

# Focused and Casual Interactions: Allowing Users to Vary Their Level of Engagement

Henning Pohl<sup>1,2</sup>

FG Mensch-Computer-Interaktion  
University of Hanover,<sup>1</sup> Germany  
Henning.Pohl@hci.uni-hannover.de

Roderick Murray-Smith<sup>2</sup>

School of Computing Science  
University of Glasgow,<sup>2</sup> Scotland  
Roderick.Murray-Smith@glasgow.ac.uk

## ABSTRACT

We describe the *focused-casual continuum*, a framework for describing interaction techniques according to the degree to which they allow users to adapt how much attention and effort they choose to invest in an interaction conditioned on their current situation. Casual interactions are particularly appropriate in scenarios where full engagement with devices is frowned upon socially, is unsafe, physically challenging or too mentally taxing. Novel sensing approaches which go beyond direct touch enable wider use of casual interactions, which will often be ‘around device’ interactions. We consider the degree to which previous commercial products and research prototypes can be considered as fitting the focused-casual framework, and describe the properties using control theoretic concepts. In an experimental study we observe that users naturally apply more precise and more highly engaged interaction techniques when faced with a more challenging task and use more relaxed gestures in easier tasks.

## Author Keywords

Casual interaction; deliberate interaction; sensing; interaction techniques; peripheral interaction; foreground/background

## ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Theory and Methods, Interaction Styles*

## INTRODUCTION

The current generation of computers is designed under the assumption that users are fully engaged with them when interacting. This is especially true of mobile, touch-enabled devices, which require users to watch the screen, as the lack of non-visual feedback makes it hard to interact in any other way, forcing them to focus on the interaction.<sup>1</sup> However, such full engagement with their devices is not possible or desirable for users at all times, e.g., because users are physically, mentally, or socially inhibited. Users might at any given moment also be engaged in a different task (e.g. social interaction, safety-critical or work-related tasks), otherwise distracted, tired, or

<sup>1</sup>Users might memorize some common touch sequences but without tactile landmarks for support, will have high recall error rates [21].

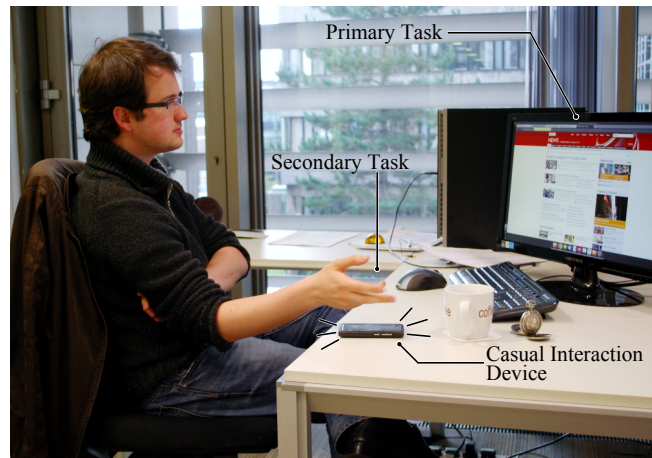


Figure 1. Daniel is sitting at the desk, reading an article, when his phone rings. He dismisses the call with a wave above the device as he does not feel like taking care of it now.

feeling lazy, reducing their ability and desire to focus on what the designer might consider the foreground interaction.

Instead of requiring users to be highly engaged with their devices, we propose a class of interaction techniques that allow users to choose the level of interaction depending on the situation, so that they can also choose to *casually* interact with them. Giving users the freedom to choose their desired level of engagement is empowering and has the potential to make their interactions more suitable for a given context (such as different social settings), as well as more relaxing, physically less demanding and safer to use.

In this paper, we provide an overview of casual interactions and how they enable users to be laid back, while retaining control. We relate this to examples of existing designs, and present an evaluation, showing how users, when given the opportunity to do so, take the chance to relax control when able to, while taking it back when required to do so by a task of higher difficulty or importance. We conclude with challenges to the research and design worlds in creating such focused-casual systems.

## CASUAL INTERACTION SCENARIO

Consider the user shown in Figure 1. He is busy reading an article online when his phone rings. While he is not sure who is calling, he still does not feel like answering. Because he is focused on his primary task, he chooses not to devote a lot of attention to his phone. He thus dismissively and inattentively waves over the device to signal his intent not to take the call. In this way he signals two things to his phone: (1) he is not particularly interested in

the call right now and (2) his dismissal is casual in nature. The second aspect transfers some powers to the phone. Had the user picked up the phone, glanced at the caller and then dismissed the call, it would be a clear indication that the call was not to be accepted. The casual dismissal, on the other hand, is more vague in execution, is based on missing information and therefore leaves some room for decision making to the phone. Imagine the user's boss is calling. In this case the phone could insist on continuing to ring, prompting the user to give the call a second look. Casual interaction enables users to engage more for fine control or engage less, either reducing the number of states they can drive the system to, or handing over more autonomy to their devices.

### CASUAL INTERACTION

Interface designers often assume that users focus on their device when interacting, but this is often not the case. In fact, there are many scenarios where users are not able to, or do not want to, fully engage with their devices. In general, inhibiting factors can be divided into (1) physical, (2) social, or (3) mental subgroups.

*Physical reasons* users can not fully engage with a device are often closely related to questions of accessibility. Users who can not use one arm (permanently or temporarily), for example, can not dedicate one hand to holding a device while the other interacts with it [39]. Thus, they can not achieve the same level of speed/accuracy as unencumbered users, and the engagement with their devices is limited by those constraints. In general, physical impairment inhibits our achievable accuracy, for example Keates et al. showed that movement error in mouse usage is larger for motion-impaired users [28]. In some cases, users can obtain physical relief, and overcome strain or fatigue by changing the level of interaction. For some temporary impairments, e.g. those caused by wearing gloves, users can choose themselves between giving up control (sometimes completely) or relieving the impairment (taking off the gloves) for a more full engagement (but cold hands). Casual interaction systems try to enable interaction even in such situations where users can not engage manually as needed in focused interaction.

*Social reasons* are mostly concerned with the question of how much engagement with a device is acceptable in a given setting. While taking a phone out of the pocket and checking stock quotes is perfectly acceptable when alone, this interaction would be considered rude when performed during a date. While social settings limit some interactions, they can also strengthen others. Users might, for example, be more inclined to perform interactions, that make them look relaxed and laid back. The desire to seem as if one is not putting in too much of an effort in order to be seen as *cool* has, e.g., been described for 14–16 year old students by Warrington et al. [54]. We postulate that casual interaction is desirable in social settings, because it can reinforce some users' desired image of being in control yet laid back and relaxed about it. The view of casual interaction in the social sphere fits well into the concepts of *front stage* and *back stage* by Goffman [15]. He notes that people change their behavior based

on what they image people around them are thinking about them—putting on an *everyday performance* when on the *front stage*. As people put on a character, they are more likely to engage in behavior that they see as potentially appealing. Note that casual interactions are, however, not necessarily more *subtle interactions*, as understood in [11], as the casual action might be more overt than a focused one.

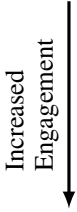
*Mental reasons* are primarily issues of distraction. Users might be engaged in a primary task, leaving little attention for other interactions. For example, Bakker et al. found that even primary tasks shifted back and forth between the center and the periphery of users' attention [5]. This suggests that there is room for interfaces enabling users to carry over this behavior for their interactions. Exhaustion after a day of work also leads to effort-avoiding behavior. In general, the capacity to make active choices declines over the day (a concept known as *ego-depletion* [6]), suggesting that users could especially benefit from casual interactions later in the day.

A commonplace in these scenarios is that the users still *want* to engage with their devices, just sometimes with lower effort and precision or in a more sporadic fashion. Casual interaction mechanisms should allow users to control the *level* to which they engage—they do not want to give up control completely. These are also not purely casual systems: the interaction space is continuous, spanning from very focused to very casual interactions.

### Levels of Engagement

When interacting with devices, the level of engagement with the device will differ depending on the situation (whether the current constraints were physical in nature or related to social or mental state).

Although the ranking will vary depending on the context and constraints, for interactions with smartphone-like devices a qualitative ranking of interactions by engagement could be:

- 
- Changes in the environment.
  - Presence of user in proximity.
  - Undirected waving above the device.
  - Turning the device on its face to silence a call.
  - Multiple finger pat on the device.
  - Controlled pointing on the device.

The high-engagement extreme describes very focused interactions, in which a fully committed user is giving her entire attention to the interaction, and actions and responses are tightly coupled in time. Playing games often falls in this category. On the other end of the scale are interactions that are of secondary or incidental nature. For example, muting an alarm clock by tapping it anywhere, or turning over a phone to reject a call can be done without giving the task too much attention. There can even be levels of casual interaction within otherwise highly focused settings – the popularity of kinetic scrolling in touchscreen devices is partly because after an initial flick, the user can pull back their finger, reducing their engagement level, and wait to see where they end before stopping the scrolling, or giving it a further impetus.

### Casual Interaction from a Control Theory Perspective

So far we have presented an informal description of the differences between focused, engaged interaction and casual interaction. We can make this difference more quantitative and objective by characterizing the casual–focused continuum in control theoretic terms, where we view the closed loop between user and computer as a dynamic system [25]. At every moment, humans are controlling a number of states [40], and this can be used to infer where they are paying attention, and what they are controlling [57]. Motor control studies have demonstrated that the precision of the response will depend on the uncertainty of the state information the human perceives (the more uncertain the human’s awareness of state, the smoother their responses tend to be) [31, 49].

In order to create an appropriate closed-loop behavior, designers can alter the sensing and feedback mechanisms in the phone to allow both—user and software—to adapt. This allows the user to control those aspects of their environment which appear most appropriate to the situation.

#### *Characterizing the Changes in Control*

*Focused interaction* will typically involve the user sampling the feedback frequently and accurately, and responding rapidly via accurate, high-bandwidth inputs to changing stimuli displayed on the computer and will take place over extended periods of continuous control. In information theoretic terms, the user will have a high channel capacity from their outputs to their senses, a measure called *empowerment* in [30], and we will see predictable responses to stimuli with a high level of consistency in the information rate of the interaction.

*Casual interaction* will tend to break or degrade elements of this control loop in a number of possible ways. The interactions may be *intermittent*<sup>2</sup>, i.e. the user perceives output from the device in a variable manner (e.g., some channels may be obscured by environmental disturbances, or are only occasionally actively sampled by the user) or only provides input occasionally, or subject to delays. The information flow in interaction is much more sporadic and unevenly distributed over time. There may be asymmetry in the interaction loop, where the user provides good quality input but has low-bandwidth output, or vice-versa, where the input is very noisy or uncertain, while the feedback is high-quality. If any of these features is present, the ability of the user to control the system decreases, and their channel capacity from output to input declines—they are less empowered with the specific aspect of the system examined.<sup>3</sup>

#### *Dealing with Reduced Control Levels*

Given the reduced control potential (fewer bits of input and output information and longer or varying delays) in the casual setting, there are different ways the joint system can function. Either (1) the system remains a direct control interface, but adapts such that there are fewer states that can be entered in a casual interaction mode, or (2) it moves to a higher-level of a hierarchical control system. In the hierarchical mode, user

<sup>2</sup>In control theory: discrete or non-proportional control [25].

<sup>3</sup>Note, however, that they are only less empowered with *this particular loop*. The decrease in attention required may be allowing them to attend to more important aspects of their environment.

inputs are changed to higher-level reference values which are achieved by lower-level controllers or agents.

In casual control systems, users voluntarily limit their input to either a reduced, or a more abstract level of control temporarily, while preserving their option to step in and take over at any given time. Flemisch et al. coined the term “*H-metaphor*”, using the example of a horse rider’s interaction with his horse as an analogy for the handover of autonomy in computer systems, as the certainty of control varies [14]. If a rider uses certain, frequent and deliberate movements, the horse follows movements exactly; as the control becomes more vague, the horse resorts to familiar behavior patterns and will take over more of the control. This notion of being able to ‘loosen or tighten the reins’ of interaction with an intelligent device is likely to be vital in the creation of future human–computer interaction systems, and is linked with, but not identical to, the casual–focused continuum. The ability to move effortlessly between casual and engaged interaction will be a keystone of the approach. This will depend on input sensors, output displays, inference of the meaning of user actions and interaction metaphors that users can grasp.

### Input Sensing Requirements

Currently, mobile devices mostly rely on touch sensing, while sometimes also offering motion-sensing or vision-based interaction. Interaction which depends on touch forces a degree of engagement which can be undesirable. Devices can be hard to reach at times or users can be inhibited from directly touching their devices for social acceptability reasons. Devices offering casual interaction modes will typically sense input from users over a wider range of modalities than in traditional, focused interaction. In this section we touch on a number of technologies we see as enabling for such devices.

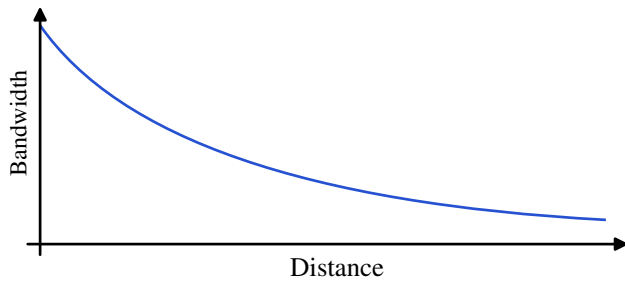
#### *Sensing technologies*

*Magnetic sensing* has been used to sense instrumented (e.g. wearing a ring) fingers for gestural interactions [29] and clicking [17]. *Nenya* in particular addresses this issue and offers the socially acceptable modality of twisting a tracked ring for analog input [3]. *Vision-based systems* sense users away from the device, and the technology can be integrated into laptop touchpads [9], diffuse illumination systems [41] or directly into screens [20], but tend to be sensitive to lighting changes in practical use.

*Capacitive sensing* is usually used for tracking fingers directly on a surface. However, fingers can still be sensed slightly off the surface, but those readings are noisier and thus less accurate, as demonstrated by Rogers et al. for touch (*Finger-Cloud* [44]) and full finger poses (*Anglepose* [45]). They also showed that by combining coarse long-range sensing with finer close-range sensing, capacitive sensors can be used to determine the 3D position of the finger on top of a device.

*Electromyographic sensing* is a less explored option, but has previously been used as well to detect coarse gestures above the surface [7]. Similarly, *electric field sensing*<sup>4</sup> is well suited for gesture recognition but less so for pointing. Casual interaction devices could also use the user’s *Bodyspace* [50] for limited bandwidth, but convenient interactions.

<sup>4</sup>Available, e.g., in Microchip Technology’s MGC3130 as *GestIC*.



**Figure 2.** The bandwidth of user interactions generally shrinks with increasing distance to devices. Users are, e.g., less precise in pointing at an object 10 m away than at an object 10 cm away. This dropoff also holds true for perceiving visual feedback. Additionally, most sensing methods deteriorate with increased distance to the sensor.

Devices tailored for casual interaction will have to cover a larger range at varying fidelities. Thus, such systems will often use a combination of the above mentioned sensing options, e.g., smart homes with instrumented living spaces can integrate information from a range of sensors, and on mobile devices a combination of capacitive sensing and IR distance sensors as an inexpensive and easy to build option. In *Hover-Flow* six such IR distance sensors provided a rough image of the space above the device and were shown to enable recognition of one-handed sweep and rotation gestures [32].

### Output feedback

In the same way that casual interactions benefit from sensing at a distance, the feedback modalities used can be structured such that users can be provided with feedback at a range of levels of attention. Combinations of tactile, audio and visual feedback can be used to provide the appropriate amount of information at appropriate frequencies, to allow the user to control the interactions at a suitable level of engagement. This will tend to go from lower-precision feedback for casual interactions, possibly with sampling delays, to higher resolution, more responsive feedback for focused interaction.

There are some existing feedback mechanisms that can be used in casual interactions, e.g., ambient and peripheral displays [35], or glanceable displays [47]. Such displays can also be incorporated into textiles for feedback on-the-go [53]. Qin et al. have added LEDs to the back of mobile devices in an effort to provide ambient lighting for extension of the display and interaction spaces [42]. In their *Damage* prototype, Williams et al. put LED-studs in a bracelet to facilitate group communication [55]. Users could make their corresponding studs on other bracelets glow in any color, enabling them to send non-disruptive and ambient messages. Harrison et al. investigated different light behaviors for a single LED as a means to communicate system state to a user [16].

Coupling input sensors to infer whether the user was able to sense the feedback provided by device outputs (e.g., was background noise such that they could hear the device, was display visible to user, etc.) will be an important feature for inference about how casual user responses are. As in the initial scenario, the meaning of a call rejection is different if the user knew who was calling or not.

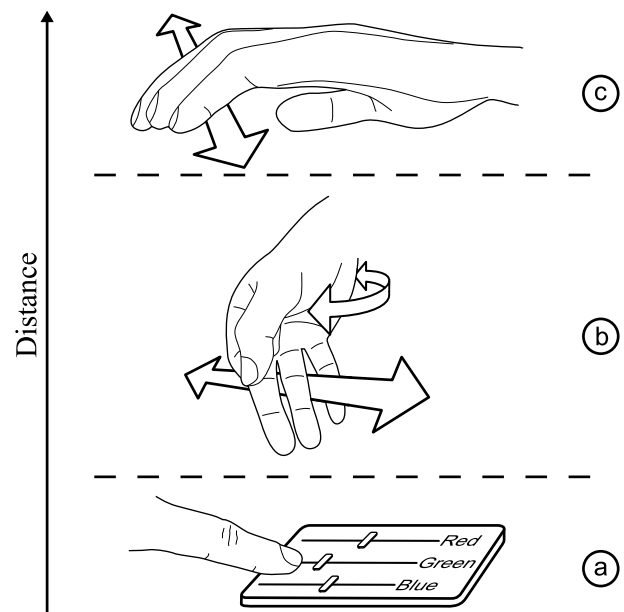
### Interaction Bandwidth

Depending on how users interact with a device, the available interaction bandwidth changes. Drawing on a touch-screen, for example, results in much more information (bits per second) than pressing a physical button. Bandwidth can be limited due to the sensing technology used, or due to user behavior. A bottleneck in the feedback channel (e.g. due to disturbance) can also affect the interaction bandwidth.

Capacitive sensing is an example of a sensing technology which provides high accuracy on the surface while offering reduced accuracy in above-the-surface sensing. In many other cases (but not all), distant interactions suffer from reduced sensing capability and increased user error (see Figure 2). Magnetic force, for example, behaves according to the inverse cube law, resulting in fast deterioration of sensing quality. In interaction on the surface additional opportunities for differentiation arise, e.g., by using sound features [37].

User behavior plays a role, e.g., as users have more control over movements they perform on a surface (i.e. with touch feedback) than in the air. For multi-layer interaction above the surface, [51] suggests that layer thickness should be around 4 cm for users to adequately be able to interact.

To enable users to interact with a system at different control levels requires adapting input handling. As introduced earlier, we can either reduce the degrees of freedom we control, or provide goals and autonomy to lower-level controllers. We see several general techniques for reducing bandwidth requirements for a given interface. These include: reducing field fidelity, spatial remapping, combining fields and dynamic combinations such as gestures. An example application of such techniques, where users could change the lighting of a mood light using different levels of input complexity, is shown in Figure 3.



**Figure 3.** Users are able to control the color of a moodlight by (a) precisely picking a color using sliders on the surface of the device, (b) manipulating the sliders together to control brightness coupled with slight rotation for changes in hue using above-the-device interactions, (c) in-air control of a one-dimensional hue slider at fixed brightness.

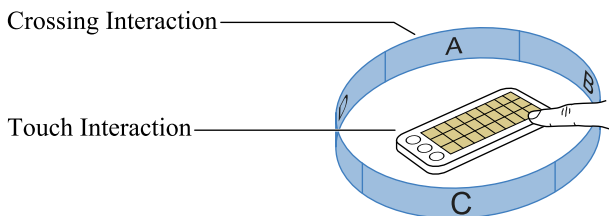


### Reducing Field Fidelity

Control of sliders on the surface allows for accurate parameter changes. If the precision for this kind of interaction is missing, e.g. above the surface, one can choose to reduce the number of available slider values. This preserves the direct control of sliders and only slightly limits the user.

Similar reductions can be applied to other kinds of widgets. In the area of text entry there already exist several techniques allowing users to type using *sloppy* or constrained input, such as *SHARK*<sup>2</sup> [33] or *SAK* [34]. For selection widgets (e.g. buttons) user input could be handled as probabilistic [56, 46].

*Spatial Remapping*: Sensing is not limited to the space directly above the device. Instead, sensing capability for many sensors extend in a rectangular frustum from the device. Future devices might include hemispherical sensing, extending the covered space even more. The shape of the sensing frusta results in an increase of covered sensing space when further away from the device. This allows to (a) fit more targets into the space or (b) make targets larger. As sensing accuracy also deteriorates with distance, increasing the target size seems more appropriate. Figure 4 shows an example of a target remapping (in combination with a field reduction), where 27 surface targets are mapped to 4 crossing targets in the air. For the 2D case, Accot and Zhai have shown that such a remapping (pointing to crossing) results in similar task times at lower error rates [2]. We hypothesize similar results for 3D remappings.



**Figure 4.** Device offering crossing and touch modalities. Due to the comparatively lower precision, the crossing interaction applies a *combining fields* approach as well.

*Combining fields*: In the example interface shown in Figure 3, interaction up close allows users to change a color by modifying each color component separately. While this provides precise control, this might not be necessary in all situations. Instead of setting a specific color, users often merely desire to set a *mood* in the range from gloom to vivid. Thus, the three-dimensional control of color components can be abstracted to a one-dimensional one that only controls mood. In effect, by combining fields, an interface tries to capture the *essence* of an interaction and provide an abstraction for it. Note that users can still choose to fully engage and get full control over the color sliders.

### Gestures

Gestures offer an alternative way to invoke device functions, trading spatial accuracy for dynamic consistency. *Whack gestures* are one example of making the main tasks of an interface available for inattentive invocation [23]. Other devices use the physical properties of the device form to enable casual

interactions, e.g. the Nokia X5-01<sup>5</sup> is a square phone with a curved base. When the device is spun on its base on a table, the music player triggers a ‘jukebox shuffle’ event—a playful, low-precision, low-effort gesture. In general, gesturing can be used to decouple user commands from the devices themselves. Systems exist that, e.g., instrument shoes [4] and clothing [43] for gesture sensing, and with the success of the Microsoft Kinect, vision-based sensing of device-free movement is becoming more prevalent.

### Agents—making lower-level systems autonomous

With the previous techniques, users still had direct control, if at a lower level of engagement. Users might instead choose to relinquish control even more and have a lower-level autonomous system take over. Instead of picking a color directly (high bandwidth), they would only indicate a desire for change—an action that can be communicated over a low bandwidth channel. Related examples include Syntonetic’s *Moodagent* app,<sup>6</sup> where a user can slide mood levels to generate playlists, rather than picking individual music tracks. By inferring a user’s engagement level, casual interaction systems could estimate the utility of invoking an agent vs. requiring more user input, similar to previous work of Horvitz [22].

### RELATED WORK

Casual interfaces build on existing work from areas such as ubiquitous computing, proxemics and gestural interaction. We present an overview of techniques that could be applied in a casual interaction system and previous work looking at varying-intensity interactions.

### Interactions in Hoverspace

Sensing users’ hands when they are not touching a device has been the basis for several previous interaction techniques. Spindler et al. investigated interaction with lenses in above-the-surface layers [48]. In particular, they showed the inversely proportional effect of layer thickness on task time. Earlier work on interaction layers by Subramanian et al. focused on pen interaction [51]. In particular, they looked at how to correct for drift and what kind of selection technique is most appropriate for in-the-air interaction. Building on that work, Kattinakere et al. show how Accot and Zhai’s steering law [1] applies to in-the-air interactions [27]. In *Mockup Builder* Araùjo et al. provide a concrete example of a continuous interaction space, combining touch on the surface with finger tracking in the air [12]. This, e.g., allows intuitive extrusion of 3d shapes by finger movement through the air. Hilliges et al. implemented pinching and grasping techniques for seamless interaction on and above a tabletop system [18]. Movement and shape of users’ hands in the space just above a device were monitored by Jackson et al. to enable manipulation of 3d objects [24].

While the above work investigates input of different modalities, none of them investigate those in the context of casual interactions. We build on some of the existing techniques, e.g., investigating hand gestures similar to the ones proposed by Kratz et al. [32].

<sup>5</sup> [http://www.developer.nokia.com/Devices/Device\\_specifications/X5-01/](http://www.developer.nokia.com/Devices/Device_specifications/X5-01/)

<sup>6</sup> <http://www.moodagent.com>

### Seamless Interaction

There has been some research into providing interfaces that allow for multiple ways to interact and seamless switching between those. Vogel and Balakrishnan outline the design principles for public displays that offer four interaction zones from ambient to personal interaction [52]. They note how interaction can adapt based on the level of attention of the user (e.g. angle of the user to the display). Their prototype also enables all interaction levels at any point, giving users the freedom to choose how much to engage. They look at engagement mostly as a function of distance, while casual interfaces have a user-centric view on engagement. Similarly, Clark et al. allowed users to interact with a display using touch from up close and pointing from further away [10]. They combined this with a zoomable interface approach, enabling fine grained selection up close with a more coarse range of options when at a distance.

While Vogel and Balakrishnan were interested in interactions with displays, we investigate interactions with personal portable devices, which have their own set of challenges. *AirFinger* enabled moving dragging and tracking operations from the surface to the space above the surface [59]. Yu et al. then associated different control-display (C/D) ratios to each layer, thus allowing for more control on the device and more sensitivity above. Their work is limited to dragging and tracking though, while we look at a wider range of inputs.

A comprehensive overview of techniques for on and above the surface interaction was provided in [36]. While they included a large number of techniques, they did not look at how those techniques could support casual interactions.

### Levels of Interaction

When thinking about different levels of interaction, we are building upon Buxton's foreground/background model [8]. Buxton makes a distinction between intentional interactions (in the foreground) and those that happen in the periphery of a user's attention (in the background).

Hinckley et al. investigate how sensors in mobile devices can be used to implement foreground and background interactions [19]. Here background interactions are those performed without much attention or on-the-fly, for example, automatically rotating the screen content when a device rotation is detected. The distinguishing characteristic between the two is the level of users' direct attention to the interaction.

Dix looks at background interactions that happen unbeknownst and secondary to an intended primary interaction, which he refers to as incidental interactions [13]. In his framework, incidental interactions, as users become experience with them, eventually transition to expected and finally intended interactions. It is thus the users' experience that determines the level of user intent, instead of users' freely choosing their desired level of engagement as in casual interactions.

Ju et al. extend the foreground/background model with a measure of initiative for their implicit interaction framework [26]. Around the example of the proactive *Range* whiteboard system, they present three implicit interaction techniques that help users keep in control of the proactive behavior.



Figure 5. (a) Ceiling mounted Kinect and (b) SHAKE SK7-ExtCS1 capacitive sensing extension board used for the study prototype. This setup allowed for the necessary on and above the device sensing.

Olivera et al. investigated how tangibles help users perform secondary interaction tasks [38]. They showed that such peripheral interactions are less taxing with tangibles than with GUIs. Wolf et al. looked at how microinteractions can support secondary, less demanding, tasks while users are engaged in a different primary one [58]. In contrast to the other works mentioned above, here two simultaneous foreground interactions are used instead of a mix of the two grounds.

While all those previous systems make a distinction between foreground and background interaction, this is mostly seen as a binary choice. The idea of background interaction assumes that the interaction is implicit and happens with no or, at most, only slight intention of the user. In this framework the transitions between background and foreground are often made by the systems themselves, e.g. by popping up alert windows. Instead, casual interaction is all about the users' choice of the level of engagement. Casual interaction is never fully in the background, but also is not the primary foreground activity. The users' choice dictates the level of ground an interface provides.

### STUDY: USERS VARY ENGAGEMENT WITH TASK DIFFICULTY

We now present a small user study to explore and evaluate one basic principle of casual interactions: when given the freedom to pick the level of engagement, a user's choice will be influenced by the level of control needed. We have developed a prototype in line with the aspects touched upon in the Input Sensing Requirements section for this evaluation.

### Prototype

For the qualitative evaluation of casual interactions, we have designed a prototype (shown in Figure 5) that allows for above-the-device interaction as well as precise touch interaction. A SHAKE SK7 sensor pack with SK7-ExtCS1 capacitive sensing extension board from SAMH Engineering Services is used as the foundation. It provides 24 capacitive sensor pads (each 7 mm<sup>2</sup> in size) in a 4 × 6 arrangement (57 mm × 36 mm overall). Those sensors provide accurate on-the-device touch tracking and also can be used to track

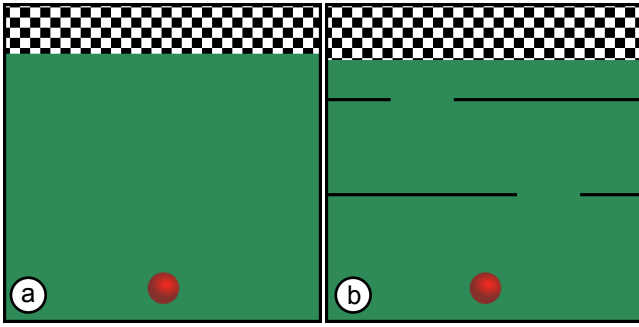


Figure 6. Examples of an (a) easy and (b) harder level used in the evaluation.

input up to 1 mm away from the surface. In the future, combining different pad sizes could enable sensing at much larger distances, while preserving the capability for accurate on-the-surface touch tracking [44]. Future devices could also incorporate range sensors for coarse above-the-device tracking [32]. In our current prototype we use a Kinect sensor to emulate such future devices. The Kinect is mounted on the ceiling, looking down on the sensor pack. We use computer vision techniques to extract the hand from the depth image by thresholding and connected-components analysis. Sensor data is combined on a PC, mapping the Kinect data to the coordinate space of the device.

### Study

In our discussion of casual interaction we have so far assumed that people are interested in letting go during interactions sometimes. We designed an experiment to test this, by observation of peoples behavior while playing a game. We hypothesize that with increasing index of difficulty, players will take more control of the interaction, while being more casual in easier levels.

To test this we use an obstacle course navigation type game that requires players to steer a ball to a target area. On the way, players need to stay away from the walls. If contact with a wall is made, the ball is reset to the starting position. Figure 6 shows two example levels with different indices of difficulty. The ball in the game has momentum and is simulated using a physics engine.

We compute the index of difficulty for each level according to Accot and Zhai’s steering law [1]. For navigation along a curved path  $C$ , parameterized by  $s$  and with width changing as  $W(s)$ , the path’s index of difficulty is thus defined as:

$$ID_C = \int_C \frac{ds}{W(s)}$$

While there is a linear relationship of the index of difficulty to task time, this aspect was not of interest for this experiment. We did not ask participants to be fast, nor did we show a timer or any ranking based on time. This was done to prevent participants from trying to finish levels as fast as possible. Instead of changes in task time, we were interested in changes of engagement depending on the varying index of difficulty.

For control of the ball, participants had to use our casual interaction prototype system (shown in Figure 5). This allowed them to exercise three levels of control:

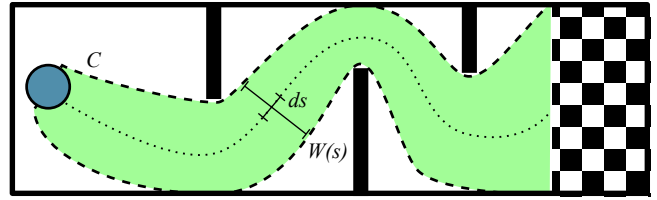


Figure 7. The index of difficulty for a game is defined as the integral of the reciprocal of the width along the required path.

**Touch** allowed for the most precise control. After touching the surface, any finger movement directly results in a corresponding ball translation (at a low  $C/D$  ratio). By using clutching, the ball can be moved to any point in the level.

**Hover** provides a higher  $C/D$  ratio rate control mechanism. Here users move their hand in the space just above the device. The zero point is set to the vertical axis through the center of the device. Deviation from this center results in a force of proportional amplitude being applied to the ball.

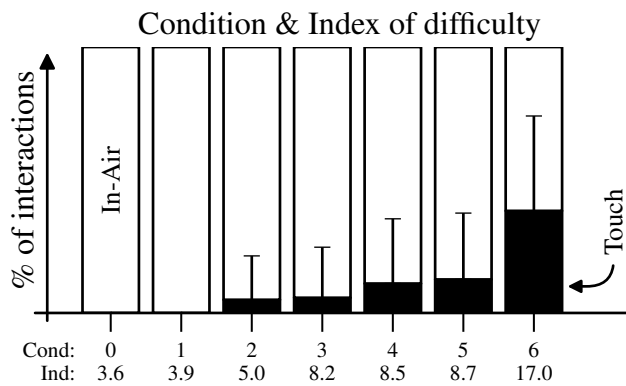
**Gestures** can be used further atop the device, beyond the range of hover interactions. Here, users indicate a general *direction* they would like the ball to move in. This is the least precise of all three methods. Execution of a gesture results in an impulse of constant strength and in the indicated direction being applied to the ball.

While the direct touch mode allowed for more precise control, it required participants to closely engage with the device. The hover option on the other hand is less precise, but requires no such close engagement. Finally, the gesture option allows participants to potentially “solve” a level with one ballistic action, but it is also the least precise option available—offering no feedback control ability.

We recruited 12 participants (2 female, age 24–48  $\bar{x} = 29.75$ ,  $\sigma = 6.37$ ) from Glasgow University’s School of Computing Science. They were given a chance to familiarize themselves with the game in a special training level. Participants played freely until they felt confident in their ability to control the game and switch between techniques. Afterwards, they played through a randomized sequence of 7 levels. Participants were instructed to decide themselves how much they wanted to engage with the game at any given situation. They were also assured that obstacle collisions were not a problem in this study and that they would not be evaluated based on their performance. As mentioned earlier, we made sure they did not feel the need to finish a level as fast as possible.

We find that for low difficulty levels, participants exclusively used hover and gesture control (see Figure 8). Only at the highest index of difficulty (corresponding level shown in Figure 6b), was a noticeable number of interactions performed by direct touch control. Note, though, that *hover* can be regarded as the ‘default’ mode. When a user’s hand is near the device this registers as a hover interaction. Most users occasionally let go of control and allowed the ball to move around, as there was no penalty for obstacle collisions. The ratio of touch interactions shown here should therefore be regarded as a lower bound.

Looking at the number of times participants used gesture control for a level (shown in Figure 9), we see a similar



**Figure 8.** The study showed an increase in the share of direct touch interaction for higher indices of difficulty. Ratios are computed per participant and averaged over all participants. Error bars correspond to 1 std dev, y-axis from 0% to 100%.

trend with the number of gesture invocations declining with increasing index of difficulty. Condition 5 here is somewhat of an outlier, because even though the path to the target area was tight, a direct vertical movement was sufficient to reach the target. While participants often took several attempts to get the direction right, most succeeded in reaching the target that way. This indicates that Accot and Zhai's steering law, while a good initial estimate, is not sufficient for capturing the actual level of difficulty for the set of input techniques used here. It is also interesting to note that for a case such as condition 5, if users can get away with multiple attempts at a low-effort gesture, they appeared to prefer this over a higher engagement interaction style.

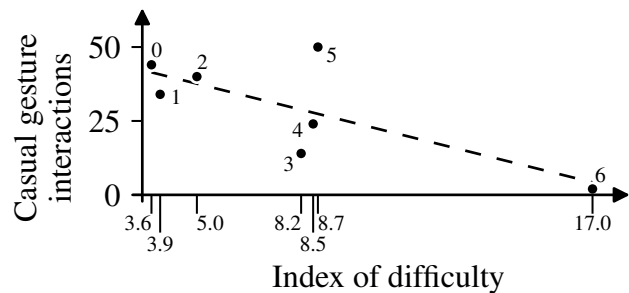
Nonetheless, participants are adapting their behavior according to the posed level of difficulty. This supports one basic premise of the casual–focused continuum: users are willing to fall back to relaxed input if the level of control is still sufficient to achieve their desired goal.

#### CHALLENGES FOR DESIGNERS AND RESEARCHERS

We have presented a conceptual framework for thinking about how interfaces could be placed on a focused–casual continuum. When designing for casual interactions, designers have to think about a number of aspects. They will have to, e.g., decide (1) Whether the levels of interaction will be on a continuum, or a set of discrete levels. If discrete, how many levels of interaction they want to support? (2) What set of interactions is appropriate for a given device, and how to develop interaction metaphors which can stretch across focus levels? How do you support control when you can reach smaller state spaces with decreasing engagement level, while having a consistent user model of the system? (3) How different levels of engagement can be differentiated? (4) What mechanisms could facilitate fluid switching between casual and engaged interaction, can these be made robust?<sup>7</sup> (5) How will we introduce casual systems to users successfully?

In research, a wide number of questions around casual interaction open up. (1) Can we propose concrete quantitative

<sup>7</sup>Around-the-device interaction leads to an elevated risk of unintended activation. The increase in active volume means there is a higher chance generating unintended interactions. Casual interaction systems will have to include safeguards tuned to the specific system to prevent such accidental activations.



**Figure 9.** The study showed a decrease of gesture usage for higher indices of difficulty. Data points show gesture usage aggregated for all participants and condition number.

metrics, based on levels and variability of information rates, which would allow us to classify a given interaction mode as focused or casual? Could this be used to provide measures throughout the interaction in bits/s? (2) Should we add questions about casual interaction to subjective user feedback questionnaires? (3) Can we demonstrate that allowing users to vary their engagement level leads to an improved user experience? (4) What methods are appropriate for investigating the social aspects of casual interaction? How culturally and device-specific are these likely to be?

We will illustrate these issues with the example of TV viewing, where users are not physically inhibited, but strong mental and social factors are at play. We could easily imagine, e.g., that after a day at work, a user is less inclined to make active choices and is more willing to give up some control. Control in this scenario is mostly relevant in the context of channel choice and volume control, although with increasingly sophisticated Internet TV, such as BBC's iPlayer, Netflix or Apple TV, more complex interaction becomes necessary. Current remotes actually already offer something along the lines of our engagement concepts. Some remotes offer a 'complicated' panel, which can be hidden, leaving simpler interaction possibilities. Furthermore, users can jump to a channel immediately using a number pad. However, this comes at a higher selection cost (smaller, less distinct buttons) and might require multiple button presses. Instead, they can use increment and decrement buttons, stepping through the channels at a lower engagement cost.

In this case, designers need to decide whether anyone can interact with any of the remote controls, or with the TV itself, at a distance. How many levels would be appropriate (and easy to explain) in this context? How do we develop interaction metaphors which can stretch across focus levels? Casual interaction devices could extend this by giving a simplified interaction when the user is not pressing remote buttons. They could have had swipes for channel changes, but no direct control of channel numbers, or could extend the swipe metaphor to allow for larger channel steps based on the intensity of a swiping gesture. If the system had a user model from prior behavior, and could recognize the user, then either the system could move to an autonomous recommendation mode, where swipes just corresponded to 'next recommendation', or the user could stay in direct control, but given a likelihood function for each channel, the control space could be distorted to make likely channels easier to acquire when only low bandwidth input is provided. How would the feedback



mechanisms change for different engagement levels? Would this make clear to the users why the channel changed, when one of them was waving his hands about something else?

As researchers investigating such a proposed system, we would like to be able to measure the interactions in such a case and quantify the level of focus needed to use the system. When performing user studies we would need to be able to relate subjective feedback and the objective level of focus displayed in user behaviors during the study. We would also want to see how easily the user could control the level of engagement at any point in time, and whether this itself led to frustration. Studies of social aspects of the system should include the changes in overt behavior when multiple people are in the room compared to a single user. It will be interesting to see whether making the engagement level 'socially attentive' would be a benefit, e.g. after one person speaks, the system might be primed to increase the likelihood of interpreting a hand gesture as a volume down request.

### CONCLUSION

The ability for users to elegantly adapt their level of engagement with a computer, in a context-sensitive manner, is going to be of increasing relevance in a world populated by many interactive systems. Especially when analyzing the design of multimodal, interactive systems with around-device interaction potential. This paper proposes designing systems explicitly to support *casual interaction*, which enables users to decide *how much* they would like to engage with interactive systems. We used a control-theoretic framework to describe the features of lower levels of engagement with a system (intermittent sampling of inputs and outputs, variability in delays in responding, lower-accuracy input and output, and the relationship to hierarchical control systems). We illustrated an aspect of user adaptation to interaction difficulty in an experiment which allowed users to achieve their goals at a range of levels of engagement, providing some initial, exploratory results showing that users do indeed change their engagement depending on the difficulty of the task at hand.

We hope that this paper will encourage designers to consider how they can instrument devices and design interactions such that users can choose to vary their level of engagement. We anticipate that it will also lead to new developments in measures of interaction in the HCI research community.

### ACKNOWLEDGEMENTS

Nokia's *Global Donations* scheme provided financial support. Julie and John Williamson, and Antti Oulasvirta provided valuable feedback.

### REFERENCES

1. Accot, J. and Zhai, S. Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. *Proc. CHI '97*. 295–302.
2. Accot, J. and Zhai, S. More than dotting the i's - Foundations for crossing-based interfaces. *Proc. CHI '02*, 4. 73–80.
3. Ashbrook, D., Baudisch, P., and White, S. Nanya: Subtle and Eyes-Free Mobile Input with a Magnetically-Tracked Finger Ring. *Proc. CHI '11*. 2043–2046.
4. Bailly, G., Müller, J., Rohs, M., Wigdor, D., and Kratz, S. ShoeSense: A New Perspective on Gestural Interaction and Wearable Applications. *Proc. CHI '12*. 1239–1248.
5. Bakker, S., van den Hoven, E., and Eggen, B. Acting by hand: Informing interaction design for the periphery of people's attention. *Interacting with Computers*, 24(3), 2012:119–130.
6. Baumeister, R.F., Bratslavsky, E., Muraven, M., and Tice, D.M. Ego Depletion: Is the Active Self a Limited Resource? *Journal of Personality and Social Psychology*, 74(5), 1998:1252–1265.
7. Benko, H., Saponas, T.S., Morris, D., and Tan, D. Enhancing Input On and Above the Interactive Surface with Muscle Sensing. *Proc. ITS '09*. 93–100.
8. Buxton, B. Integrating the Periphery and Context: A New Taxonomy of Telematics. *Proc. GI '95*. 239–246.
9. Choi, S., Han, J., Kim, S., Heo, S., and Lee, G. ThickPad: A Hover-Tracking Touchpad for a Laptop. *Proc. UIST '11*. 15–16.
10. Clark, A., Dünser, A., Billingham, M., Piumsomboon, T., and Altimira, D. Seamless Interaction in Space. *Proc. CHI '11*. 88–97.
11. Costanza, E. *Subtle, intimate interfaces for mobile human computer interaction*. Ph.D. thesis, Citeseer, 2006.
12. De Araujo, B.R., Casiez, G., and Jorge, J.A. Mockup Builder: Direct 3D Modeling On and Above the Surface in a Continuous Interaction Space. *Proc. GI '12*. 173–180.
13. Dix, A. Beyond intention - pushing boundaries with incidental interaction. *Proceedings of Building Bridges: Interdisciplinary Context-Sensitive Computing*. 1–6.
14. Flemisch, O., Adams, A., Conway, S.R., Goodrich, K.H., Palmer, M.T., and Schutte, P.C. The H-Metaphor as a guideline for vehicle automation and interaction. Tech. Rep. NASA/TM-2003-212672, NASA, 2003.
15. Goffman, E. *The Presentation of Self in Everyday Life*. Anchor, New York, NY, 1959.
16. Harrison, C., Horstman, J., Hsieh, G., and Hudson, S. Unlocking the Expressivity of Point Lights. *Proc. CHI '12*. 1683–1692.
17. Harrison, C. and Hudson, S.E. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. *Proc. UIST '09*. 121–124.
18. Hilliges, O., Izadi, S., Wilson, A.D., Hodges, S., Garcia-Mendoza, A., and Butz, A. Interactions in the Air: Adding Further Depth to Interactive Tabletops. *Proc. UIST '09*. 139–148.
19. Hinckley, K., Pierce, J., Horvitz, E., and Sinclair, M. Foreground and Background Interaction with Sensor-Enhanced Mobile Devices. *ACM Transactions on Computer-Human Interaction*, 12(1), 2005:31–52.
20. Hirsch, M., Lanman, D., Holtzman, H., and Raskar, R. BiDi Screen: A Thin, Depth-Sensing LCD for 3D Interactions using Light Fields. *ACM Transactions on Graphics*, 28(5), 2009:159:1–159:9.
21. Hoggan, E., Brewster, S.A., and Johnston, J. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. *Proc. CHI '08*. 1573–1582.

22. Horvitz, E. Principles of Mixed-Initiative User Interfaces. *Proc. CHI '99*, May. 159–166.
23. Hudson, S.E., Harrison, C., Harrison, B.L., and LaMarca, A. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. *Proc. TEI '10*. 109–112.
24. Jackson, B., Schroeder, D., and Keefe, D.F. Nailing Down Multi-Touch: Anchored Above the Surface Interaction for 3D Modeling and Navigation. *Proc. GI '12*. 181–184.
25. Jagacinski, R.J. and Flach, J.M. *Control Theory for Humans: Quantitative approaches to modeling performance*. Lawrence Erlbaum, New Jersey, 2003.
26. Ju, W., Lee, B.A., and Klemmer, S.R. Range: Exploring Implicit Interaction through Electronic Whiteboard Design. *Proc. CSCW '08*. 17–26.
27. Kattinakere, R.S., Grossman, T., and Subramanian, S. Modeling Steering within Above-the-Surface Interaction Layers. *Proc. CHI '07*. 317–326.
28. Keates, S., Hwang, F., Langdon, P., Clarkson, P.J., and Robinson, P. Cursor Measures for Motion-Impaired Computer Users. *Proc. Assets '02*. 135–142.
29. Ketabdard, H., Roshandel, M., and Yüksel, K.A. Towards Using Embedded Magnetic Field Sensor for Around Mobile Device 3D Interaction. *Proc. MobileHCI '10*. 153–156.
30. Klyubin, A., Polani, D., and Nehaniv, C. Empowerment: A Universal Agent-Centric Measure of Control. *IEEE Congress on Evolutionary Computation '05*. 128–135.
31. Körding, K.P. and Wolpert, D.M. Bayesian decision theory in sensorimotor control. *TRENDS in Cognitive Sciences*, 10(7), 2006:319–326.
32. Kratz, S. and Rohs, M. HoverFlow: Expanding the Design Space of Around-Device Interaction. *Proc. MobileHCI '09*. 4:1–4:8.
33. Kristensson, P.O. and Zhai, S. SHARK<sup>2</sup>: A Large Vocabulary Shorthand Writing System for Pen-Based Computers. *Proc. UIST '04*. 43–52.
34. Mackenzie, S. and Felzer, T. SAK: Scanning Ambiguous Keyboard for Efficient One-Key Text Entry. *ACM Transactions on Computer-Human Interaction*, 17(3), 2010:11:1 – 11:39.
35. Mankoff, J., Dey, A.K., Hsieh, G., Kientz, J., Lederer, S., and Ames, M. Heuristic Evaluation of Ambient Displays. *Proc. CHI '03*. 169–176.
36. Marquardt, N., Jota, R., Greenberg, S., and Jorge, J.A. The Continuous Interaction Space: Interaction Techniques Unifying Touch and Gesture On and Above a Digital Surface. *Proc. INTERACT '11*. 461–476.
37. Murray-Smith, R., Williamson, J., Quaade, T., and Hughes, S. Stane: synthesized surfaces for tactile input. *Proc. CHI '08*. 1299–1302.
38. Olivera, F., García-Herranz, M., Haya, P.A., and Llinás, P. Do Not Disturb: Physical Interfaces for Parallel Peripheral Interactions. *Proc. INTERACT '11*. 479–486.
39. Oulasvirta, A. and Bergstrom-Lehtovirta, J. Ease of Juggling: Studying the Effects of Manual Multitasking. *Proc. CHI '11*. 3103–3112.
40. Powers, W.T. *Behavior: The control of Perception*. Aldine de Gruyter, 1973.
41. Pyryeskin, D., Hancock, M., and Hoey, J. Extending Interactions into Hoverspace Using Reflected Light. *Proc. ITS '11*. 262–263.
42. Qin, Q., Rohs, M., and Kratz, S. Dynamic ambient lighting for mobile devices. *Proc. UIST '11*. 51–52.
43. Rekimoto, J. GestureWrist and GesturePad: unobtrusive wearable interaction devices. *Proc. ISWC '01*. 21–27.
44. Rogers, S., Williamson, J., Stewart, C., and Murray-Smith, R. FingerCloud: Uncertainty and Autonomy Handover in Capacitive Sensing. *Proc. CHI '10*. 577–580.
45. Rogers, S., Williamson, J., Stewart, C., and Murray-Smith, R. AnglePose: Robust, Precise Capacitive Touch Tracking via 3D Orientation Estimation. *Proc. CHI '11*. 2575–2584.
46. Schwarz, J., Mankoff, J., and Hudson, S.E. Monte Carlo Methods for Managing Interactive State, Action and Feedback Under Uncertainty. *Proc. UIST '11*. 235–244.
47. Sellen, A., Eardley, R., Izadi, S., and Harper, R. The whereabouts clock: early testing of a situated awareness device. *Proc. CHI '06*, CHI EA '06. 1307–1312.
48. Spindler, M., Martsch, M., and Dachsel, R. Going Beyond the Surface: Studying Multi-Layer Interaction Above the Tabletop. *Proc. CHI '12*. 1277–1286.
49. Strachan, S. and Murray-Smith, R. Bearing-based selection in mobile spatial interaction. *Personal and Ubiquitous Computing*, 13(4), 2009:265–280.
50. Strachan, S., Murray-Smith, R., and O'Modhrain, S. BodySpace: Inferring body pose for natural control of a music player. *Proc. CHI EA '07*. 2001–2006.
51. Subramanian, S., Aliakseyeu, D., and Lucero, A. Multi-Layer Interaction for Digital Tables. *Proc. UIST '06*. 269–272.
52. Vogel, D. and Balakrishnan, R. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. *Proc. UIST '04*. 137–146.
53. Wakita, A. and Shibutani, M. Mosaic Textile: Wearable Ambient Display with Non-emissive Color-changing Modules. *Proc. ACE '06*. 48:1–48:7.
54. Warrington, M., Younger, M., and Williams, J. Student Attitudes, Image and the Gender Gap. *British Educational Research Journal*, 26(3), 2000:393–407.
55. Williams, A., Farnham, S., and Counts, S. Exploring Wearable Ambient Displays for Social Awareness. *Proc. CHI EA '06*. 1529–1534.
56. Williamson, J. *Continuous Uncertain action*. Ph.D. thesis, Dept. of Computing Science, University of Glasgow, 2006.
57. Williamson, J. and Murray-Smith, R. Pointing without a pointer. *ACM SIG CHI*. 1407–1410.
58. Wolf, K., Naumann, A., Rohs, M., and Müller, J. A Taxonomy of Microinteractions: Defining Microgestures Based on Ergonomic and Scenario-Dependent Requirements. *Proc. INTERACT '11*. 559–575.
59. Yu, C., Tan, X., Shi, Y., and Shi, Y. Air Finger: Enabling Multi-scale Navigation by Finger Height above the Surface. *Proc. UbiComp '11*. 495–496.