Designing and Theorizing Co-Located Interactions

Thomas Reitmaier
Dept. of Computer Science
University of Cape Town
treitmaier@gmail.com

Pierre Benz *
Dept. of Computer Science
University of Cape Town
benz.pierre@gmail.com

Gary Marsden
Dept. of Computer Science
University of Cape Town
gaz@cs.uct.ac.za

* Note: Benz is joint-first author with Reitmaier. Please abbreviate citation to Reitmaier & Benz *et al*.

ABSTRACT
This paper gives an interwoven account of the theoretical and practical work we undertook in pursuit of designing co-located interactions. We show how we sensitized ourselves to theory from diverse intellectual disciplines, to develop an analytical lens to better think about co-located interactions. By critiquing current systems and their conceptual foundations, and further interrelating theories particularly in regard to performative aspects of identity and communication, we develop a more nuanced way of thinking about co-located interactions. Drawing on our sensitivities, we show how we generated and are exploring, through the process of design, a set of co-located interactions that are situated within our social ecologies, and contend that our upfront theoretical work enabled us to identify and explore this space in the first place. This highlights the importance of problem framing, especially for projects adopting design methodologies.

Author Keywords
co-located interaction; co-presence; design research; theory

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Communication is about humanness. It is what we do, who we are, and what makes us into a society. Because communication is so central to our being we take to new media and communication channels, as they give us new or different means of doing what we love to do: expressing ourselves and telling our stories. But not all views of communication are equal; Harper critiques a pervasive vision of the communicating human, which many researchers within CHI and ubicomp have adopted – the one that emphasizes the mechanics of human communication and can "overlook the humans who are doing the communicating" [13,p.49]. Here, communication acts are thought of as messages that are transferred through a medium — be it face-to-face or on a phone — from one person to another. To be sure, this model has proven generative potential, but the deeper problem, as Harper alerts us, is that "the vision used to orient design is of a world that is not the same as the one real people populate" [13, p.240]. Instead, we need to figure "communication between people [as] a performance that ties people together (or throws them apart) in various ways" [13, p.247]. Seeing communication as a performance highlights that it is not just about *what* we say, but *how*, and *where* we say it. These different characterizations of communication, as messages and performances respectively, help us understand why research agendas tend to focus on communication with absent others, and now provoke us to think about how to design for more meaningful co-located interactions.

In fact, an intrinsic assumption around most research in mobile design is that the devices we use most commonly – mobile phones – are *tele*phones, implying interaction at a distance. But research also shows that people wish to interact with their handsets when co-located [14]. This is especially true in regions with high network costs, such as Africa, where many users own a handset, but cannot afford to send a text, let alone share multimedia data [24]. To address these issues, some of our earlier work looked to create an image sharing system, where a user may broadcast an image from their handset over WiFi or Bluetooth to other users within signal range. Results showed that the system was well received [21], leading us to ask two key questions:

What else would people want to do when they are together? Rather than building endless prototypes, we look to theories of proxemics and social interaction to inform our ideas.

How do we support mobile co-located interactions? Whilst it is possible to imagine many types of digitally supported proxemic interaction, is it possible to build a platform on current hardware that would allow us to explore that space?

It is within the context of those questions that we undertook the work presented here. To advance this goal we first show how we *sensitized* ourselves to theories on proxemics, context, identity, and embodiment to obtain a more nuanced understanding of performative aspects of communication and co-located interactions. We explain how we used these theories as a critical lens to examine and *critique* current systems that support co-located interactions. But far from just critiquing, we also show how we applied and refined this lens to *generate* and map out a new set of co-located interactions on mobile devices. We then explain how we integrated insights, sensitivities, and critical stances, derived from theory, to *explore* these through the process of

CHI 2013, April 27–May 2, 2013, Paris, France.

design, culminating in the design of two probes that express and explore key concepts. We present these aspects of our work in four sections, *sensitizing*, *critiquing*, *integrating & generating*, and *exploring*. In choosing this structure, we depart from conventional practice, where related work, usually discussing theory and related systems, is presented as an upfront baseline that is often bracketed out from 'actual' research. Here we show how theory, related systems and even our own previous work were central *throughout*: a constant, yet productive, site of struggle. Our theoretical understandings, drawn from a variety of different intellectual disciplines, emerged as we engaged with them at the different stages of our research. So beyond generating and exploring co-located interactions, it is this story that we tell.

SENSITIZING

Focusing on performance-based aspects places the importance of communication at the heart of the human condition, and destabilizes the assumption that the message of a communication act is somehow more important than how it is bound to context, time, and identity. This broader view opens up a path for research that looks at how co-located people might want to produce and engage with the media stored on their phones. Here, we turn to theory to sensitize ourselves to profound concepts such as *context, time,* and *identity* to understand how we might design for more meaningful co-located interactions, as well as how we shouldn't.

Understanding co-located practices

Two major conceptual lenses used to describe and analyze co-located situations and interactions are Adam Kendon's *F-formations* and Edward Hall's *proxemics*.

Kendon calls the space used in the course of an activity a transactional segment [19, p.3]. So, the transactional segment for a person watching TV incorporates the cone-like space between the couch and the TV. In co-located situations the transactional segments of the people involved typically overlap. A common space is established in which a common activity, for instance looking at a photo or having a conversation, can take place. In analyzing the spaces formed in the course of co-located interactions, Kendon noticed how people "orient and space themselves in relation to one another directly reflects how they may be involved with one another" [19, p.3] In Kendon's terms, cooperative spatial arrangements sustained over time are called *formations*. And formations with a shared transactional space are called *F-formations* (Figure 1). These consist of three segments: the *o-space*, or inner 'sharing-space', the *p-space*, which provides placement of the participant's bodies, and the *r-space*, which serves as the separator between the activity of the interaction and the outside world.

The power of *F-formations* lies in their simplicity; they are widely applied as a conceptual lens to analyze interactions between people. This can be done quickly and does not require knowledge of the nature of the interaction taking place within the F-formation. Using F-formations we can learn about a space, by analyzing where people are and aren't interacting. For instance, if you overlook a public space from above you can see all kinds of different F-formations being formed and dissolving again and, over time, gain an appreciation its interactional properties.

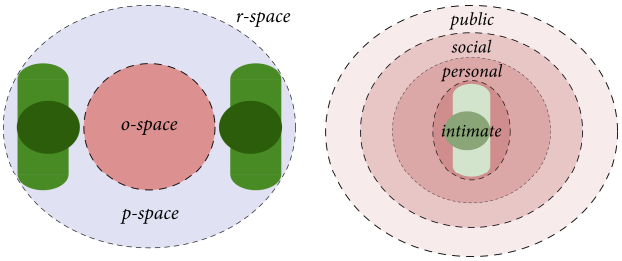

Figure 1. Kendon's *F-formations* & Hall's *proxemic zones*.

Another major work in this area is Hall's theory of *proxemics* [12]. Hall has coined the term proxemics for people's use and perception – through eyes, ears, nose, and skin – of space. Many of Hall's observations focus on how *fixed-feature spaces* (e.g. architecture) interact with *semifixed-feature spaces* (e.g. furniture), and how this affects people, and how they interact with one another. While our experience of space certainly depends on the interplay of fixed- and semifixed-feature spaces, for Hall the most significant category of spatial experience is a mostly unstated, informal space: the distances maintained in encounters with others. He proposes that a person does not end at his or her skin, but is surrounded by a series of expanding and contracting fields: a space, whose size and properties vary on account of culture (more generally) and personal relationships. These can be classified into four, discrete proxemic zones of Figure 1: *intimate*, *personal*, *social*, and *public*. While many of his observations seem commonsense, they relate to a behavior – our use and perception of interpersonal space – that largely lies outside of our awareness; we take it for granted. Hall uncovers these through his descriptions of causes and effects, highlighting how these are interpreted. For instance, people who are angry will move in close to make their point, just as people who are amorous will move in close to express affection. And it's not just touch, but also sensing the heat of another person that combine as we feel intimacy. But the very same sensations (touch & heat) can make you feel claustrophobic in a packed train, so we keep our muscles taught to maintain *our* space.

Researchers who have grounded their work in Hall's proxemic theory and terminology [e.g, 11] acknowledge, as we also suspected in our readings, that Hall's theory is at best suggestive to design and that "we just don't understand the HCI of proxemics" [11, p.50]. But we can find inspiration in Hall's approach that *reflexively* integrates observations. Harper advocates for a similarly reflexive approach to understand everyday actions, such as a son brewing his mother a cup of tea to be ready for her when she gets home. We don't need theories to explain the significance of such acts: "we [need] to use the expertise about the world that we [gain] by living *in* that world" [13, p. 194].

Understanding boundaries of people & time

If we only focus on certain aspects of communication, such significances can easily be lost. Much research to date has focused on mediating the 'bodily mechanics' of communication through computers, such as capturing and conveying gestures or glances [13]. While, bodily mechanics are important, focusing solely on them creates boundary around individuals. Hutchins argues to soften such boundaries and advocates that "the proper unit of analysis is […] not bounded by the skin or the skull. It includes the socio-material environment of the person" [16, p. 292]. Similar arguments apply to time. Experience, as Hutchins notes in a later article, "is not only multi-modal, but also multitemporal or temporally extended in the sense that it is shaped both by memories of the past (on a variety of time scales ranging from milliseconds to years) and by anticipation of the future (over a similar set of time scales) [17, p.432]. We integrate our memories of the past, experience of the present, and anticipation of the future when we engage with others [34] and with artifacts of our world [17]. Rigid boundaries of time or bodies cut through lines of interaction and obscure relevant phenomena, such as the significance of a cup of tea brewed in anticipation of mother's arrival.

Understanding identity

If we soften boundaries of bodies and time, then important implications follow for identity that contrast with most, if not all, ontologies in computer science. These start with an instance or individual – an anatomical, self-referential unit – and associate a unique identity to it, to discriminate one instance from another. They further group instances together to form sets, classes, collections, etc. based on common attributes. To be sure, this is a powerful model, but it is also one that views identity rigidly. While identity is something unique to us, it also "implies a relationship with a broader collective or social group of some kind" [3, p.1]. In softening borders, the question then becomes how does identity interact with the broader collective over time? Buckingham argues against rigid views because: "who I am (or who I think I am) varies according to who I am with, the social situations in which I find myself, and the motivations I may have at the time" [3, p. 1]. In discussing different disciplinary orientations towards identity, he alerts us that identity is not something we *posses*, or something we *are,* but is something we *do*. It is something that comes into being in dialogue between self and other.

One of the more prominent theorists on identity is Erwin Goffman, who sees identity as something that is performed [10]. Goffman approached identity through the metaphor of theater to describe how individuals use their physical and social surroundings to present themselves to the world and, in turn, how the 'world' interprets their 'performance'. He noted that the relationship between the performer and the audience is one of *impression management*; through their (inter)actions performers project themselves, whether intentionally or unintentionally, to an audience that interprets their (inter)actions. So both parties are involved in negotiating a 'performance'. Goffman also classified two types of performance '*regions*' that a performer has access to and is used to maintain their impressions, namely the *front-stage* and the *back-stage*. The front stage is an attempt by the performer to give the appearance that their performance is their de facto standard. Clothing, posture, speech, gesture, and expressions, can also affect performances. In turn, the backstage is the region where the performer can openly contradict front-stage performances, drop their front and "step out of character" [10. p. 70]. While Goffman is not without critics [3], he, like Hall, was a perceptive observer of social interactions. He elevated the world of social interactions from the obscurity of plain sight, by giving us a vocabulary to talk about and observe it.

Understanding context

The discussions so far show that co-located interactions, like communication acts in general, are highly contextual. But as Dourish alerts us, it is a context of a particular nature [7]. Especially face-to-face, context is an interactional, dynamic, occasional property that arises from activity. This characterization departs from how we usually delineate context, as something that is stable and can be measured and encoded without reference to the activity at hand. Within any dialogue what is and what isn't contextually relevant cannot be established a-priori; a comment can trigger a memory and lead the dialogue down a different path. So something that wasn't contextually relevant before, now is. In fact, negotiating context is a very *ordinary* achievement [7]; we do it almost automatically often without noticing. This interactional view of context suggests inquiries and designs that focus not on how to re-present context, but rather on flexibly accommodating changing contexts.

Harper argues that to sensitize ourselves to the performative nature of communication acts, then we must see these along three dimensions: "the how of the act itself (the bodily skills involved), the who of the act (where one needs to be alert to the intentions of the actor themselves and how the undertaking of some act conveys a sense of identity or self for that person and to the audience or recipient of the act), and third, the where of the act (the location of its performance)" [13, p.245]. Or, in short, context, communication, and identity are enmeshed in *performance*.

Understanding photographic (co-located) practices

To show how this enmeshing works in practice, we identified (mobile) photography as a salient practice in which the theories we have discussed take hold in a way that also relates them to (digital) media. Photos usually depict unique, memorable, happy, events and rarely the routine, sad, or ordinary [28]. In a sense they re-present an idealized self [10]. Photos with family or friends contribute to a sense of belonging and of self. Related practices of photo viewing and sharing re-produce and mediate deeply rooted social practices such as gift giving [25] and storytelling [36].

Kindberg et al examined how 34 people in the US and UK use their camera phone, their intentions at the time of capture, and subsequent usage patterns [20]. The authors distinguish these intentions along two dimensions: *affective-functional;* and *social-individual*, where social use is further broken down into *co-located viewing* and *absent sending/viewing*. Their study shows affective photos outnumbering functional, and social outnumbering individual photos. Most image sharing happened in the moment and on the phone's screen, and rarely via Bluetooth or MMS. Results also show that post hoc sharing didn't happen nearly as often, because "the time and effort one must put into sending these 'gifts' [is] difficult to achieve in the moment" [20]. Some participants simply 'hadn't gotten around to it yet,' forgot or lost the impulse to share. When post hoc sharing did happen, it often involved storytelling.

Chalfen further analyzes how people communicate and tell stories with photos [4]. He contends that photographers are reluctant to create self-containing visual narratives. The narrative remains in the head of the photographer: "a picture may be 'worth a 1000 words' … [but] pictures don't literally 'say' anything—people do the talking" [4, p.70]. This does not mean that photos always need to be accompanied by a narrative especially if shared with close-knit, yet absent friends [20]. But it does illustrate why post hoc, co-located sharing is a practice of enduring importance [36]. Photos, especially those we cherish, re-present memories, usually of "moments of positively valued change, marked by parties, official recognition, or public celebration" [4, p. 96]. In this sense, photos link to identity, group belonging, and presentation of self. The stories we tell around photos do more than contextualize a photo but also contribute to our biographical 'narratives' – the stories that explain ourselves to ourselves [3]. Face-to-face sharing lets people adapt their presentation of photos, and with it their presentation of self, to their social surroundings. Such face-to-face sharing is a "dynamic, improvisational construction of a contingent situated interaction between story-teller and audience" [36, p.1073]. This explains why applications, explicitly designed for *tele*communication of photos [e.g. 27] still find their use in co-located situations.

CRITIQUING

In the last section we looked at co-located practices and developed a performative lens of co-located interactions by sensitizing ourselves to concepts of identity, context, and time. This lens destabilizes simpler, yet prevalent, views of interacting humans. Here, we show how we applied and refined our sensitivities by critiquing related work through this lens. We identify two examples of related work that are representative of how co-located or proxemic interactions are commonly conceptualized. We intend for discussions to be productive, a chance for us to contextualize and interrelate what we learned and to identify opportunities for design. But we also recognize that criticism, an embraced part of more mature design disciplines, can be perceived as negative and unhelpful, so to set the tone for a constructive discourse, we start by critiquing our own previous work.

Mobile Digital Stories

In our previous work we designed a mobile digital storytelling system to suit the needs and functions of rural African communities. Informed by ethnography and technology experiments involving storytelling, we implemented a design workshop to involve users in a rural community in South Africa's Eastern Cape in the design of a mobile digital storytelling system. Using this method, we created a mobile digital storytelling prototype to suit the needs of rural users. Details of our design process and the resulting prototype, that can flexibly accommodate different digital storytelling techniques, have been published elsewhere [30]. For our critique we focus on how we field-tested the prototype in a rural community in Kenya. During this formative evaluation, we didn't consider *performative* aspects of story*telling* and focused instead on how people *created* stories – how they took photos, recorded audio, and stitched them together into digital stories – on our prototype. While we also gained more diverse experiences and recorded these in our field notes and pictures, our unit of analysis during storytelling activities was centered largely on the prototype and the person interacting with it. It was only after sensitizing ourselves to theories of distributed cognition [16] and situated action [34] that we understood the lines of interaction that were being cut. People did not merely create, nor tell, digital stories, they *performed* them. One storyteller paid little attention to the prototype when recording her story and instead looked deep into our eyes. Users tailored performances to specific (co-located) audiences; they engaged within their physical and with their social surroundings, often in co-located and collaborative creativity. While we were able to reflect on our activities with more appropriate analytical lenses, we only begun to understand the meanings users created well after the fact [29]. This illustrates how narrow disciplinary orientations obscure important aspects of co-located interactions on mobile devices/applications, especially how these are constituted in and inseparable from physical and social contexts of use.

Mobiphos

Mobiphos is a novel interface that supports photo capture and automatic co-located, synchronous sharing within a predefined group [5]. Reading that paper, we suspected that privacy would be a major issue; but to our and the researchers surprise it was not. Because photo-capture is automatically linked to co-located sharing, the consequences of people's actions are apparent – all members of the group will be able to see all the photos you take. So, people adapted their photo taking behavior to take this into account. Users created meanings by matching the possibilities of the technology to their ongoing goals, on the fly; instead of worrying about privacy they adapted their photo taking behaviors and at different times collaborated or took mischievous and funny photos for other to see.

While Mobiphos can provide an enchanting experience, it works better for certain genres of photography, such as tourist photography [4]. This is also the scenario in which Mobiphos was deployed. Likewise, the tourist is an established social role [10], that carries with it conventions and norms that also relate to photo taking, for instance that photos tend to be of places and objects. The design and deployment of Mobiphos also sets an explicit timescale that is limited to the duration of the co-located 'camera recreation.' What happens before? What happens after? How do groups form? How do people leave or join in later? What happens to the photos and how is privacy negotiated then?

Proxemic Interactions

Greenberg et. al recognize the importance of spatial relationship in proposing and formulating an admittedly speculative vision of ubicomp, namely *proxemic interactions*: "just as people expect increasing engagement and intimacy as they approach others, so should they naturally expect increasing connectivity and interaction possibilities as they bring their devices in close proximity to one another and to other things in the ecology" [11, p.44]. Proxemic interactions are triggered by sensing relationships between people and digital as well as non-digital objects. These relationships are characterized and measured along five dimensions: *distance*, *orientation*, *movement*, *identity*, and *location*. In developing different scenarios and applications, they show how Hall's proxemic zones (see Figure 1) [12] can: "regulate implicit and explicit interaction; trigger such interactions by continuous movement or by movement of people and devices in and out of discrete proxemic regions [Hall's zones]; mediate simultaneous interaction of multiple people; and interpret and exploit people's directed attention to other people and objects" [1, p.121]. The researchers have implemented a number of prototypes, making use of 'fine-grained knowledge' of these dimensions by tracking them through a motion capture system. Most prototypes are rule-based. The proxemic presenter augments traditional presentation tools, and shows presenter notes when the person presenting turns towards the display. The proxemic media player pauses when a person starts to read a book. It can interpret pointing gestures to allow users to select different media items, and displays the movie title when another person enters the room. These applications are intuitively appealing, and show that Greenberg et al. are correct that devices should react to proxemics. But we are concerned that such information is encoded in a representational view of context [7]. Likewise, Goffman teaches us that people behave differently in relation to different contexts. We are concerned that such proxemic interactions are operationalized along too ridged and narrow a definition of identity that "uniquely describes the entity" [11, p.44]. What might proxemic or co-located interactions look like, that are built upon a performative account of identity and an interactional understanding of context? And what would they look like, not in the conference or living room of the future, but on the presently ubiquitous mobile phone.

INTEGRATING & GENERATING

The theories we considered so far have largely been formulated prior to or independent of a device that is shaping the communication landscape around us: the mobile phone. In this section we revisit the theory we outlined earlier to integrate and relate these to how people (might want to) engage with each other and their mobile devices when co-located. In the process, we generate new design spaces and suggest underexplored phenomena and avenues for research.

A key aspect of our performative lens is Goffman's theory on the presentation of self. It has been widely applied to analyzing how we present ourselves on social media outlets, such as Facebook and MySpace, where we write ourselves into being [2], using text, but also media. In everyday co-located interactions, however, our bodies, not our profiles, are the focal point of that performance. We use gesture, speech, and facial expressions; augment them with clothing styles, in order to project who we are [10]. Combining these perspectives, leads us to question what role media plays in our co-located 'performances'? To do so, we first need to understand how media links to performative aspects of identity? Perhaps a better way of understanding identity is when we are stripped of it. In *Asylums,* Goffman describes how this can happen when mental patients are institutionalized [9]. Names are replaced by numbers; our clothing with a uniform; our hair gets shaved off; our pockets are emptied; and the small paraphernalia that we carry on our person – wallet, often containing pictures, briefcase, purse, handbag, book, etc. – are also removed. These items form a person's *'identity kit'*, crucial items for the management of a personal front [9]. For some, the mobile phone, as a physical object, is an intimate part of their identity kit: the brand, the color, and how it is accessorized. But the mobile is more than a physical device; it is a repository of information and histories [22]. So beyond physical devices, how do mobiles figure into our identity kits? What does the stuff on the device – the playlists, gallery, call and SMS log, etc – say about us? What role do these personal repositories play in how we present ourselves face-to-face?

To answers these questions we revisit Goffman's concepts of impression management through front and backstage regions. Looking at mobile phones and the applications we use on them, we might say that they are *front stage* devices, as they provide "insights into our tastes, our style of consumption, and perhaps our allegiance to certain groups" [22,p. 96]. But we might also say the mobile is a *back stage* device, as it has "evolved into a significant repository of personal information" [22, p.97], often containing sensitive information which might, to use Goffman's terminology, contradict the performer's front. But it isn't a purely back stage device either, as we draw upon this repository on applications, like Facebook, Flickr, messaging, and email, to present aspects of ourselves to the people we communicate with, enabling and sustaining a *front stage* performance. This dual characterization of the device also explains, why when co-located, people are reluctant to let go of their mo-

biles, opting instead to show a photo to their friends while holding onto the device. We argue that a better way of figuring the mobile is as a *back stage* device that interfaces with the *front stage*, both asynchronously using Facebook and Email or synchronously when co-located.

One of the main implications of a device that functions in both *front* and *back stage* regions, like the mobile, is that maintaining the device in co-located situations can be problematic. Consider the example of displaying photos from the phone's gallery to a co-located friend. Since photos are usually displayed chronologically, the act of going through one's photos together could easily become uncomfortable, if the presenter accidentally shows a photo that, although chronologically next, shows the presenter in an entirely different context. Or it can lead to overload, with the presenter having to manage the presentation, as well as the performance, resulting in the presenter 'breaking their performance' to ensure that only correct images are shown. We identify a need for mobile devices and applications to accommodate both *front* and *back stage* management, especially when used in relation to co-located contexts.

We must also consider what happens in the situation itself. The act of coming together, the essence of co-located interactions, is not a virtual activity but a physical one; we do it in relation to our bodies. As such, we revisit Kendon's *F-formations,* as they explain how groups form and are maintained during interactions. Our aim is not to replicate physical formations, but to find *inspiration* in them. Beyond physical characteristics of group formations, we are inspired by Kendon's descriptions of how groups form and change over time. Consider if you want to join a group of chatting people. By moving towards the group you make a bid for entry. But the often circular arrangement of groups creates a boundary, forcing you to wait along its periphery. The group must accept your bid and does so by *making space* for you. Groups do this almost automatically and fluidly by changing their placing, posture, formation, and perhaps even their conversation topic. We find inspiration in the fluid and visible dynamic of this practice. If our mobile devices could reflect this 'coming together' of people and devices, what are suitable gestures for forming or bidding entry into such a group? Once established, how would mobiles adapt and reconfigure themselves to reflect such a coming together? What are suitable interface metaphors and interactions? Without disrupting the fluid dynamic of co-located interactions, how might mobile applications augment such encounters? How are such encounters sustained? How do they change? Do they live on after dissolving?

When thinking about co-located interactions, it is easy to see these as isolated events. But the bonds between people remain intact, as they move between periods of absence and presence – rhythms that punctuate life. Or as Suchman reminds us, interactivity, or engaged participation with others, does not just require a *presence*, but also an *autobiography*, and a *projected future* [34]. Far from isolated events, we look forward to our get-togethers. To be sure, some get-togethers are spontaneous, but rarely completely unexpectedly as a substantial degree of ordinariness characterizes our lives [7]. During get-togethers, we use our mobiles to share a past experience, a piece of our autobiography so to speak [36]. So why can't we use our mobiles to project into the future and support such practices? Can a system be designed for people to draw upon the media they produce and consume while being mindful of absent others, but rather than sharing in the moment or forgetting altogether [20], project into the future? That is, to anticipate future presence and the joy of sharing photos face-to-face might bring, where intersubjectivity is richer [15], and where we can co-orient [26] towards and make sense of the media together?

EXPLORING

All of these questions provoke us to think about what meaningful co-located interactions on mobile devices might look like. We identify two, to a degree interrelated, design spaces that highlight different aspects of our discussions so far. As they are shaped by our performative lens, they share many characteristics – how they approach issues of identity and context; and how they interweave with existing practices – but differ along a few dimensions that we outline next.

The first, which we call *Share Face2Face,* assumes that the natural co-located sharing pattern is not necessarily a pattern of sending files, but a pattern of co-consumption and co-orientation. Within this class the natural sharing gesture is *show-and-tell*, in which a small group of people co-orient themselves towards the mobile's screen to look at a photo or listen to a song, together. *Share Face2Face* does not see these events as isolated, but as something that people might look forward to. As such it has a softened boundary of time, and will explore how to enhance these encounters, by allowing people to draw upon the media that they produce and consume while apart and being mindful of others at a different time and place; and then to bring this act of mindfulness into a co-located situation, at a later time or in a different place. This extended timeframe renders *Share Face2Face* into a design space that is more deliberate, slow, and curatorial; that leaves room for pausing and pondering; and allows people to anticipate and look forward to future presence. It is not about sharing across distance in the right-here-and-now, but sometime in the future, when the time is right, when we re-connect face-to-face.

Our second interpretation challenges the pervasive cloud-computing paradigm. People take to the cloud because services, such as Facebook, Flickr, Soundcloud, allow them to discover and share media and engage with it socially. Much ink has been spent on performative aspects of social media use and the difficulties of portraying oneself in front of a largely imagined audience while retaining privacy. But in co-located situations the audience is very real and it does not require us to imagine a context in which someone might view media we make available. As a consequence, some of the issues surrounding privacy could be offloaded into the

real world, or negotiated on-the-fly by more closely matching media we might want to, or want others to, engage with to the situation at hand. We envision *Cloudlets* to be hyper-localized, ad-hoc instantiations of the cloud that are more or less independent of cloud infrastructure, but nevertheless provide similar services and opportunities for both virtual and physical engagements surrounding media. As an underlying infrastructure *Cloudlets* can make applications aware of, and provide opportunities for engaging with, nearby people and devices. Such co-located engagements surrounding media happen in real-time, affording rich ways to interact with our data when co-located. For instance, music-playing applications can allow friends to broadcast selected songs and albums to one another, virtually merging collections, highlighting similar tastes, or providing ways to discover new music. Calendars applications could show past or future common events and match free time for co-located people to schedule future get-togethers. In short, just like Cloud services annotate our data with semantics, *Cloudlets* aims to make our mobiles more aware of, and provide opportunities for engaging with, co-located people.

If we look at these design spaces closely, how they interweave with social practices, how they are informed by a theoretical scaffolding of a variety of intellectual disciplines, we see a wicked problem; our formulation of the situation is an integral part to addressing it [8]. As such, it warrants a research through design (RtD) approach [38]. While crossing disciplinary boundaries during our ongoing theoretical ponderings are an invaluable resources that help us to observe, talk, and think about co-located interactions, in choosing a RtD approach we should also remain true to *our* disciplinary orientations, namely interaction design and computer science. In crossing boundaries, it is easy to become dismissive of our own skill set, which in comparison to the perspectives and insights that *others* bring to the situation appear unremarkable and ordinary. As reflective and reflexive design researchers, we embrace the unique perspectives and skill sets that we bring to the situation and recognize how we generate knowledge. As interaction designers, we have developed a deep understanding of what software, interfaces, and sensors can and cannot do. These understandings and skill-sets color our interpretations of theories. We agree with Gaver when he says, "the practice of making is a route to discovery" [8, p.942]. Thus, the prototype [35] presents a chance to express, and through the process of prototyping, further our understandings of these co-located interaction spaces. We experiment with technological *possibilities* – new arrangements of people, contexts, and technologies – to build the *right* thing [37], to change and possibly disrupt behavior [31] and see what reality could become [23]. We design and implement prototypes to understand and explore these spaces in the lab and through unfolding activities of co-operative design-in-use [35]. In the following discussion, we call prototypes that have been designed and deployed in this fashion *probes*. We chose a probing approach because the possibilities afforded by current technologies don't support co-located interaction in a way that resonates with our earlier discussions. By building probes, left strategically incomplete and flexible, we engage with the more synthetic aspects of design, that "allows for a richer more situated understanding" [8, p.942] – not just in theory, but in practice by presenting users with real usage contexts [18]. In the following two sections we introduce these probes and how they express our understandings of design spaces we identified above.

Figure 2. The *Share Face2Face* probe.

Share Face2Face

In looking for pragmatic design solutions [24] – ones that don't require adding infrastructure – we explored ways to convey our operative image of *Share Face2Face* on existing camera phones. Researchers studying camera phones have identified that the impulse to share is strongest in the moment [20]. Looking at how media is shared, posted, or 'Bluetoothed' on current devices we identified that all camera phones have built in sharing mechanisms. In most cases, these can be accessed from the contextual menu of a particular media item, revealing a list of sharing options: Bluetooth, Email, Facebook, or Messaging. To encapsulate our operative image, we wanted to present a similar option to users called *Face2Face* and created a suitable icon (see Figure 2). To keep our probe simple and interpretatively flexible [33], we created a dummy application on an Android phone that hooks into the Android's built in sharing mechanism. As a consequence, the Share Face2Face probe was displayed next to media sharing technologies such as Email, Bluetooth, and Facebook, which we hoped would make users reflect on how they currently share media [32]. If indeed a piece of media needs to be shared in the right-here-and-now, or if sharing it face-to-face later on might bring about a different experience, for instance by seeing the reaction of a friend's face when they look at a photo or listening to a song together. While we intend to use this probe on a larger scale, to first test out its feasibility we first went around the university campus during lunchtime to informally interview about 10 students.

We first showed them the concept, which we accessed through phone's contextual share menu from an image we

took earlier. The *Share Face2Face* concept immediately caught the imagination of the students we interviewed. They talked about how face-to-face sharing of a photo is different experience from Facebook: "*facial expressions say more than likes*." But when queried about how they might like to use such an application, their answers were characterized by *vagueness*. In their view such an application would be useful in general, but we only elicited a few specific scenarios, such as showing a botanist friend a picture of a flower a student saw on a recent hike. Reflecting on these results we realized that we failed to consider how the probe only takes on *specific* meanings in relation to *specific* images, evoking *specific* feelings. Clicking on the Face2Face sharing option can be seen as lodging an intent to share this media item face-to-face. This reminded us of Suchman's theory of *Plans & Situated Action* – commonly misinterpreted as a theory of only situated action, neglecting the important imaginary and discursive practice of planning [34]. It is through planning that we project into the future and imagine how things might be. Such sharing *intents* – or plans – are an important resource within the situated action or practice of face-to-face image sharing. This suggests a discourse exploring the relationship between plans and situated action; how do these sharing intents relate to, and might be useful within, face-to-face encounters; how might they change with time? To explore this relationship within the *Share Face2Face* design space, we subsequently expanded our initial probe with a diary functionality to let users create text and voice notes that link to the specific media items that they intend to share face-to-face. This will allow us to explore more specific scenarios and intents to help us further understand this design space.

Cloudlets

In order to investigate whether *Cloudlets* are a suitable means for engaging with media in co-located situations, we needed to test whether current mobile phones actually support such an infrastructure, before we could tackle some of the more interesting questions that surround *Cloudlets*. So, we prototyped an ad-hoc infrastructure that allows three devices to interconnect via Bluetooth in a seamless manner. To render this infrastructure visible to users, we developed two probes. The first probe was a messaging application that broadcasts a message to each connected device as it is being typed; and the second probe was a photo-displaying application, which displayed selected photos from connected devices in a vertical list of photo-streams, designed to be similar to most mobile photo-sharing applications, such as Flickr. The probes were designed to demonstrate the real-time nature of the content being shared and were not feature-filled applications. This was done to supplement an experimental session, as we were unsure how participants would approach the technology or what features it would require. As such, the aim of the experiment was to stimulate discussion with the users and investigate whether they could understand and envision other applications that make use of this underlying infrastructure.

With the probes in place, we then performed an experiment to test if our more theoretical ponderings did indeed have practical relevance, as well as to identify whether the basic concept behind *Cloudlets* – people and their devices coming together to create a localized cloud – is something that resonates with users. Empathizing with the participants, we did not explicitly engage with the more theoretical concepts surrounding identity and performance. Because, taken out of context, Goffman's characterisation of *front* and *back-stage* can be misunderstood to imply that the *back-stage* is somehow more honest and the *front-stage* somehow less.

The experiment included a total of ten participants, which were split in two groups of five. The participants' ages ranged from 18 – 22 years and contained a total of three groups of friends and four strangers. This format was chosen to target an age group comfortable with mobile sharing and to understand the application features needed due to both friend and stranger relations. To get participants talking, we first discussed their current practices of sharing media in co-located situations. Most of them used cloud-based services, such as BBM, WhatsApp, and Facebook, or displayed the mobile phone screen to others in a show-and-tell gesture. When questioned about Bluetooth, they mentioned that it was too slow and clumsy to use, reverting instead to cloud services or show-and-tell gestures. They did note that they would only hand their mobile device to another if it was a small gathering of people or if they knew the person very well; but even then it was assumed that the 'shared' would only view the content and not do anything else with the device. We followed this with an introduction and demonstration of the probes. We encouraged participants to interact with the applications and asked questions pertaining to the probes. The probes caught the imagination of the participants; they envisioned several co-located scenarios surrounding media, such as photo-sharing, lecture-material broadcasting and mobile music-streaming. They were also surprised to learn that the probes used only Bluetooth; a technology that they never imagined could function in the manner demonstrated. This was encouraging, as it not only demonstrated that participants could envision novel uses for *Cloudlets*, but also exemplified that participants want richer co-located sharing experiences surrounding media than currently possible.

We then questioned participants about applications classes, such as co-located photo-sharing or an overlapping group calendar. In relating their everyday experiences to these classes, participants saw their usefulness and could imagine using them, but quickly raised issues relating to privacy and identity management: who has control over who sees what; how they gain access to it; how they limit or specify what they are sharing and with or to whom; what others do with the content; who owns the content; how do they enable and disable this technology; how do you specify who can connect to your device; etc. These concerns were amplified when the participants were asked to collaboratively sketch out two applications, a calendar and a photo-sharing appli-

cation, that used Cloudlet infrastructure. Overall, participants wanted more control over content and sharing granularity. The experiment indicates that *Cloudlets* resonated with participants; they find the possibilities for richer co-located sharing exciting, and our theoretical ponderings on privacy and identity management are of central importance.

FUTURE WORK

Moving forward, we are interrogating the scenarios proposed by users, and unpacking the subtle distinction between what is given (implying giving ownership) and what can only to be viewed while co-located. Beyond static read/write distinctions within file-management, this suggests a more refined, social conception around permissions and how acts or gestures might operationalize these. After all, co-located interactions surrounding media render these into social objects. Comments on group formation are leading us to consider more dynamic privacy models around groups, as well as more appropriate interface metaphors, allowing users to better negotiate privacy and identity management. We are developing these through further experiments around refined probes, to further situate and conceptualize *Cloudlets* and how they might be utilized in practice.

In fact, the functionality we see emerging in Cloudlet applications could well be based on set theory. In essence people bringing together data in co-located situations want to find the intersection, union and difference of such data: we might want to find songs we have in common (intersection) or dates when we are all free (difference). By adding set theory based operators to our Cloudlet applications, we can explore new forms of co-located interaction. So we postulate that it is not just social theory, but something as abstract as set theory, that can be used as a lens to understand and generate new forms of co-located interaction.

CONCLUSION

Co-located or any type of social interactions are characterized by their ordinariness [7]; we take them for granted. These interactions are enmeshed with unwritten rules, feelings, expectations, anxieties, and experiences [10,12]. So, we drew upon theory to develop analytical lenses and sensitize ourselves to key practices, sometimes rendering these visible to us in the first place. We have characterized this approach as a constant, yet productive site of struggle that depends on one pervasively misunderstood activity: *reading* [6]. When we speak of reading we see it as an act of habitation rather than consumption. Just like a renter furnishing her apartment with objects, acts, and memories, by continually engaging with (or reading) theoretical texts at different stages of our research, we have made these texts our own. We used analytical lenses derived and refined through reading to understand and critique approaches and conceptual foundations of related work including our own. These understandings showed us that we needed a better way to think about co-located interactions on mobile devices. By interrelating theories we uncovered that such interactions are situated primarily in a *social* ecology with devices, not in a device ecology with people. Our dialogic engagement with theory allowed us to develop such grounding principles. We have identify two design spaces *Cloudlets* and *Share Face2Face*, and have begun to explore them through the process of design, which so far has differed from more traditional user-centered design (UCD) processes.

In his widely cited paper on context, Dourish notes that if we are to sensitize ourselves to different disciplinary orientations on profound concepts such as context, identity, and communication, this not only implies "a change to the ways in which we go about designing technologies, but also a change to the technologies that we design" [7]. This is more than a change in process, because the results of a PD or UCD process can still be stable, static, and closed [7]. It was after all "a circular move of writing a cognitivist rationality onto machines and then claiming their status as models for the humans" [34, p.259] – a view that is not only prevalent in academic but also in everyday discourses. While we firmly believe in the philosophies of PD and UCD, in practice we have so far not adhered to them. Although we are currently transitioning into more traditional UCD cycles, the discussions presented thus far characterize a different, first step that doesn't fit into UCD. We call it *trying to understand*. In our account, we have articulated – sometimes explicitly and other times deliberately implicitly – this *theory on design* [38] in the process of uncovering implications for the design of co-located interactions. If we, as the CHI community, take design seriously, this should happen in parallel. We aren't saying that every project should engage with theory in such a manner, but it was the right approach within this project, and is an emerging trend in HCI [31]. We also believe that it is an approach that could benefit other projects that latch onto more tacit aspects of our everyday lives. This is why we report and interrelate them here, because *theory on design* makes more sense in relation to *specific* design problems [8]. To develop research through design as a methodology and to hold ourselves accountable to the decisions we, while struggling with, nevertheless made, we respond to Zimmerman et al.'s call to action by documenting the *whole* process, showing "how theories from other disciplines were integrated" and beginning with the crucial first step: *problem framing* [38, p.316]. But further than theory on design, we hope that our discussions, interpretations of theory in relation to co-located interactions, probes, and technology experiments surrounding the probes, will serve as placeholders that open a new and fruitful design space: the space of co-located interactions that fit primarily into our social ecologies and not just our device ecologies. We hope that others will join us to explore (and theorize) this space.

ACKNOWLEDGEMENTS

We thank participants for their time and efforts, Yvonne Rogers for helping us develop our understandings of theory, and reviewers for their insightful discussions.

REFERENCES

1. Ballendat, T., Marquardt, N., & Greenberg, S. Proxemic Interaction. *ACM Intl. Conf. on Interactive Tabletops and Surfaces – ITS'10*, ACM (2010), 121–130.
2. boyd, d. Why Youth (Heart) Social Networking Sites. In D. Buckingham, ed., *Youth, Identity, and Digital Media*. MIT Press, Cambridge, MA, 2008, 119–142
3. Buckingham, D. Introducing Identity. In D. Buckingham, ed., *Youth, Identity, and Digital Media*. MIT Press, Cambridge, MA, 2008, 1–22.
4. Chalfen, R. *Snapshot: Versions of Life.* University of Wisconsin Press, Madison, WI*,* 1987.
5. Clawson, J., Voida, A., Patel, N., et al. Mobiphos: A Collocated–Synchronous Mobile Photo Sharing Application. *MobileHCI '08*, ACM (2008), 187–195.
6. de Certeau, M. *The Practice of Everyday Life*. University of California Press, Berkeley, CA, 1984.
7. Dourish, P. What we talk about when we talk about context. *Pers. and Ubi. Computing 8*, 1 (2004), 19–30.
8. Gaver, W. What should we expect from research through design? *CHI '12*, ACM (2012), 937–946.
9. Goffman, E. *Asylums*. Anchor Books, New York, 1959.
10. Goffman, E. *The Presentation of Self in Everyday Life*. Anchor Books, New York, 1959.
11. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M. Proxemic Interactions *interactions 18*, 1 (2011), 42–50.
12. Hall, E.T. *The Hidden Dimension*. Anchor Books, New York, 1966.
13. Harper, R. *Texture: Human Expression in the Age of Communications Overload*. MIT Press, Cambridge, MA, 2010.
14. Harper, R., Regan, T., Mosawi, K.A., & Rouncefield, M. Trafficking: Design for the Viral Exchange of TV Content on Mobile Phones. *MobileHCI'07*, ACM (2007), 249–256.
15. Hollan, J. and Stornetta, S. Beyond Being There. *CHI '92*, ACM (1992), 119–125.
16. Hutchins, E. *Cognition in the Wild*. MIT Press, Cambridge, MA, 1995.
17. Hutchins, E. Enaction, Imagination, and Insight. In J. Stewart, O. Gapenne & E.A. Di Paolo, eds., *Enaction: Towards a New Paradigm for Cognitive Science*. MIT Press, Cambridge, MA, 2010, 425–450.
18. Kangas, E. and Kinnunen, T. Applying user-centered design to mobile application development. *Communications of the ACM 48*, 7 (2005), 55–59.
19. Kendon, A. Spacing and Orientation in Co-present Interaction. In A. Esposito, N. et al. eds., *Development of Multimodal Interfaces: Active Listening and Synchrony*. Springer, Berlin, 2010, 1–15.
20. Kindberg, T. Spasojevic, M. Fleck, R. & Sellen, A. The Ubiquitous Camera: An In-Depth Study of Camera Phone Use. *Pervasive Computing 4*, 2 (2005), 42–50.
21. Kun, L. & Marsden, G. Co-present photo sharing on mobile devices. *MobileHCI'07*, ACM (2007), 277–284.
22. Ling, R. *New Tech, New Ties: How Mobile Communication is Reshaping Social Cohesion*. MIT Press, Cambridge, MA, 2008.
23. Löwgren, J. & Stolterman, E. *Thoughtful Interaction Design.* MIT Press, Cambridge, MA, 2004.
24. Marsden, G., Maunder, A., & Parker, M. People are people, but technology is not technology. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 366*, 1881 (2008), 795–804.
25. Mauss, M. The Gift: The Form and Reason for Exchange in Archaic Societies. Routledge, London 2011.
26. McLeod, J., & Chaffee, S. Interpersonal approaches to communication research. *American Behavioral Scientist*, 16, (1973), 469-499
27. Naaman, M., Nair, R. & Kaplun, V. Photos on the Go. C*HI '08*, ACM (2008), 1739–1748.
28. Shove, E., Watson, M., Hand, M., and Ingram, J. *The Design of Everyday Life*. Berg, Oxford, UK, 2007.
29. Reitmaier T. 'She looked deep into our eyes': reflections on cross-cultural practice. *Knowledge Management for Development Journal 7*, 3 (2011), 327–339.
30. Reitmaier, T., Bidwell, N.J., and Marsden, G. Situating digital storytelling within African communities. *Int. J of Human-Computer Studies 69*, 10 (2011), 658–668.
31. Rogers, Y. *HCI Theory: Classical, Modern, and Contemporary*. Morgan Claypool, 2012.
32. Sengers, P., Boehner, K. et al. Reflective design. *Proc. Critical computing – CC'05*, ACM (2005), 49–58.
33. Sengers, P. & Gaver, B. Staying open to interpretation. *DIS '06*, ACM (2006), 99–108.
34. Suchman, L. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge Univ. Pr., NY, 2007.
35. Suchman, L., Trigg, R., and Blomberg, J. Working artefacts: ethnomethods of the prototype. *The British journal of sociology 53*, 2 (2002), 163–179.
36. Van House, N.A. Collocated photo sharing, storytelling, and the performance of self. *Int. Journal of Human-Computer Studies 67*, 12 (2009), 1073–1086.
37. Zimmerman, J., Forlizzi, J., and Evenson, S. Research through design as a method for interaction design research in HCI. *CHI '07*, ACM (2007), 493–502.
38. Zimmerman, J., Stolterman, E., and Forlizzi, J. An Analysis and Critique of Research through Design. *DIS'10*, ACM (2010), 310-319.