It Takes Two (To Co-View): Collaborative Multi-View TV

Mark McGill†
m.mcgill.1@research.gla.ac.uk

John Williamson‡
jhw@dcs.gla.ac.uk

Stephen A. Brewster†
stephen.brewster@glasgow.ac.uk

† Glasgow Interactive Systems Group ‡ Inference, Dynamics and Interaction Group
School of Computing Science, University of Glasgow, Glasgow, G12 8QQ, Scotland, UK

ABSTRACT
This paper investigates how we can design interfaces and interactions for multi-view TVs, enabling users to transition between independent and shared activity, dynamically control awareness of other users activities, and collaborate more effectively on shared activities. We conducted two user studies, first comparing an Android-based two-user TV against both multi-screen and multi-view TVs. Based on our findings, we iterated on our design, giving users the ability to dynamically set their engagement with other users activity. We provide the foundations of a multi-user multi-view smart TV that can support users to transition between independent and shared activity and gain awareness of the activities of others, on a single shared TV that no longer suffers the bottleneck of one physical view. Through this we significantly improve upon a user's capability for collaborative and independent activity compared to single-view smart TVs.

Author Keywords
Multi-user; TV; Multi-view; Displays; Engagement;

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
There has rarely been a technology of such prevalence as the TV. With over ~52.2 million [20] of them in the UK alone, they continue to be a central component in our home lives. A report by Ofcom [21] showed that 91% of UK adults view TV on the main set each week, and underlined the importance of the living-room TV specifically 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". The TV is often a social medium, with one survey suggesting that over 52% of live viewing and 56% of time-shifted viewing is shared, predominantly with one other person [26]. However, its use is often supplemented or entirely supplanted by other devices, for multi-tasking, co-viewing or private viewing of content. It is both this proliferation of TVs, and the rapid uptake of other devices (such as tablets and phones) which confirm a fundamental problem of the TV: shareability. We use TVs because they offer large, accessible, HD displays which enhance our media experiences. However, this naturally disposes users against sharing the display: split-screen and picture-in-picture approaches force users to attend to distracting content they may not necessarily wish to, whilst compromising the existing experience through obscuring, downscaling, or compromising the aspect ratio of, the content being consumed. Additionally, they offer no privacy considerations. Personal devices circumvent these issues, guaranteeing the user full use of a semi-private display.

TVs also have issues regarding interactivity in shared contexts. Consumer approaches such as Chromecast[1] offload interaction from the TV display onto secondary devices, using the TV as a terminal for showing selected content or mirrored activity. These approaches allow for both independent (device-based) and shared (TV-based) activity, and some awareness of what others are doing. However, when second screening (using a secondary device alongside the TV) it can be problematic to actively or passively share activity with others [14], whilst interacting with others on the same TV display is often poorly facilitated [13] and potentially distracting, with no capability for truely independent views. Additionally, the phones, tablets or other devices being used are often inferior to the TV in some important respects, for example, in terms of size, casual accessibility to others in the room, and socialization, with users in their own private "digital bubble" [8]. Finally, 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", whereby 24% of polled couples resorted to going into different rooms in order to watch TV separately [19].

These problems arose because of a fundamental limitation of the TV: it has one shared physical view. However, this technological limitation is being overcome, with existing consumer TVs capable of multiplexing many separate views in what is often termed "multi-view" [10]. These allow users the capability to consume content independent of others in the room, whilst retaining the same shared focal point e.g. one user might be watching sport whilst the other watches the news. However, we can imagine interactions that go beyond this e.g. allowing users to privately investigate details on a film they are watching, with the capability for others to switch over to see these details if they so wish, all without having to resort to a secondary devices. This paper examines how we can design interfaces and interactions for multi-view TV usage, enabling users to transition between independent and shared activity, dynamically control awareness of other

[1] google.co.uk/chrome/devices/chromecast/

TVX'15, June 03 – 05, 2015, Brussels, Belgium

ACM 978-1-4503-3526-3/15/06...$15.00
http://dx.doi.org/10.1145/2745197.2745199

users' activities, and collaborate more effectively on shared activities. Our first study compares an Android-based two-user smart TV against both multi-screen and multi-view displays in a collaborative movie browsing task. Based on our findings, we iterate on our design, giving users the ability to transition between casual (viewing both views) and focused (viewing only one view) modes of usage, and dynamically set their engagement with other users' activities. This work provides a foundation for multi-user multi-view smart TVs that can support both collaborative and independent activity, and transitions between activities on a single shared TV.

BACKGROUND

Multi-View Displays

Multi-view displays are displays that are capable of providing two or more independent views to two or more users. There are a number of technologies that are capable of achieving this aim[4] e.g. Lenticular displays, using sheets of lenticular lenses atop a standard LCD screen allowing for different views based on gaze angle; Parallax-barrier or Masked Displays, employing masks (e.g. singular portholes[7]) in order to control what subpixels are viewed at a given angle. Today's state-of-the-art multi-view technologies are active-shutter displays, which have high refresh rate, low pixel-persistence[2] 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 levels of crosstalk[3] 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. Active shutter displays are relatively commonplace currently, being used predominantly for consumer 3DTV. As such there are a number of multi-view capable 3DTVs on the market e.g. LG "dual-play"[4] and Samsung "Multi-View" displays[5] support either two-view/two-person gaming, or the capability to independently consume different media, with transitions managed via a physical switch on the 3D glasses. This synchrony between 3DTV and multi-view TV is likely to continue: consumer glasses-free 3DTVs demonstrated thus far rely on lenticular displays. As a consequence, displays capable of supporting a number of independent views based on gaze angle, without the need for glasses, might soon be a reality in consumer households.

Usage Versus Single And Multi-Display Groupware

Multi-view displays can be used by single users or groups and have a number of advantages over comparable smart TVs in each case. 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 [12]. In multi-user contexts, they have been used to support single display privacyware [24], independent and collaborative activity on table-tops, e.g. Permulin by Lissermann *et al.*[10] which supported two users sharing a 120Hz two-view display, or Permulin's precursor[1], and independent views in groups such as in the case of C1x6[9] which employed multiple projectors in order to achieve a 12-view 360 Hz display allowing for 6 stereoscopic views.

It is in terms of multi-user use that multi-view displays have the most potential. In Single Display Groupware (SDG) and Multi-Display Groupware (MDG), multiple users have to either share display resources, or split attention across different displays. In SDG, they typically share use of the interface presented, e.g. [13], or partition the display in order to accommodate multiple interfaces or activities, e.g. [31]. In MDG they leverage additional displays in order to provide elements of task independence, perhaps moving activities to a secondary display, or a personal private display, or even having multiple shareable displays such that there are more display surfaces to present or interact on e.g. [18].

Multi-view displays have the potential to combine the advantages of both SDG and MDG. SDG provides a shared focus of attention and thus activity [5], which has been shown to significantly improve users' ability to collaborate [29]. MDG allows for task independence and selective or casual awareness. For example in Lunchtable [18] where group work could be spread between multiple displays, allowing for an element of independence, but with activity still be visible to the group as a whole. Multi-view displays have the capability for both independent operation and collaboration, with a shared focus of attention throughout. However, unlike in SDG and MDG, transitioning between independent and collaborative states, and gaining mutual awareness of the activity of others (e.g. through glancing, peeking, peripheral vision) must be explicitly designed for, as users no longer have the ability to manage their visual attention via gaze. This is a significant problem with respect to collaboration and coordination, as systems utilizing multi-view displays must actively communicate the requisite information to allow users to gain awareness of group activity.

Permulin[10] attempted to address this issue by providing a set of behaviours that enabled users to selectively gain a level of awareness of their partners' activities on a collaborative table-top. This was achieved through providing the ability to have both private views and a shared group view which could contain private information, as well as the ability to peek at a collaborators' private views to facilitate activity awareness. Permulin exemplified both why multi-view displays have great potential for collaboration. It provided a shared-focus workspace with the ability to collaborate or operate independently, whilst also demonstrating the problems faced in trying to provide the capability to transition between shared and private views. Their display management behaviours were heavily reliant on use of the touch surface table-top display. We would suggest that more generalised behaviours for managing full use of the display, and transitions between available views, are required if multi-view displays are to be usable in home contexts.

[2]Pixel persistence: the time it takes a pixel to transition from its current state to its next state

[3]Crosstalk: the extent to which one image is retained into the next image e.g. where one view is a car and the other a boat, crosstalk would be manifested as the boat being visible (from a faint outline to wholly superimposed) in the car view, and *vice versa*.

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

[5]samsung.com/us/video/tvs/KN55S9CAFXZA

Shared-Use TV: From Devices To The Display

The private "digital bubble" [8] of device usage offers a problematic barrier to socialization and interaction, with mobile phone use in particular having significant anti-social connotations [28]. Efforts have been made to penetrate this bubble, for example Lucero *et al* [11] proposed mobile collocated interactions, whereby users would "take an offline break together", pooling their device resources toward "shared-multiuser experiences". They aimed to create the capability for joint attention, whilst enforcing a break from online socialization, appropriating mobile device displays for passing photos around a table. This emphasis on shareability and joint attention is important as it underlines how co-located interactions are made to be more effective, through the ability to share awareness and activities. This link between awareness and our capability to collaborate has been a frequent topic of discussion within CSCW [6].

However, mobile devices are not necessarily the most shareable displays in the room. McGill *et al.* [14] 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* [25] discussed the scale of displays relative to users' visual angle and distance, noting that the scale of the display must match the social interaction space.

There are a variety of ways in which TV displays can be shared, from proxemic approaches employed by Ballendat *et al* [2] where the display adapted to the angle and proximity of whomever was interacting with it, to approaches where the sharing was based on social behaviours [13]. However, in all cases, interfaces and interactions have likely been compromised by the fact that there is inevitably only one TV view to be shared, limiting the potential for independent activity.

Finally, there is the concept of engagement to consider. Pohl & Murray-Smith's focused–casual continuum describes interaction techniques according to the degree to which they allow users to determine how much attention and effort they choose to invest in an interaction i.e. the ability to adapt how engaged they are [22]. This concept has rarely been explored within the space of TV displays. Whilst it is beneficial to be able to give users the ability to dictate their level of awareness and engagement in others activity, there is an inherent difficulty in varying engagement when physically sharing a display. Multiple users could be attending to, and interacting with, the display without the ability to personally control their level of engagement with others' activity, e.g. employing split-screen approaches.

Summary

Multi-view capable TVs have the potential to facilitate multi-user collaborative use in home contexts, removing the reliance on less social and shareable second screen interactions. As such, this paper investigates how we can design generalised multi-view interfaces that can support transitions between independent and collaborative tasks, whilst providing the capability for awareness regarding on-going activity occurring across available views.

STUDY 1 - MULTI-USER MULTI-VIEW TV

Given the potential of multi-view displays, the aim of our first study was to design, develop and evaluate a fully functional Android-based multi-view TV. Throughout this paper we chose one important limitation: that we would be investigating only the visual component of such a system, and not the audio. Enabling per-user audio whilst retaining the ability to hear and converse with others is an area of active research, with solutions ranging from bone-conductance headphones, to directional sound-beams (e.g. BoomRoom [17]) and it is reasonable to expect these systems being incorporated into future multi-view displays. The study had the following aims:

- To allow users to gain awareness of each others' activity through a simple set of behaviours by which they could transition between virtual views without compromising in terms of distraction, aspect ratio and utilized screen area;

- To determine how multi-view TVs compared to single-view TVs in terms of perceived workload, usability, and ability to collaborate.

- To determine the extent to which users were aware of each others' activity and how close this was to their optimum level of awareness.

In order to accomplish this, we designed and built a two-view (meaning two interactive virtual views), two-user (meaning the system supported two independent physical views made up of whatever we wish to render of the virtual views) multi-view system with the capability to allow two users to transition between collaborative and independent activity. An overview of this design can be seen in Figure 1.

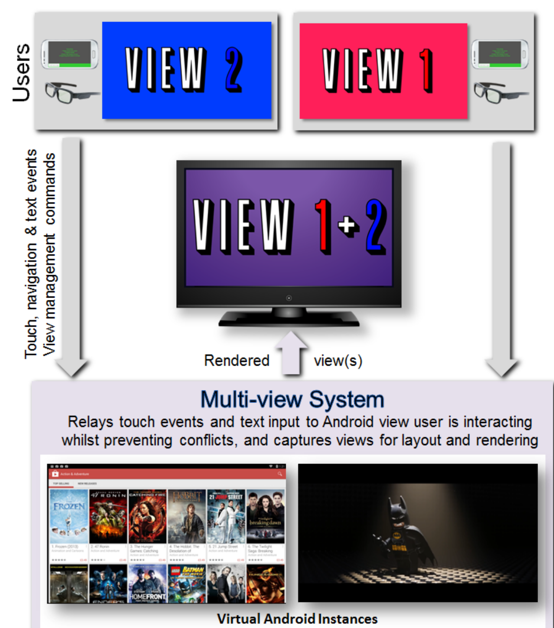

Figure 1. Overview of multi-view system in both studies. Here two users can have completely independent physical views (labeled View 1 and View 2) made up of however we wish to render our virtual Android views, with inputs routed appropriately.

We provided users with two touch gestures (enacted via a touchpad; see *Implementation*) to switch between the two available virtual views. The *transition* gesture switched the user between the two available virtual views, at which point they were free to interact with the current view. The *peek* gesture allowed the user to switch to the view they were not currently interacting with for so long as they performed the gesture, at which point they would return to their current interactive view. Through these behaviours, we hypothesized that users would be able to adequately determine their awareness of each others' activity, transitioning between independent and collaborative states, and gaining awareness of what activity their partner was performing, if they felt the need.

Implementation of Multi-view Display

To provide users with a fully-functioning multi-view TV we realised that the typical approach of implementing software capable of allowing users to only perform a given task (e.g. implementing a multi-view photo browsing application) would not be representative of smart TV usage. Thus, we built a generalised, ecologically valid multi-view system that would give users capabilities above and beyond current smart TV capability, allowing them to interact commonly used consumer applications. Given the adoption of Android into the smart TV area, we believed that building a system utilizing multiple emulated Android devices would best approximate this. As such, we used instances of Genymotion[6], a high-performance x86 Android emulator, running Android 4.x.

To present users with entirely separate views, which could be of the same virtual Android device, or different devices, depending on the users current display settings, we utilized nVidia 3D Vision, an active-shutter IR transmitter for the PC, coupled with an nVidia graphics card performing stereoscopic rendering at 120Hz, 60Hz worth of "left" eye frames, and 60Hz worth of "right" eye frames. To provide users with independent views, we needed to be able to present only the "left" eye frames to one user, and the "right" eye frames to another. This was achieved using Youniversal active-shutter glasses[7] which had the capability to be set into a "2D" mode where only one of the left or right frames of the 3D image was allowed through both eyes. Our emulator screen-capture software then rendered a stereoscopic image, such that the left image constituted of whatever view we wished to provide one user, and the right image whatever view we wished to provide the other user. This gave users the ability to view separate Android emulators (hereafter virtual views), or transition to the same virtual view, all without affecting their partner's physical view. To minimize crosstalk, we utilized a 24" BenQ XL2411T Display which supported nVidia LightBoost, resulting in little to no perceptible ghosting between views; this was important as it meant that awareness could only be gained through our multi-view behaviours and mechanisms, not through inadequacies in the technology.

To interact with the Android virtual views, we used Samsung S3 phones as touchpads, rendering coloured cursors which matched the colour of the user's touchpad on whichever view they were interacting with. Additionally, when occupying a view, a coloured eye would be rendered in the bottom right corner, to allow users to be aware of when they were both sharing the same view. These touchpads supported a simple set of gestures: dragging one finger moved the on-screen cursor, tapping one finger made a selection; dragging two fingers performed a scroll gesture; tapping four fingers caused a *transition* action, whilst pressing four fingers performed a *peek* action for so long as the fingers were on the touchpad. Additionally the physical back, home, and application switcher buttons were mapped to the same functions in the emulator. Text input was provided via the onscreen keyboard. These interaction events were sent to our software then routed to the appropriate Android virtual view via the Android Developer Bridge.

Experimental Design

The study design incorporated three Conditions: *(1) Single display* with one LCD display and one shared virtual Android view, as a comparative baseline for a standard smart TV; *(2) Two displays* with two LCD displays with a virtual Android view on each, allowing us to measure the default level of awareness of each others activity as users could transition between views by gaze; *(3) Multi-view display* with a single LCD display providing two independent physical views, each displaying either of two virtual Android views depending on the users usage of the system (see Figure 2).

For our task we chose movie browsing, a loosely coupled and ecologically valid collaborative task that commonly occurs on TVs e.g. collaborative searching for entertainment in [15]. Movie browsing can be performed independently or together, but the eventual outcome (having to select acceptable movies to the group) necessitates collaboration. Users were instructed to browse a given set of categories of movies in the Google Play store application, with the task of selecting movies to watch together with mutual friends for the du-

[6] genymotion.com/

[7] xpand.me/products/youniversal-3d-glasses/

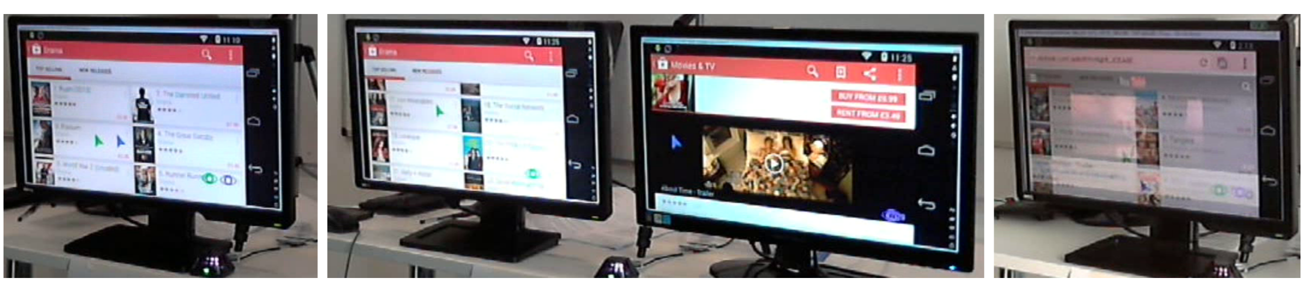

Figure 2. Left: Condition 1, single display with one virtual view. Middle: Condition 2, two displays, each with its own virtual view. Right: Condition 3, multi-view display when viewed without active-shutter glasses. This supports two independent physical views (and thus two users), constituting of whichever Android virtual view each user wishes to interact with.

ration of each Condition. Three categories were selected for each Condition, with users instructed they could browse them however they saw fit. Additionally, users had the capability to watch trailers (with the instruction to moderate trailer viewing time to under a minute per trailer) and use a selection of other applications if they so wished, namely the Chrome web browser and the IMDB app. Users were tested for 15 minutes per Condition in a within-subjects design, and there were 9 pairs, 18 users in all (mean age=23.6, SD=5.5, 16 male, 2 female) recruited from University mailing lists as pairs that knew each other (e.g., friends, family, etc.).

To determine the effects on users' abilities to collaborate effectively, we utilized post-condition questionnaires derived from a previous collaborative TV [14]. We also measured the effect our systems had on workload (NASA TLX) and usability (System Usability Scale (SUS) [3]). Additionally, users were asked to rank the Conditions in order of preference.

To establish the default / optimal level of awareness of each others' activity, for the *two displays* Condition we recorded and analysed video footage of each participant, coding timestamps regarding which display the participant was looking at, if any. These timestamps, along with logs of viewing in the multi-view display Condition, were parsed such that we could accurately compare the viewing behaviour across Conditions. Where applicable, Gini coefficients were calculated. These are a measure of inequality used for analysing viewing distribution in previous studies[30, 14]; 1 denotes maximum inequality i.e. $100-0$ or $0-100$, and 0 maximum equality i.e. a 50-50 distribution when dealing with two items. As our use of Gini coefficients typically involves two comparison points, for both studies in this paper we also used directed Gini coefficients where applicable, whereby we encode the direction of the inequality such that 100-0 would resolve to 1, whilst $0-100$ would resolve to -1 (meaning the Gini coefficient resolved to a measure of distance between two points).

Results

Where appropriate a repeated-measures ANOVA (GLM) or Friedman test with *post-hoc* Wilcoxons was performed, green indicates $p < 0.05$. We found significant differences between the Condition 1 (single display) and Conditions 2 (two displays) and 3 (multi-view display). Conditions 2 and 3 were superior in terms of capability to collaborate (e.g. WS-1, MO-1), ability to work independently (WS-2), and workload/usability (see Table 1). However there were no significant differences between Conditions 2 and 3, with Condition 2 typically having only moderately higher mean scores.

User rankings (see Figure 3) again showed significant differences between Condition 1 and Conditions 2/3, with both conditions ranked better. There was no significant difference between the mean rankings of Conditions 2 and 3, however Condition 2 was ranked better than Condition 3. Given that the two display condition provided users with the ability to attain their optimal level of awareness with respect to their partners activity, this implies that our behaviours for managing awareness failed to match this standard.

Question	Condition: 1: Single Display	Condition: 2: Two Displays	Condition: 3: Multi-view Display	Friedman Test	Wilcoxon *Post-hoc* ($p < 0.05$)
WS-1: We were able to collaborate effectively	3.11 (1.81)	4.94(1.21)	5.00 (0.77)	$\chi^2(2) = 16.0, p < 0.01$	1-2, 1-3
WS-2: We were able to work independently to complete the task	1.94(1.47)	5.67(0.49)	5.33(0.49)	$\chi^2(2) = 31.5, p < 0.01$	1-2, 1-3
WS-3: It was easy to discuss the information we found	4.39 (1.65)	5.50 (0.62)	5.39 (0.78)	$\chi^2(2) = 7.61, p < 0.05$	None
WS-4: We were able to work together to complete the task	3.94 (1.70)	5.28 (1.07)	4.78 (1.44)	$\chi^2(2) = 7.4, p < 0.05$	1-2
WS-5: I was able to actively participate in completing the task	3.83 (1.425)	5.61 (0.50)	5.33 (0.77)	$\chi^2(2) = 21.4, p < 0.01$	1-2, 1-3
MO-1: How well did the system support collaboration?	2.56 (1.72)	4.72 (1.18)	4.78 (0.88)	$\chi^2(2) = 17.2, p < 0.01$	1-2, 1-3
MO-2: How well did the system support you to share particular information with your partner?	3.94 (2.01)	4.61 (1.75)	5.17 (0.92)	$\chi^2(2) = 1.82, p = 0.4$	NA
MO-3: I was able to tell when my partner was looking at what I was browsing?	4.89 (1.60)	5.17 (0.92)	5.39 (0.61)	$\chi^2(2) = 0.383, p = 0.83$	NA
MO-4: How well did the system support you to see/review what your partner was talking about?	4.83 (1.25)	5.33 (0.69)	5.50 (0.62)	$\chi^2(2) = 5.57, p = 0.06$	NA
WE-1: The system was helpful in completing the given task	3.11 (1.68)	5.06 (0.94)	5.06 (0.87)	$\chi^2(2) = 20.8, p < 0.01$	1-2, 1-3
WE-2: I was aware of what my partner was doing	5.39 (0.85)	5.00 (1.33)	4.67 (0.97)	$\chi^2(2) = 9.48, p < 0.01$	None
PE-1: My partner was aware of what I was doing	5.28(0.96)	5.06 (1.06)	4.56 (1.10)	$\chi^2(2) = 9.49, p < 0.01$	None
TLX: Overall Workload	38.50 (24.70)	19.40 (16.00)	22.20 (15.40)	$\chi^2(2) = 10.6, p < 0.01$	1-2, 1-3
SUS: System Usability Scale	58.10 (22.20)	83.30 (14.30)	78.90 (13.80)	$\chi^2(2) = 13.2, p < 0.01$	1-2, 1-3

Table 1. Questions from [14]: (WS) WebSurface[27], (MO) Mobisurf[23], (WE) WeSearch[16], (PE) Permulin[10]. Questions were 7-point Likert scale (results range from 0-6, higher is better). TLX is from 0 (lowest) to 100 (highest), SUS is from 0 (worst) to 100 (best). Means with standard deviations are presented across Conditions. A Friedman test was conducted with *post hoc* Bonferroni corrected Wilcoxon tests.

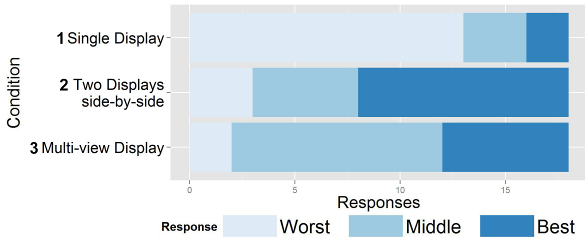

Figure 3. User ranking – Friedman test $\chi^2(2) = 10.3$ p < 0.01, *post hoc* Bonferroni corrected Wilcoxon test showed differences between 1-2 and 1-3

Viewing and Interaction

Examining the viewing patterns and behaviours exhibited in Conditions 2 and 3, we find significant differences in terms of viewing behaviour (see Table 2 and Figure 4). This difference is visualized in Figure 4, where we can see that in Condition 2 ~50% of overall viewing and ~90% of viewing instances were accounted for in viewing instances which lasted under 10 seconds; in comparison, Condition 3 demonstrates that users relied on much longer views, showing a clear difference in behaviour.

	Condition		RM-Anova
	2	3	
Mean Duration of Views (secs)	3.39 (3.51)	40.64 (37.40)	$\chi^2(1) = 16.6, p < 0.01$
Gini: Interaction	0.839 (0.27)	0.641 (0.34)	$\chi^2(1) = 3.75, p = 0.053$
Gini: Viewing	0.394 (0.233)	0.447 (0.306)	$\chi^2(1) = 0.356, p = 0.55$

Table 2. Mean (SD) viewing and Interaction comparison between Conditions 2 and 3. Gini coefficients show equality regarding how likely users were to view or interact with either Android view, 1 is maximum inequality, 0 is maximum equality.

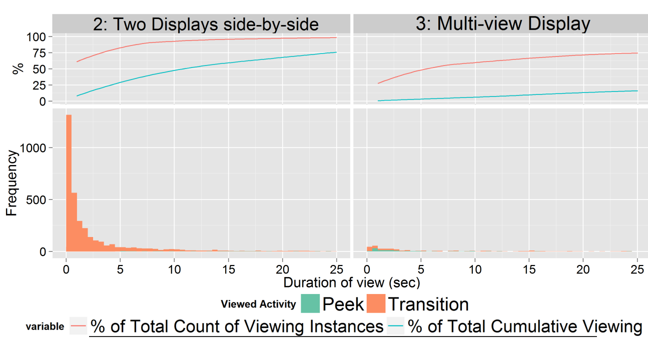

Figure 4. Individual viewing behaviour across participants. Bottom: Histogram (0.5 second bins) counting number of instances of viewing at a given duration. Top: Graph presenting percentage of overall cumulative viewing and percentage of overall number of viewing instances.

In terms of how this viewing was accomplished in our multi-view display, Table 3 demonstrates that our *transition* behaviour was utilized for the majority of this viewing, with the *peek* gesture accounting for only ~5% (~32 seconds) worth of viewing on average. Given that the peek gesture was intended to allow quick and casual viewing of a partners activity, the lack of usage evidenced in Figure 4 suggests that this gesture, whilst utilized, was not sufficient for providing casual awareness.

	Viewing Mechanism		RM-Anova
	Transition	Peek	
Mean Total Viewing (SD)	566.8 (36.4)	32.9 (36.4)	$\chi^2(1) = 146, p < 0.01$
Mean Duration of Views (SD)	45.98 (36.3)	8.22 (18.3)	$\chi^2(1) = 13.5, p < 0.01$

Table 3. Mean (SD) viewing for Condition 3 (multi-view display) broken down by whether a transition or peek resulted in said view.

With respect to how likely users were to view or interact with (i.e. perform touchpad or textual actions on) either virtual view (see Table 2) there were no significant differences between Conditions 2 and 3. There was a bias toward equality with respect to interaction with the multi-view display, however this was likely due to the fact that once a user performed a transition in Condition 3, they were free to interact with the view they had transitioned to. In Condition 2, these transitions were typically managed by gaze, thus users would have to explicitly perform the transition gesture to then interact with this view. This suggests an interesting benefit of multi-view displays when coupled with touchpad remote controls: inputs can always be routed to the view the user is attending to. For a MDG system to accomplish this would require gaze tracking, a different input modality or additional effort on behalf of the user to manage which display they were interacting with, effort which the results of Condition 2 suggest users were unlikely to undertake.

We also asked about the acceptability of using a shared audio space. We found that being able to hear audio coming from both displays was less acceptable in the multi-view Condition compared to the two displays (Condition 2 mean=3.50 sd=2.15; Condition 3=2.61, sd=1.85, 0=Unacceptable, 6=Acceptable; no statistically significant difference).

Discussion

Our results demonstrated that a multi-view TV is preferable to a single-view TV, which is not entirely surprising: as much as we can design an interface for multi-user use, the physical bottleneck of having to share the display inevitably negatively affects performance. The comparison between our multi-view display and the two physical displays did however demonstrate some marked differences not in how well users perceived their ability to collaborate or gain awareness of each others activity, but in how this awareness was accomplished. The two physical displays in Condition 2 were used to facilitate a casual and continual awareness of the activity of the other participant, through a multitude of shorter glances at each display. In contrast, the multi-view condition featured much longer views of each virtual view. Whilst we attempted to facilitate the ability to gain casual awareness through the *peek* gesture, this difference in viewing behaviour suggests that casual awareness is more readily accomplished by gaze, and not through system functionality. Whilst having two displays is marginally preferable to multi-view, it is unlikely that this would be an acceptable configuration in the home, thus these results suggest we must design to accommodate for casual awareness.

STUDY 2: CASUAL AWARENESS IN MULTI-VIEW TV

The results of our first study raised a significant question. If perceived awareness and ability to collaborate was not significantly different between the two-display and multi-view conditions, but the way in which this awareness was accomplished was (with much shorter glances between displays), should we attempt to enable this more casual, continual gaze based awareness, and how? Incorporating continual and casual awareness necessitates a compromise with respect to distraction due to other user's activity. Some aspect of the user's physical view must be used to provide this awareness. This goes against one of the primary aims of our initial study, which was to develop a set of behaviours that would allow for management of multiple views whilst not compromising the users current physical view in terms of distraction, aspect ratio and utilized screen. To study this, we designed a system that could answer the following questions:

- How much of their physical view are users willing to sacrifice to gain a casual awareness of other virtual views?

- Given the ability to transition between a casual awareness mode and a fullscreen mode, how would users appropriate such a system? Would they rely on only one mode, or use both, and if so to what degree would they use both modes?

We designed two additions to our previous multi-view TV system, applying the concept of the casual–focused continuum [22] to awareness. The first was to give users the ability to vary their engagement with others by directly controlling how much of their personal physical view was given up to awareness of what is happening in virtual views other than that which they are currently interacting with (see Figure 5). This was accomplished through the use of a slider on the touchpad (see Figure 6). At its extremes, it would devote the majority of the user's physical view to either to the virtual view the user was interacting with, or the other available virtual view; as the slider moves to the center of the touchpad, the user's physical view would begin to be split evenly between both virtual views.

Figure 5. Example of two users in the dynamic split-screen mode, with different levels of engagement with each others activity. The user's currently interactive virtual view is always on the right of the physical view.

We anticipated that this mechanism could encompass a variety of behaviours, from selecting an appropriate ratio between the virtual views as a one-off, or repeatedly employing the slider to dynamically change the ratio between the virtual views as and when required, for example allowing users to be aware of a trailer their partner might be watching in the other virtual view. Through this, we hoped to establish if there were any norms with respect to how much of the physical view users were willing to give up for casual awareness. It is important to note that the aspect ratio of the content being viewed was preserved at all times, thus resulting in portions of the screen remaining unused, as can be seen in Figure 6.

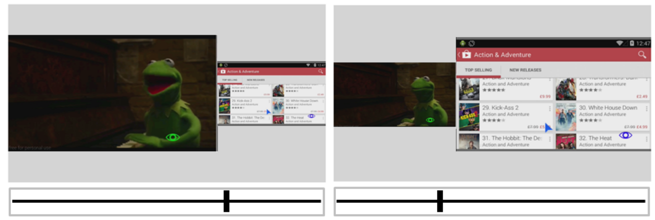

Figure 6. Example of the dynamic split-screen slider design. Here we see a user's physical view (shaded grey) being transformed Left: from a bias toward the currently non-interactive virtual view on the left; Right: to a new bias toward the interactive virtual view on the right.

The second addition was the ability to transition between this casual awareness mode and the fullscreen / fully-focused awareness mode that was the multi-view display in the previous study. As such, we incorporated a 3-finger tap gesture that would allow users to switch between the casual awareness mode, utilizing whatever screen ratio it was previously set at, and the fullscreen awareness mode. In both modes, the *transition* and *peek* behaviours functioned as before; in casual awareness mode, these actions resulted in the two virtual views swapping positions for that user.

Implementation And Experimental Design

The implementation was the same as the first study, aside from the two additional interactions. Transitions between modes, use of the slider and transitions between views were all animated, with changes to the slider affecting the rendering in real-time. Users could interact with only one virtual view at a time; this interactive view was always to the right of the user's screen, and signified with a grey border.

For this study, we had three Conditions. They were *(1) Multi-view display* which was the fullscreen multi-view display from the previous study, serving as a baseline for new iterations of our multi-view design; *(2) Dynamic Split-Screen Multi-view* which was a display that provided only the casual awareness mode; and *(3) Selective Multi-view* which provided users with the ability to switch between the modes from Conditions 1 and 2 using a 3-finger tap. As the aims of this study were primarily investigating how users would appropriate a system which supported both casual and fullscreen awareness behaviours, we chose not to counter-balance all Conditions. Instead, we counter-balanced with respect to Conditions 1 and 2, before moving on to Condition 3. This was done so that users received significant training with respect to using the fullscreen and casual awareness systems before using the dual-mode system in Condition 3. With respect to measures, all transitions between views and modes were logged in order to see both users default behaviour in each condition, and how they appropriated our selective multi-view system. The same task design and post-Condition questionnaires were utilized as from the previous study, with the addition of asking users how distracting they found their partner's activity and how in control they felt regarding awareness of their partner's activity. Users had access to the same set of applications as before. There were new 7 pairs of participants, 14 users in all (mean age=26.4, SD=3.3, 14 male) that again knew each other (friends, family etc.), recruited from University mailing lists.

Results

In terms of our questionnaire analysis from the previous study, we found that whilst the fullscreen Condition was often rated the poorest in terms of ability to collaborate, awareness, and distraction there were no significant differences between Conditions. Additionally, there were no significant differences with respect to workload or system usability.

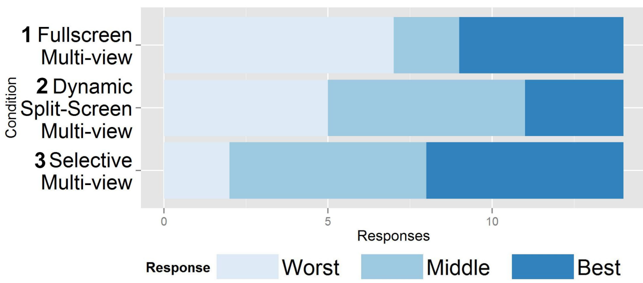

Figure 7. User ranking (lower is better) ordered by mean ranking - Friedman test $\chi^2(2)$ = 1.71 p = 0.42.

There were no significant differences with respect to user rankings (see Figure 7), however, there was a somewhat dichotomous split between users preferring either the selective mode or the fullscreen mode. Similarly, with respect to the proportion of viewing and interaction between the virtual views, there were no significant differences (see Table 4).

	Condition			
	1	2	3	RM-Anova
Interaction	0.73 (0.29)	0.65 (0.29)	0.75 (0.35)	$\chi^2(2) = 1.39, p = 0.5$
Viewing	0.47 (0.25)	0.57 (0.29)	0.55 (0.32)	$\chi^2(1) = 1.39, p = 0.5$

Table 4. Mean (SD) Gini coefficients for viewing and interaction. Gini coefficients show how likely users were to view or interact with virtual view. 1 is maximum inequality, 0 is maximum equality.

Casual vs. Fullscreen Awareness

Figure 9 details how the usage of our selective multi-view system compared to our comparative baselines. Here we see a surprisingly even split between behaviour usage in our selective multi-view system. Every capability, aside from the peek gesture, was utilized to a similar degree. Significantly, the most utilized function was our gesture for switching between fullscreen and dynamic modes. Transitions between virtual views occurred in both modes, however somewhat diminished in the dynamic mode, supplanted by use of the slider for enacting changes in screen ratio.

	Viewing Mechanism		
	Dynamic Mode	Fullscreen Mode	RM-Anova
Mean Total Viewing (SD)	206.0 (212.0)	274.0 (212.0)	$\chi^2(1) = 2.23, p = 0.136$
Mean Duration of Views (SD)	26.6 (33.3)	30.5 (34.2)	$\chi^2(1) = 0.291, p = 0.589$

Table 5. Viewing for Selective Multi-view display, broken down by whether the display was in Dynamic or Fullscreen mode.

Indeed users appeared to split their viewing between the Dynamic and Fullscreen modes relatively evenly, as evidenced in Table 5. In examining this split per user in Figure 8, we can see that the majority of users split their viewing time between modes equally. However, there were 3 users who somewhat favoured fullscreen mode and 3 who almost entirely favoured fullscreen mode.

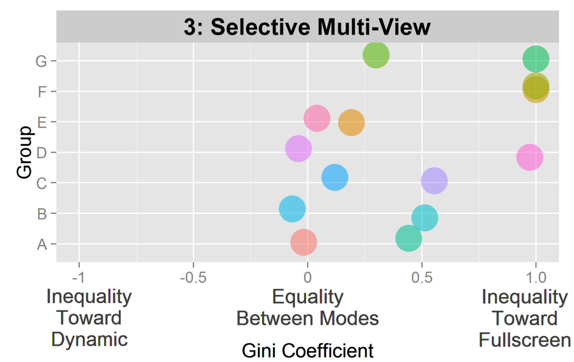

Figure 8. Directed Gini coefficient viewing in Dynamic and Fullscreen modes by group and participant (coloured) for Selective Multi-view. As an example, -1 indicates complete inequality toward the Dynamic mode, meaning users spent the entire duration in that mode. Jitter was added to Group axis in order to allow overlapping pairs to be differentiated.

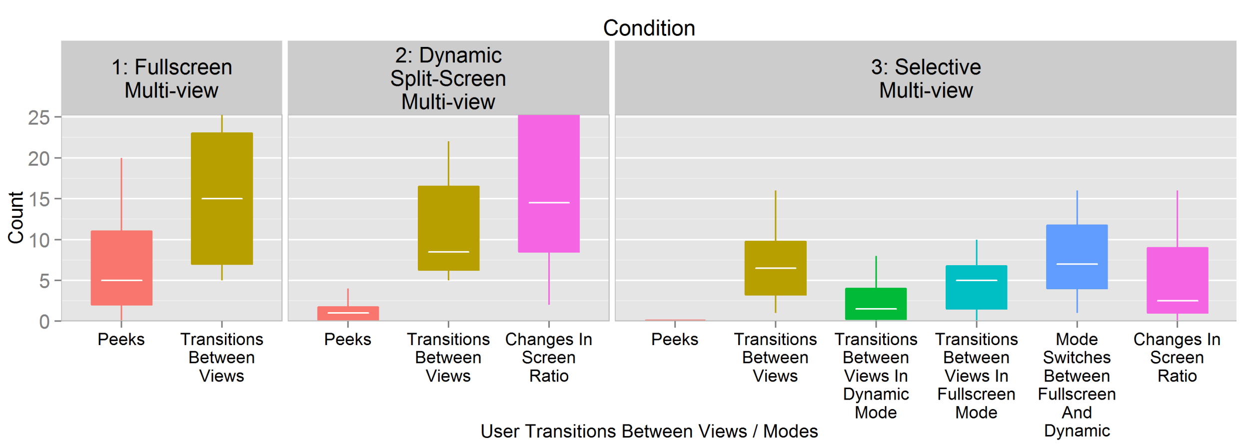

Figure 9. Boxplot of inter-quartile range of display management actions available to users: peeks (a non interactive look), transitions between views (moving between virtual views), changes in screen ratio (a slider manipulation), and mode switches between fullscreen and dynamic states.)

Usage of Casual Awareness Mode

Figure 10 visualizes the usage of the slider bar to show how much of the display the user was willing to dedicate to casual awareness. In Condition 2, we can see two clear peaks, meaning that users were typically moving between using ~8% and ~31% of the width of the display for casual awareness. In Condition 3, there was a much wider variety of usage, with peaks at 7%, 20%, 43%, 67% and 95%.

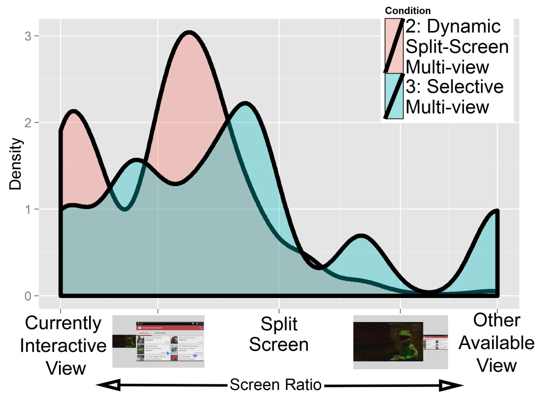

Figure 10. Kernel density plot of probability of distribution of slider values, determining the ratio by which the two virtual views are displayed. Left is biased toward the view the user is interacting with, right is biased toward the other available view, typically used by their partner. Condition 2 peaks at 8%, 31%**; Condition 3 at** 7%, 20%, 43%, 67%, 95%

Discussion

Our results indicate some interesting behaviours regarding how much of the display users were willing to allocate to awareness of others' activity. Users of the selective multi-view display dynamically varied awareness of their partners activity, the majority of the time dedicating between 7% and 43% of the display to this, but occasionally dedicating the majority of the display to awareness, whilst either retaining the ability to interact (the peak at 67%), or forfeiting interaction entirely by making the interactive view essentially non-visible (95%). We suggest that this approach could be used to determine empirically how much of a given display should be used for casual awareness (likely varying based on the physical properties of the display). However, given the dynamic usage exhibited it would be worthwhile to expose this functionality to users, if not in a continuous form then perhaps a discrete slider moving through derived ratios.

With respect to how users appropriated our selective multi-view system, our management behaviours were utilized in both casual and fullscreen / focused modes, with some users reporting that, in the fullscreen mode, having the ability to transition between views was conveniently like having a "previous channel" button. Notably, three users were entirely unwilling to use the dynamic mode, instead remaining in fullscreen mode for the duration of the Condition. This suggests that in a consumer multi-view system, the ability to transition between views without compromising the maximal rendering of content on the display is an important property. However, there is also significant value in incorporating the ability to be casually aware of the activity of other views, for example when performing independent activity but with some shared aspect such as video content.

GENERAL DISCUSSION & FUTURE WORK

Through our two studies we have demonstrated a viable design for a two-user, multi-view TV display. Our initial multi-view display was significantly better than the single shared display in terms of the ability to collaborate and operate independently, demonstrating a set of behaviours which allowed users to effectively share usage of the TV display whilst minimizing the impact on each others' physical view and capability to interact effectively. However, a viewing comparison between our multi-view display and an ideal awareness display using two TVs indicated significant differences in terms of how this awareness was accomplished, with much shorter casual glances occurring in the ideal case.

Given this, we iterated upon the design of our multi-view TV display, incorporating mechanisms to allow users to transition between casual and focused states, and dynamically determine their level of engagement when in a casual state. The usage of this "selective" multi-view system confirmed the importance of both modes, demonstrating that given the ability, users will transition between modes and vary their engagement with others' activity in both modes. In the fullscreen mode, engagement was varied through transition gestures, whilst in the casual awareness mode users dynamically varied their engagement through use of our view slider for controlling the amount of display given over to casual awareness.

With respect to future work, there are a number of interesting areas. Examining the usage and effect of multi-view displays across different collaborative tasks might have a significant impact on the results, and thus the contexts in which multi-view might prove most useful. There is also the issue of effectively using the physical display area when in casual awareness mode. Our approach involved scaling virtual views whilst preserving aspect ratio, resulting in parts of the screen being under-utilized. We would suggest that for certain types of content (e.g. video) it may be acceptable to truncate this content, presenting only a portion of the virtual view sufficient to provide awareness whilst maximizing usage of the screen. Similarly, there may be ways of communicating sufficient awareness in a more discrete and unobtrusive fashion e.g. textually.

There are additionally questions regarding scale and appropriation: in scaling the interactions up to support more than two views, and more than two users, how many views/users are manageable before the complexity undermines the benefits of such a TV? And what kind of social impact might such a display might have in the home. Would users transition from devices to the the more shareable TV for some subset of their activities? Answering this would require a longitudinal deployment and the availability of an HD, glasses-free multi-view display, however this paper does provide the foundations for such future work.

Finally, for multi-view to work across a breadth of media tasks, solutions will be required to either help users in managing a shared audio space (preventing potentially frustrating conflicts), or provide personal and private audio spaces in an acceptable manner (i.e. excluding previously used solutions such as in-ear headphones that impact socialization).

CONCLUSIONS

This paper has presented two studies iterating upon the design of an Android-based two-user multi-view TV. Through this process, we have established a set of functionality necessary for users to be able to operate independently and collaboratively using a multi-view TV. We suggest that given the findings of this paper, multi-view TVs should ideally support both transitions between views (and thus independent and shared activity), and transitions between focused fullscreen usage, and usage supporting casual awareness of other pertinent activity. Furthermore, there appears significant merit in giving users the ability to dynamically determine their requisite level of awareness based on their engagement with others' activities in this casual awareness state. This research demonstrates that multi-view TVs have the potential to supplement or supplant the secondary device usage that is now commonplace in the home, bringing interaction and activity back toward a shared-and-shareable focal point, the TV.

ACKNOWLEDGMENTS

This work was supported in part by Bang & Olufsen and the EPSRC. This publication only reflects the authors' views.

REFERENCES

1. Agrawala, M., et al. The two-user Responsive Workbench. In *Proc. SIGGRAPH '97*, ACM Press (Aug. 1997), 327–332.
2. Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction. In *Proc. ITS 2010*, ACM Press (2010), 121–130.
3. Brooke, J. SUS-A quick and dirty usability scale. *Usability evaluation in industry* (1996).
4. Dodgson, N. A. Multi-view autostereoscopic 3D display. In *Stanford Workshop on 3D Imaging*, Stanford University (2011).
5. Gross, T. Supporting Effortless Coordination: 25 Years of Awareness Research. vol. 22 (June 2013), 425–474.
6. Gross, T., et al. User-Centered Awareness in CSCW-Systems. *International Journal of Human-Computer Interaction 18*, 3 (2005), 323–360.
7. Kitamura, Y., et al. Interactive stereoscopic display for three or more users. In *Proc. SIGGRAPH '01*, ACM Press (Aug. 2001), 231–240.
8. Kreitmayer, S., Laney, R., Peake, S., and Rogers, Y. Sharing bubbles. In *Proc. UbiComp '13 Adjunct*, ACM Press (Sept. 2013), 1405–1408.
9. Kulik, A., et al. C1x6: A stereoscopic six-user display for co-located collaboration in shared virtual environments. In *Proc. SIGGRAPH Asia '11*, ACM (2011), 188:1–188:12.
10. Lissermann, R., et al. Permulin: mixed-focus collaboration on multi-view tabletops. In *Proc. CHI '14*, ACM Press (2014), 3191–3200.
11. Lucero, A., et al. Mobile collocated interactions: Taking an offline break together. *interactions 20*, 2 (Mar. 2013), 26–32.
12. Matusik, W., Forlines, C., and Pfister, H. Multiview user interfaces with an automultiscopic display. In *Proc. AVI '08*, ACM Press (May 2008), 363.
13. McGill, M., Williamson, J., and Brewster, S. A. How to lose friends & alienate people. In *Proc. TVX '14*, ACM Press (June 2014), 147–154.
14. McGill, M., Williamson, J., and Brewster, S. A. Mirror, mirror, on the wall. In *Proc.TVX '14*, ACM Press (June 2014), 87–94.
15. Morris, M. R. Collaborative search revisited. In *Proc. CSCW 2013*, ACM Press (2013), 1181–1192.
16. Morris, M. R., Lombardo, J., and Wigdor, D. WeSearch. In *Proc. CSCW 2010*, ACM Press (2010), 401–410.
17. Müller, J., Geier, M., Dicke, C., and Spors, S. The boomroom: Mid-air direct interaction with virtual sound sources. In *Proc. CHI '14*, ACM (2014), 247–256.
18. Nacenta, M. A., et al. The LunchTable. In *Proc. PerDis 2012*, ACM Press (2012), 1–6.
19. Newton, T. Digital Divorce: 24 per cent of couples watch on-demand TV in different rooms. recombu.com, 2013.
20. Ofcom. Communications Market Report, 2012.
21. Ofcom. Communications Market Report, 2013.
22. Pohl, H., and Murray-Smith, R. Focused and casual interactions: Allowing users to vary their level of engagement. In *Proc. CHI 2013*, ACM (2013), 2223–2232.
23. Seifert, J., et al. MobiSurf. In *Proc. ITS 2012*, ACM Press (2012), 51–60.
24. Shoemaker, G. B. D., and Inkpen, K. M. Single display privacyware. In *Proc. CHI '01*, ACM Press (2001), 522–529.
25. Terrenghi, L., et al. A taxonomy for and analysis of multi-person-display ecosystems. In *Personal and Ubiquitous Computing*, no. 8 (2009), 583–598.
26. Thinkbox. TV Together: a very social medium, 2013.
27. Tuddenham, P., et al. WebSurface. In *Proc. ITS 2009*, ACM Press (2009), 181–188.
28. Turkle, S. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, Inc., Jan. 2011.
29. Wallace, J. R., et al. Investigating teamwork and taskwork in single- and multi-display groupware systems. In *Personal and Ubiquitous Computing*, no. 8 (2009), 569–581.
30. Wallace, J. R., et al. Collaborative sensemaking on a digital tabletop and personal tablets. In *Proc. CHI 2013*, ACM Press (2013), 3345–3354.
31. You, W., et al. Studying vision-based multiple-user interaction with in-home large displays. In *Proc. HCC 2008*, ACM Press (2008), 19–26.