ($n = 156$) they found that with respect to these existing management behaviours, that control was a commodity:

"First come, first serve", "passing control around", "negotiation e.g. asking for control" , and "turn taking" were the most used strategies, with "hierarchical (an individual is typically in control)" , "scheduled blocks for sharing control of the TV", and "taking the control from whomever currently has it" falling behind. This supported the view that control of these systems is a commodity or resource in and of itself. As the person currently in control plays a large part in dictating events, if you acquire control, you might be reticent to relinquish it; societal norms of fairness may, however dictate that strategies be introduced to accommodate other's wishes and uses, hence passing control, turn taking and negotiation feature.

Based on these existing behaviours, they attempted to establish if existing behaviours for managing control were applicable in situations where the bottleneck of a single remote control was no longer relevant, if users could self-mediate control or whether everyone being capable of providing input at any time was acceptable, and what new behaviours could be build upon the concept of virtualizing control. As such, they examined 10 different control schemes in an EPG browsing task using a discrete-control interface with one shared cursor (see Figure 6), broadly categorised as either "one user in control at a time" (hereafter "one user"), and "multiple users concurrently" (hereafter "everyone"). The "one user" schemes were based on existing behaviours: **passing**, **taking** and **turn taking**. Additionally, a variant of

Fig. 6 "Mediation of control" UI [30]. Three Android devices were used as remote controls to a Windows Media Center interface (pictured right). The bottom left / right buttons changed function depending on the condition being evaluated, offering up basic controls for mediating who is in control of the TV EPG interface. Users shared discrete control of a single cursor.

passing / taking was introduced: ***lending***, essentially a hierarchical means of managing control where control could be lent out, and revoked, from an individual with authority. The **control** condition also fell into this category, being one remote control physically shared amongst participants.
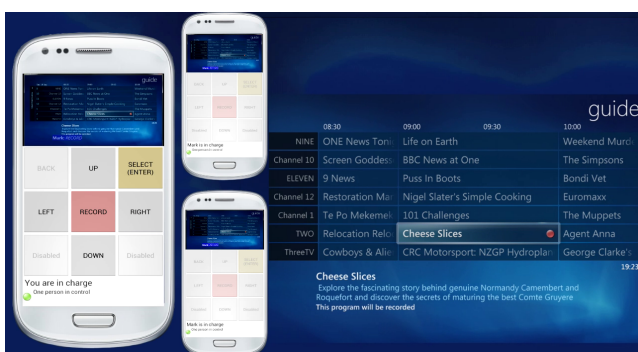
The "everyone" schemes were introduced on consideration that, if everyone could potentially be in control of a single-user interface, would an amount of self-organisation take over, thus demonstrating that system-based mediation of control was not necessary? As such, conditions were added allowing for **everyone** in control, **subsets** of control (where different group members had control of different functions, thus requiring cooperation), **hierarchy** (where one member's input would override that of the others), **plurality** (where selection decisions were based on majority votes but navigation was concurrent) and **blocking** (where members could selectively and temporarily block each other from control).

What they found was that fundamentally groups could not self-mediate control; when everyone was in control, the group could not effectively manage use of the shared interface, leading to destructive inputs and confusion; in contrast, schemes built upon existing behaviours that enforced a "one user at a time" constraint (such as passing / lending / taking control) were at worst comparable to having one physical shared remote control. This work had implications for the rise in device-based inputs for example, with Android phones using IR-transmitters allowing for exactly this kind of usage scenario to occur; in the worst case, when users wished to interact concurrently via such mechanisms, their use of the TV interface would be both frustrating and wildly ineffective.

More broadly, this work suggested that we do not need to rely on concurrent use interfaces in order to facilitate social and shared use; existing behaviours for sharing use could be re-purposed and virtualized, allowing users to retain the familiar interfaces and mental models they had grown used to whilst enabling the use of new input mechanisms and modalities. This is an important principle to consider: whilst we can facilitate multi-user use through designs for concurrency, we may lose something in the process (e.g. behaviours for managing control, such as a parent taking control away from a child), and we may introduce more complexity than is necessary or wanted.

4.2 Decreasing the isolating effects of multi-screening: Awareness & Shareability

Shared use of a TV interface only accounts for a portion of the activity in the living-room due to the uptake of multi-screening. The question then is how can we further extend the social, shared-use capabilities of the TV in order to benefit these new behaviours. McGill *et al.* [31] suggested that the TV could be used to diminish the "digital bubble" effect

of multi-screening, through the utilization of previously discussed screen-mirroring technologies. They proposed that the mediation of control behaviours discussed previously for sharing a single-user interface could be adapted to instead share use of a single-user resource in the form of the TV display, allowing users to self-manage their mirrored use of the TV through passing / taking / requesting the display from whomever was currently mirroring to it.

They hypothesized that this system would lead to increased awareness of a groups activity whilst preserving an individuals ability to operate independently and privately. Awareness within collaborative systems is a key issue and considered a necessary train for effective collaboration. It has been studied for many years, with a variety of definitions [15]. The most relevant interpretations are those of Greenberg *et al*. [16], and Schmidt [48]. Greenberg *et al*. [16] discussed the concept of workspace awareness ("one persons understanding of another person's interaction with a shared workspace") [50], specifically artefact ("what objects are they working on") and action ("what are they doing") [16, 15] awareness. Schmidt [48] referred to awareness in terms of actors; actors both monitor activities, and display activities perceived as relevant to their colleagues. The effect of this is improved situation awareness, which in turn allows for more effective collaboration [16].

McGill *et al*. evaluated a prototype shared screen mirroring system (see Figure 7), where users interacted with smart phones which provided the capability to manage mirrored use of the TV via on-screen display (OSD) buttons. They compared this shared screen mirroring system to mobile device usage only (no TV), and a control screen-mirroring setup whereby one device was permanently mirrored to the TV, representative of current usage of consumer screen mirroring. With respect to multi-screen device usage, they found that even though users were seated next to each other, physically sharing the device view occurred infrequently, underlining the "digital bubble" effect. In contrast, the conditions where the TV was utilized as a medium for sharing activity significantly improved a groups perceived ability to collaborate.

More fundamentally however, they also showed the necessity of incorporating management behaviours that reflected the social needs of the users of the TV. In the single mirrored device control condition, they found that users independence was compromised: the person whose activity was mirrored to the TV essentially dictated the collaborative experience, and diminished the individuals capability to operate independently. This was in contrast to the shared screen mirroring system, where users readily engaged in managed use of the display as a group, taking the display when they wished to present or share activity with the group, and relinquishing when they wished their activity to remain private, with no impact on their capability to operate independently.
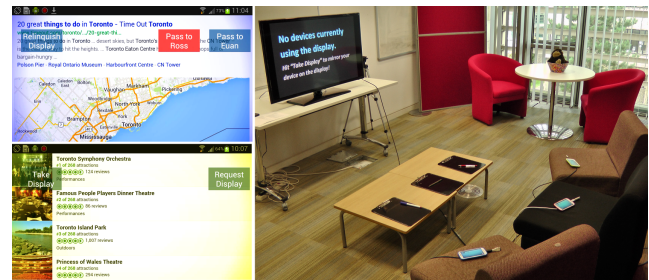
Fig. 7 Shared screen mirroring system [31]. Left: client UI (top is current "owner" of the TV, bottom is another participant). The coloured glow around the edges was unique to each device, whilst overlayed semi-transparent buttons enabled management of the shared mirroring TV. Right: Living-room-like social space used for conducting study; users were seated such that they could physically share device views if they so wished, whilst everyone had a direct view of the TV.

This work emphasized both the isolating effect of multi-screen usage, and the potentially detrimental effect of technologies that do not take into account the social context of usage. Furthermore, it provided as a counterpoint a system capable of bursting the "digital bubble" of multi-screening; in taking an existing technology (screen-mirroring) and augmenting it with some basic social management behaviours (mediation of control), users could appropriate the TV in ways which reflected the social context of usage; mirrored use of the display could be passed amongst the group as and when required, with the express purpose of utilizing the accessibility and shareability of the TV to enhance the groups ability to interact with each other. In doing so, their usage of both the TV and multi-screen devices could be enhanced significantly, and the isolating effects of multi-screen usage diminished.

There are however caveats to such usage. A follow-up paper [33] examining the effects of shared screen mirroring found that sharing the TV in this way led to increasingly fragmented viewing of any on-going TV content. Whilst isolation was decreased, the primary function of the TV was potentially compromised. In order to decrease this fragmented viewing, automatic screen-division was used, whereby when a device was to be mirrored the TV would layout the content streams being consumed appropriately, for one to three content streams, as pictured in Figure 8.

The authors additionally devised a mechanism for aiding users to manage when the display slipped into split-screen or fullscreen modes by inferring a focus of attention based on the content being consumed. For example, if a device was being mirrored alongside TV content, and that device started to play a video, the system would pause the TV content and focus on the mirrored video for so long as it played. This paper also examined an alternate approach to user-led sharing of device activity, proposing "passive screen-mirroring", whereby devices within a geo-fenced area (e.g. in the same room) could be made available to be viewed by others on-

Fig. 8 The screen mirroring system from [33]. Devices featured a permanent on-screen UI for managing what was being viewed, whilst the TV automatically managed the layout accordingly, presenting TV and mirrored device content simultaneously if requested.

demand. Whilst the results were mixed, with passive screen-mirroring proving problematic in terms of acceptability due to privacy concerns, it is important to note that awareness does not have to be dictated by those undertaking the activity, but can instead be determined by the observing user, if the appropriate mechanisms are in place to determine when such activity should be made available to be viewed. We will revisit this idea of allowing users to personally vary their engagement with others' activity in the next section.

4.3 Facilitating Collaborative & Independent, Semi-Private Activity On The TV

Whilst shared screen mirroring is effective at overcoming the accessibility / shareability barrier of multi-screen device usage, there is an argument that it is addressing the symptom, not the cause: multi-screening has arisen because of a need for independent experiences and activity, a need the TV currently does not adequately fulfil. Whilst we can present multi-screen activity on the TV display, this would impact traditional usage of the TV and is thus not appropriate in many circumstances. Similarly, we could present this multi-screen activity directly to the devices in use, but this compromises device usage, and denies users of the capability to have a shared, heads-up focal point upon which they can selectively converge.

In order to diminish the role of multi-screen device usage, McGill *et al.* [32] proposed the leveraging of multi-view display technology to allow for TVs that could provide multiple, independent views to whomever was attending to the display. Multi-view displays are singular displays that are capable of providing two or more independent views to one or more users, providing the capability to allow for both independent operation and collaboration, with a shared focus of attention throughout. There are a number of technologies that are capable of achieving this aim [10] e.g. Lenticular displays, using sheets of lenticular lenses atop a standard LCD screen allowing for different views based on gaze an-

gle; Parallax-barrier or Masked Displays, employing masks, be they singular portholes [22] or a series of holes or slits [44], in order to control what subpixels are viewed at a given angle; or even modified Twisted Nematic Displays e.g. Kim *et al.* [21]. Today's state of the art multi-view technology currently is that of high refresh rate, low pixel-persistence[12] displays combined with active shutter glasses which can selectively reveal or mask frames as they are displayed. These displays offer platforms for developing gaze-angle agnostic multi-view interfaces , with low amounts of crosstalk[13] whilst retaining high frame rates and image fidelity, albeit at the expense of brightness due to the amount of time the glasses are in their "shuttered" state.

Multi-view displays already exist in consumer off-the-shelf (COTS) TVs in a limited fashion: consumer 3DTVs rely on active shutter techniques in order to convey stereoscopic left and right images to users, and typically also have options for two-view content consumption and gaming e.g. LG "dual-play"[14], Samsung "Multi-View" displays[15] and Sony "Simulview"[16] displays supporting two-view usage, with switching typically managed via a physical switch on the 3D glasses.

Multi-view displays can be used by solitary users or groups, and have a number of advantages over comparable smart TVs in each case, however more-so in multi-user usage. For example, in single-user scenarios they have been used to present different aspects of an interface based on view position, allowing users to move their head in order to peek at a menu [28]. In multi-user contexts, they have been used to support independent and collaborative activity on table-tops e.g. the Permulin[26] which supported two users sharing a 120Hz two-view display, or Permulin's precursor[2], and independent views in groups such as in the case of C1x6[24] which employed multiple projectors in order to achieve a 12-view 360Hz display allowing for 6 stereoscopic views.

McGill *et al.* leveraged this technology in its active shutter form to build a TV capable of providing two users with completely independent physical views of a set of virtual Android device views (see Figure 9). They established that the primary issue with such displays would be in designing interactions that allowed for transitions between collaborative and independent activity, and providing awareness of on-going but relevant activity on other views. An iterative design process arrived at a TV which allowed users to tran-

---

[12] Pixel persistence: the time it takes a pixel to transition from it's previous state to it's current state

[13] Crosstalk: the extent to which one image is retained into the subsequent image. For example, given a two-view multi-view display, where one view is a car, and the other a boat, crosstalk would be manifested as the boat being visible (ranging from a faint outline to wholly superimposed) in the car view, and vice-versa.

[14] www.lg.com/us/tv-audio-video/discoverlgtvs/dualplay

[15] www.samsung.com/us/video/tvs/KN55S9CAFXZA

[16] www.sony.co.uk/electronics/televisions/x9000b-series/specifications

sition between virtual views on command, or alternatively to set aside as much of the display as they wished toward an awareness of activity on other virtual views, allowing users to essentially vary their engagement with others activity dynamically, without impacting their partners usage of the TV. They found that this multi-user, multi-view TV significantly improved perceived ability to collaborate compared to a shared-use, single-view smart TV.

Fig. 9 Multi-view TV system [32]. Example of two users both in the dynamic split-screen mode, with different levels of engagement with each others activity.

Of note here is that in the near future fundamental and assumed limitations of TV displays may no longer be relevant; as suddenly as multi-screen usage came to dominate, so too might it come to be augmented with multi-view capable TV displays providing every user in the room with a unique physical view. Indeed prototype glasses-free 3DTVs have seen increasing prominence in consumer trade shows such as CES[17], with such displays having the potential to be used for glasses-free multi-view based on the gaze angle of the viewer. Displays such as these have the potential to substantially impact the problems of isolation and immersion in the living-room. With respect to immersion, there might no longer be a need to resort to a smaller device display in order to have a different media experience. With respect to isolation, if the majority of activity is conducted on the TV, then there is significant flexibility for users to transition between activities, or maintain a casual awareness of relevant activities, thus increasing the likelihood of shared experiences using the most immersive and accessible display in the room. Indeed we can even take a device-agnostic approach to activity; activity need not reside on any particular system to be displayed on the multi-view display, with screen-mirroring technology allowing for the activity of every device in the room to be observed through the TV, selectively and on-demand. Fundamentally, this allows for shared experiences without impacting any one individuals usage of

the TV display, and offers the prospect of a TV that supports multi-user usage without caveats.

4.3.1 Managing Shared Audio Spaces

Supporting multiple independent views or content streams necessitates that there be support for multiple associated audio streams to also be consumed. In consumer systems such as the aforementioned Samsung Multi-Link interface, audio focus was managed exclusively by the user; in this case the user would select which quadrant of the screen they wished to listen to, or pair a Bluetooth audio receiver up to the TV to receive audio for one quadrant in particular in the case where users wish to consume separate audio sources. This approach has been extended in literature, for example in the previously discussed shared screen mirroring follow up paper [33] a single shared audio space was used, with audio conflicts managed by the users (with the ability to selectively switch between listening to the TV and the devices), or automatically managed by the system when inferring focus (if an audio conflict was detected between the TV and a mirrored device, the TV would focus on the mirrored device whilst it was actively producing sound, pausing and thus silencing the TV in the process).

These solutions are predicated upon a single shared sound space, however where multi-view displays overcome physical constraints regarding viewing, there exist equivalent audio technologies with the capability to provide per-user audio streams. Readily available COTS solutions exist already, for example bone-conductance headphones allow users to receive audio streams without obstructing their capability to hear their environment, albeit at the cost of audio fidelity. However, as with multi-view and active shutter glasses, there are likely to be acceptability issues in requiring users to wear additional peripherals, albeit technologies such as Google Glass may lead to widespread adoption of bone-conductance technology. More generalisable, and wearable-free, solutions for per-user audio do exist within the remit of cutting edge research, for example through directional sound-beams (e.g. BoomRoom [36]), whilst existing COTS 3D audio / surround technology and mobile devices might also be utilized to create shared sound spaces, at the expense of perhaps suffering some amount of audio crosstalk.

5 Future Disruptive Technologies

Indeed the technologies and systems we have discussed thus far represent only a portion of the transformative living-room technologies with the capability to impact user behaviours in the near future, with a plethora of devices and displays likely to supplement or even replace TV usage entirely for some tasks, as demonstrated in Figure 10 where we describe

---

a devices influence over our TV usage in terms of a continuum from interaction toward wholesale replacement. These technologies have the potential to improve how we interact with the TV, augment our usage of the TV, replace our usage for some media types and contexts, or even supplant the TV entirely. We have already discussed the effect of mobile devices on our usage of TV, with their capability to move along our continuum: they can be conduits for interaction and control, sharing content and activity to the TV, or provide the personal and private experience the TV cannot.
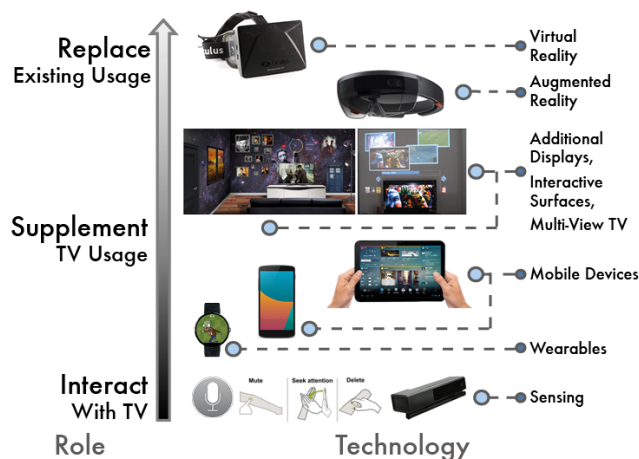
Fig. 10 An overview of new technologies likely to play a disruptive role in how we use the TV. From bottom to top, left to right - **Sensing:** Examples of how interaction with the TV might change, with voice input and gestures (e.g. skin gestures, pictured [61]), as well as an example of existing living-room sensing, the Kinect V2. **Wearables:** Android Wear smartwatch[18]. **Mobile Devices:** Phone and tablet form factors. **Additional Displays:** Left: Smart wallpaper concept [5] Right: Interactions beyond the physical TV screen [59]. **Eye-wear:** Microsoft Hololens Augmented Reality (AR) head-mounted display[19] and Oculus Rift (Virtual Reality) VR head-mounted display[20].

In terms of interaction with the TV, other technologies and form factors are already playing significant roles. Our capability to understand other forms of input, for example voice and gestures, has improved significantly. Such inputs are now commonplace in consumer smart TVs (albeit in limited forms with questionable adoption) with significant advancement occurring regarding how we design and detect gestures (e.g. fine hand gestures targeted toward TV interaction [60]). Wearable devices (e.g. smart watches) will provide new avenues for input to TV systems, increasing the likelihood of concurrent interaction.

Interactive surfaces will provide an alternative, gaze accessible, means of sharing activity in groups, circumventing some of the discussed limitations regarding screen-division (e.g. what happens when every wall is a screen e.g. BBC

smart wallpaper [5] where the wall augments the TV consumption experience with both relevant content and aesthetics, themed to match a programme being watched, or the space around the TV can be augmented e.g. [59]) allowing for TV interactions to be offloaded onto the surrounding space. Flexible and alternate form factor displays might augment particular contexts and spaces with interactivity that previously could not accommodate standard consumer displays. And Augmented Reality (AR) and Virtual Reality (VR) head-mounted displays may replace the TV entirely in terms of being the most immersive display in the room; how do we facilitate multi-user interaction and shared experiences when some users might be wearing VR headsets, whilst others might be observing this virtual activity on the TV, and how do we enable reciprocal awareness without impacting the VR experience a user is undertaking [29]? Is there still a place for a physical TV when every surface might be a display, or AR technology reaches such adoption that content can be rendered anywhere in the users' view? In considering the capabilities these new technologies introduce, we are better placed to re-evaluate what the TV is for, and how it should react to the presence and adoption of such displays.

It is not just the technology of TV viewing that may change, but also who we have the capability to watch it with. Whilst the focus of this paper is on collocated interaction, distributed at-a-distance viewing and social interaction around the TV viewing experience may play a significant part in our future usage. Given the sensing capabilities available in TVs currently, enabling synchronous viewing and communication [11] or more social interactions whilst consuming media [19] becomes feasible. Social second-screen usage will likely play an increasing role in our TV viewing, whilst our capability to be aware of who is consuming what, and when, will have increasing influence over what we consume e.g. the Netflix Android Wear application, where social recommendations can be pushed by friends to your smartwatch[21].

Much as the adoption of phones, tablets and laptops significantly changed our relationship with the TV over the course of less than a decade, so too might these new technologies, and in such cases the role and usage of the TV will continually evolve to fit into the ever-changing context of the living room. The challenge for designers is to help direct the role of TV as these new technologies see adoption, and use the shared, public, presentational resource that is the TV to benefit the social context of the living room, and the wider home. Fundamentally designers must help users to enjoy the most immersive entertainment experiences possible without being closed off to the possibilities of interacting with others and joining in other media experiences in the home; nonetheless, the TV in some form is likely to play

---

[18] www.android.com/wear/

[19] www.microsoft.com/microsoft-hololens/en-us

[20] www.oculus.com/

[21] http://www.theverge.com/2014/12/16/7401011/netflix-now-supports-android-wear-but-its-not-the-remote-control

a large part in media consumption for a significant time to come.

6 Implications For Future TV Designs

The TV plays a central role in our homes, bringing family and friends together by providing a common reference point. However, the capabilities and interactions of the modern day TV poorly reflect this reality. The work we have examined has served to outline why facilitating multi-user use of TVs is important, and how we might go about building TVs that can support this.

In the case of [30], interaction with the TV was designed to take into account social sharing behaviours, allowing for multiple users to interact with existing TV interfaces. This work demonstrates a focus on continuity, retaining the mental models, behaviours, and interfaces that users have become accustomed to. However such approaches do nothing to tackle the increasing digital isolation that is prevalent in living-rooms, nor do they adequately allow for the private and independent activity which has in part motivated the rise of multi-screening. [31] demonstrates the use of social behaviours to share mirrored use of the TV, providing the capability for awareness of other devices in the room. Indeed, this is where social behaviours are most necessary: users must have the ability to determine when, and how, this shared awareness is provided for such designs to hope to be deemed acceptable.

Shared screen-mirroring offers the possibility of diminishing the "digital bubble" effect, but at the cost of interrupting existing usage of the TV for others that may not wish to attend to the mirrored content; whilst screen-division can be employed to reduce such negative effects, again the problem of independent use proves to be a significant barrier. [32] established the viability of a new class of TV displays for circumventing this problem of independent and shared use, validating that multi-view TVs (capable of providing multiple independent and semi-private views to users) can allow for both independent and collaborative use, through designs allowing users to transition between views and between focused and casual awareness modes. This work was equally found to be applicable to mirrored device activity [33].

Each of these papers takes a different approach to solving the "multi-user" TV problem. They demonstrate how TVs can be designed to take into account the social context they inhabit: from shared use of existing TV interfaces, to using the TV as a medium by which multi-screen activity might be shared, to TVs capable of providing multiple independent physical views, allowing for both private and collaborative activity on a single shared display, and user-determined engagement with multi-screen activity. These systems exemplify how we can design TVs accessible to anyone in the room, potentially burst the isolating "digital bub-

bles" of multi-screen usage, and diminish our reliance on multi-screening for the majority of independent activity and media consumption.

These papers present different aims regarding the future role of the TV: do we continue to focus on existing interactions and interfaces, do we support multi-screen use and use the TV as a conduit for other device activities, or do we expand TV interaction and technology to support independent multi-user smart TV use and instead design mechanisms for facilitating awareness and shared experiences on-demand. Although we cannot predict which of these lines the TV will evolve along, it is clear that each of these approaches bestows new and significant capabilities to the home TV, and a TV which incorporated these designs could more readily support the social context in which it resides.

7 Overview

In this paper we have discussed existing TV and multi-screen usage, the problems this usage introduces with respect to the role of the TV in shared social contexts, and the potential impact new technologies might have on how we use, and interact with, the TV. Shared use of the TV is problematic, both from an interaction perspective, and from an inability to independently use the TV without impacting upon others usage. Users currently overcome these issues through multi-screening, but this too is problematic from a social perspective, with the potential for increased digital isolation and a lack of shareability with respect to those in the room. We have reviewed relevant literature with the aim of demonstrating how these problems can be tackled through TV interaction design, presenting ways in which multi-user use can be facilitated through shared-use and multi-view interfaces, and examining how the TV can enable increased shareability, and thus awareness, of device activity. Through this, we have demonstrated that the TV is capable of substantially more than it is currently asked to do; contrary to existing usage, it can be of increasing relevance in the multi-user, multi-screen era.

References

1. Abe, Y., Matsusako, K., Kirimura, K., Tamura, M., Nakashima, M., and Ito, T. Tolerant Sharing of a Single-user Application Among Multiple Users in Collaborative Work. In *Companion Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW 2010). 2010* (2010), 555–556.
2. Agrawala, M., Beers, A. C., McDowall, I., Fröhlich, B., Bolas, M., and Hanrahan, P. The two-user responsive workbench: Support for collaboration through individual views of a shared space. In *Proceedings of the 24th Annual Conference on Computer*

Graphics and Interactive Techniques, SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co. (New York, NY, USA, 1997), 327–332.

3. Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction: Designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, ACM (New York, NY, USA, 2010), 121–130.

4. Birnholtz, J. P., Grossman, T., Mak, C., and Balakrishnan, R. An exploratory study of input configuration and group process in a negotiation task using a large display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, ACM (New York, NY, USA, 2007), 91–100.

5. Campbell, R., Felton, R., and Hoarse, C. Smart Wallpaper. In *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video, Industry Track - TVX '14*, ACM Press (June 2014).

6. Cesar, P., and Chorianopoulos, K. The evolution of tv systems, content, and users toward interactivity. *Found. Trends Hum.-Comput. Interact. 2*, 4 (Apr. 2009), 373–95.

7. Clark, A., Dünser, A., Billinghurst, M., Piumsomboon, T., and Altimira, D. Seamless interaction in space. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, OzCHI '11, ACM (New York, NY, USA, 2011), 88–97.

8. Clinch, S., Davies, N., Kubitza, T., and Friday, A. Ownership and trust in cyber-foraged displays. In *Proceedings of The International Symposium on Pervasive Displays*, PerDis '14, ACM (New York, NY, USA, 2014), 168:168–168:173.

9. Deloitte. Perspectives on TV: dual screen, catch-up, connected TV, advertising, and why people watch TV. *MediaGuardian International Television Festival* (2012).

10. Dodgson, N. A. Multi-view autostereoscopic 3D display. In *Stanford Workshop on 3D Imaging*, Stanford University (2011).

11. Ducheneaut, N., Moore, R. J., Oehlberg, L., Thornton, J. D., and Nickell, E. Social tv: Designing for distributed, sociable television viewing. *Intl. Journal of Human–Computer Interaction 24*, 2 (2008), 136–154.

12. Ericsson Consumerlab. TV and Media: Identifying the needs of tomorrow's video consumers.

13. Google. The New Multi-Screen World (slides). *Google Think Insights* (2012).

14. Greenberg, S. Sharing views and interactions with single-user applications. In *Proceedings of the ACM SIGOIS and IEEE CS TC-OA Conference on Office Information Systems*, COCS '90, ACM (New York, NY, USA, 1990), 227–237.

15. Gross, T., Stary, C., and Totter, A. User-centered awareness in computer-supported cooperative work-systems: Structured embedding of findings from social sciences. *International Journal of Human-Computer Interaction 18*, 3 (2005), 323–360.

16. Gutwin, C., and Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work 11*, 3 (Nov. 2002), 411–446.

17. Hou, J., Nam, Y., Peng, W., and Lee, K. M. Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Comput. Hum. Behav. 28*, 2 (Mar. 2012), 617–623.

18. Hritzuk, N., Esquero, I., and Jones, K. Cross screen engagement. *Microsoft Advertising 3586*, March (2013), 1–41.

19. Huang, E. M., Harboe, G., Tullio, J., Novak, A., Massey, N., Metcalf, C. J., and Romano, G. Of social television comes home: A field study of communication choices and practices in tv-based text and voice chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, ACM (New York, NY, USA, 2009), 585–594.

20. Ingrey, M. Triple-Screening: A New Phenomenon, 2014.

21. Kim, S., Cao, X., Zhang, H., and Tan, D. Enabling concurrent dual views on common lcd screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 2175–2184.

22. Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F. Interactive stereoscopic display for three or more users. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, ACM (New York, NY, USA, 2001), 231–240.

23. Kreitmayer, S., Laney, R., Peake, S., and Rogers, Y. Sharing bubbles: Reflections on offline multi-surface scenarios. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, ACM (New York, NY, USA, 2013), 1405–1408.

24. Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., and Froehlich, B. C1x6: A stereoscopic six-user display for co-located collaboration in shared virtual environments. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, SA '11, ACM (New York, NY, USA, 2011), 188:1–188:12.

25. Lalanne, D., and Lisowska Masson, A. A fitt of distraction: Measuring the impact of distracters and multi-users on pointing efficiency. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, ACM (New York, NY, USA, 2011), 2125–2130.

26. Lissermann, R., Huber, J., Schmitz, M., Steimle, J., and Mühlhäuser, M. Permulin: Mixed-focus collaboration on multi-view tabletops. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, ACM (New York, NY, USA, 2014), 3191–3200.

27. Lucero, A., Jones, M., Jokela, T., and Robinson, S. Mobile collocated interactions: Taking an offline break together. *interactions 20*, 2 (Mar. 2013), 26–32.

28. Matusik, W., Forlines, C., and Pfister, H. Multiview user interfaces with an automultiscopic display. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '08, ACM (New York, NY, USA, 2008), 363–366.

29. McGill, M., Boland, D., Murray-Smith, R., and Brewster, S. A. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*, ACM Press (New York, New York, USA, Apr. 2015).

30. McGill, M., Williamson, J., and Brewster, S. A. How to lose friends & alienate people. In *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video (TVX '14)*, ACM Press (New York, New York, USA, June 2014), 147–154.

31. McGill, M., Williamson, J., and Brewster, S. A. Mirror, mirror, on the wall. In *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video (TVX '14)*, ACM Press (New York, New York, USA, June 2014), 87–94.

32. McGill, M., Williamson, J., and Brewster, S. A. It Takes Two (To Co-View): Collaborative Multi-View TV. In *Proceedings of the 2015 ACM international conference on Interactive experiences for TV and online video (TVX '15)*, ACM Press (New York, New York, USA, June 2015).

33. McGill, M., Williamson, J., and Brewster, S. A. Whos The Fairest Of Them All: Device Mirroring For The Connected Home. In *Proceedings of the 2015 ACM international conference on Interactive experiences for TV and online video (TVX '15)*, ACM Press (New York, New York, USA, June 2015).

34. MillwardBrown. Marketing in a multiscreen world. Tech. rep., 2014.

35. Morris, M. R. Web on the wall: Insights from a multimodal interaction elicitation study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, ACM (New York, NY, USA, 2012), 95–104.

36. Müller, J., Geier, M., Dicke, C., and Spors, S. The boomroom: Mid-air direct interaction with virtual sound sources. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, ACM (New York, NY, USA, 2014), 247–256.

37. Nacenta, M. A., Jakobsen, M. R., Dautriche, R., Hinrichs, U., Dörk, M., Haber, J., and Carpendale, S. The lunchtable: A multi-user, multi-display system for information sharing in casual group interactions. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, ACM (New York, NY, USA, 2012), 18:1–18:6.

38. Nielsen. Australian Multi-Screen Report Q1. *nielsen.com* (2014), 1–48.

39. Noah, S. Giant TVs 'bring families together', 2012.

40. NPD. Inception of Screen Sharing. *npd.com* (2013).

41. Ofcom. Communications Market Report, 2012.

42. Ofcom. Communications Market Report, 2013.

43. Ofcom. Communications Market Report, 2014.

44. Peterka, T., Kooima, R., Girado, J., Ge, J., Sandin, D., Johnson, A., Leigh, J., Schulze, J., and DeFanti, T. Dynallax: Solid state dynamic parallax barrier autostereoscopic vr display. In *Virtual Reality Conference, 2007. VR '07. IEEE* (March 2007), 155–162.

45. Pohl, H., and Murray-Smith, R. Focused and casual interactions: Allowing users to vary their level of engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 2223–2232.

46. Recombu. Digital Divorce: 24 per cent of couples watch on-demand TV in different rooms - Recombu, 2013.

47. Schmid, O., Lisowska Masson, A., and Hirsbrunner, B. Collaborative web browsing: Multiple users, multiple pages, concurrent access, one display. In *Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '12, ACM (New York, NY, USA, 2012), 141–150.

48. Schmidt, K. The problem with 'awareness': Introductory remarks on 'awareness in cscw'. *Comput. Supported Coop. Work 11*, 3 (Nov. 2002), 285–298.

49. Stewart, J., Bederson, B. B., and Druin, A. Single display groupware: A model for co-present collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, ACM (New York, NY, USA, 1999), 286–293.

50. Tee, K., Greenberg, S., and Gutwin, C. Artifact awareness through screen sharing for distributed groups. *Int. J. Hum.-Comput. Stud. 67*, 9 (Sept. 2009), 677–702.

51. Terrenghi, L., Quigley, A., and Dix, A. A taxonomy for and analysis of multi-person-display ecosystems. *Personal Ubiquitous Comput. 13*, 8 (Nov. 2009), 583–598.

52. Thinkbox. Screen Life: The View from the Sofa, 2012.

53. Thinkbox. TV Together: a very social medium, 2013.

54. Thinkbox. Screen Life : TV advertising everywhere, 2014.

55. Truscott-Howell, J. Compelling content and social conversations drives resurgence in linear TV, 2014.

56. Tse, E., Histon, J., Scott, S. D., and Greenberg, S. Avoiding interference: How people use spatial separation and partitioning in sdg workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, ACM (New York, NY, USA, 2004), 252–261.

57. Turkle, S. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, Inc., New York, NY, USA, 2011.

58. "TV Licensing". A look at the nation's changing viewing habits from TV Licensing, 2012.

59. Vatavu, R.-D. There's a world outside your tv: Exploring interactions beyond the physical tv screen. In *Proceedings of the 11th European Conference on Interactive TV and Video*, EuroITV '13, ACM (New York, NY, USA, 2013), 143–152.

60. Vatavu, R.-D., and Zaiti, I.-A. Leap gestures for tv: Insights from an elicitation study. In *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '14, ACM (New York, NY, USA, 2014), 131–138.

61. Weigel, M., Mehta, V., and Steimle, J. More than touch: Understanding how people use skin as an input surface for mobile computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, ACM (New York, NY, USA, 2014), 179–188.

62. You, W., Fels, S., and Lea, R. Studying vision-based multiple-user interaction with in-home large displays. In *Proceedings of the 3rd ACM International Workshop on Human-centered Computing*, HCC '08, ACM (New York, NY, USA, 2008), 19–26.