

Connecting in the Kitchen: An Empirical Study of Physical Interactions while Cooking Together at Home

Jeni Paay, Jesper Kjeldskov and Mikael B. Skov

Department of Computer Science / Centre for Socio+Interactive Design
Aalborg University, Selma Lagerlöfs Vej 300, DK-9220 Aalborg East, Denmark
{jeni, jesper, dubois}@cs.aau.dk

ABSTRACT

Recent research has explored the role technology might play in future kitchens, including virtually dining together, recipe sharing, augmented kitchen furniture, reactive cooking utensils and gestural interaction. When people come together in a kitchen to cook it is about more than just production of sustenance – it is about being together, helping each other, exchanging stories, and contributing to the gradual emergence of a shared meal. In this paper we present a digital ethnography of how people coordinate and cooperate in their kitchens when cooking together for the purpose of inspiring the design of social natural user interactions for technologies in the kitchen. The study is based on 61 YouTube videos of people cooking together analyzed using the frameworks of proxemics and F-formations. Our findings unfold and illustrate relationships between people’s spatial organization, their cooking activities and physical kitchen layouts. Based on these we discuss the kitchen as a design space and particularly the opportunities for social natural user interaction design.

Author Keywords

F-formations; proxemics; natural user interaction; cooking together; digital ethnography; digital kitchens; the home

ACM Classification Keywords

H5.3 Computer-supported cooperative work

INTRODUCTION

In the past decade there has been a marked increase in CSCW research centered on the home, rather than the office. This indicates an enormous potential for innovation in domestic information technology [51]. At the same time with the advent of off-the-shelf availability and affordability of motion sensing natural interaction technologies, such as Microsoft’s Kinect, perhaps we can look beyond the realms of family participation in gaming

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CSCW ’15, March 14 - 18 2015, Vancouver, BC, Canada
Copyright 2015 ACM 978-1-4503-2922-4/15/03...\$15.00
<http://dx.doi.org/10.1145/2675133.2675194>

and envision new approaches to technologies for the home, exploring new opportunities and applications. In thinking about natural user interactions as inputs to technologies in the home, it follows that we need to understand of the kinds of physical and gestural interactions that happen in co-located domestic shared activities.

Studying and analyzing human behavior in physical space, and the way we use technology and other artifacts around us, are not new ideas. As discussed by Rogers [49], a number of theories from other disciplines, new or relatively unknown to the fields of CSCW and HCI, are currently making their way into our research. Some of these, she argues, represent “the beginnings of a ‘turn to space’”, including schematic frameworks for modeling people’s interactions with each other in physical space, such as *proxemics* and *F-formations*. Developed for analyzing social interactions outside the technology domain, these theories and frameworks have recently proved to be highly valuable in relation to interaction design as this domain broadens itself, embracing the challenges of digital ecologies and pervasive and ubiquitous computing. Our approach is supported by Reitmeier et al. [48] who state: “Two major conceptual lenses used to describe and analyse co-located situations and interactions are Adam Kendon’s *F-formations* and Edward Hall’s *proxemics*”.

In past work, using proxemics and F-formations, we studied how people physically orient themselves when they are cooking together [44]. We found that the spatial organization of behavior in the context of the shared cooking activity goes beyond the kinds of formations found in purely conversational contexts, and is directly influenced by the architectural design of kitchen spaces, which structures the kinds of interactions and social activities that take place there. Cooking activities and artifacts used in cooking also contribute to the kinds of patterns that occur.

In this paper, we analyze the juxtaposition of the physical formations reported in our earlier papers [43, 44] against the different kinds of cooking activities that people are involved in to better understand the ways multiple people coordinate and cooperate when cooking, including both physical and conversational interactions. We then extend our thinking to general ideas on how our findings might be applied to the design of cooperative natural user interaction technologies to enhance the cooking experience - being assistive, social, and celebratory [13] in approach.

RELATED WORK

There is a small but extant literature in HCI and CSCW regarding the actual practices of cooking, covering a wide range of different perspectives and interests. Cooking and eating are central activities of human experience and as such are relevant to these disciplines [23]. To support the design of technology solutions for kitchens that are usable, useful and facilitate enjoyable experiences, researchers are working to provide grounded understandings of the needs and activities of cooks in different contexts [27].

Ethnographically inspired approaches such as direct observations combined with interviews are the most common techniques to understanding the current practices of people cooking and eating, such as in the work of Bell et al. [2], Chamberlain and Griffiths [7], Crabtree et al. [10] and Kerr et al. [27]. This has been done across different cultures to achieve a new perspective on the activity of cooking [2], and to understand user goals and practices in meal preparation [7, 27]. Sometimes the focus is on how people appropriate new and old kitchen technologies in their cooking [2, 37]. Other studies use indirect methods to gather data, such as the work by Hupfeld and Rodden [23] who use self reporting in the form of photographs of family meals with follow up interviews, or Clear et al. [8] who use fixed cameras mounted in kitchens for automatically recording cooking activities.

The outcomes of these “understanding” studies take a variety of forms, e.g. design strategies [2], cooking personas [27], implications for design [8, 23], all aimed at informing the design of kitchen technologies. Extending this body of work, our own research focuses explicitly on understanding cooking practices as a social activity embedded in a particular physical space. This has encouraged us to explore the ‘turn to space’ in CSCW and HCI research encouraged by, amongst others, Rogers [49] and Greenberg et al. [12], through the use of frameworks for modeling people’s interactions with each other in physical space, such as proxemics and F-formations.

Proxemics covers a topic concerned with peoples’ use of their immediate physical space in their interaction and communication with each other. It was originally developed by Edward Hall, a cultural anthropologist, in the early 1960s [16], but has recently gained renewed momentum as a framework for describing and understanding people’s interaction with devices and displays around them. In brief, Hall proposed that (in the USA) four interpersonal distances govern social interaction: *intimate* (0-0.46m), *personal* (0.46-1.22m), *social* (1.22-3.66m) and *public* (>3.66m). In the context of CSCW this provides a useful framework for analyzing and designing interactive technology for people working together. Examples of this include Obata and Sasaki’s virtual visiting system [38], Heaton’s case studies of social interaction in virtual workspaces [20], and more recently in a general interaction design context, the works of Greenberg et al. on Proxemic Interactions [12].

F-formations is part of a conceptual framework about people’s interpersonal spacing and orientation developed by Adam Kendon, a leading researcher in the study of gesture, in the late 1970s [25, 26]. The notion of F-formations helps explain how people arrange themselves spatially with respect to others, and to artifacts, in different kinds of social interactions and collaborations. It can also be used to explore the influence of physical environment on the interpersonal interactions taking place there. It was developed by Kendon in response to previous studies in proxemics in order to draw attention to the fact that “participants in occasions of interactions of all sorts may enter together into a cooperative relationship to sustain, through time, a more or less constant pattern of spatial orientational arrangement” [Kendon 2012, pers. comm.].

Briefly described, in an F-formation system, individuals have a space called a *transactional segment*, which is the space where they focus attention and manipulate artifacts. This space is defined in relationship to their lower body, and changes in size depending on the kind of activity people are doing. An F-formation is formed when the transactional segments of two or more people overlap and create a shared inner space, where the shared activity of those people occurs. This is called the *o-space* [25] or the *use-space* [26].

Like proxemics, the work on F-formations has recently attracted interest among CSCW and HCI researchers concerned with design and use of technology embedded in the physical environment, for example in the works of [21, 21, 30, 31, 34, 35, 41]. In the study by Marshall et al. [31] the F-formation system is used for analyzing social interactions between visitors and staff in an information center for tourists. They show how the presence or absence of F-formations can explain the relationship between co-located interactions and the physical environment, and how this can inform technology design. Mentis et al. [34] conduct a similar analysis but in the context of work practices in neurosurgery theatres. This leads to insight about the spatial organization of shared work practices in this specific setting, and to understanding design implications for touch-less interaction with medical technology. Other uses of F-formations in HCI include investigating the spatial relationships between humans and service robots during cooperative tasks [21], detecting F-formations from video footage for analyzing social behavior in crowded environments [21], and for supporting cross-device interaction [30].

F-FORMATIONS IN COOKING TOGETHER

Through a study of user-created videos on YouTube we analyzed how people arrange themselves spatially in their kitchens while cooking together. In previous papers [43, 44] we observed four of Kendon’s known F-formations, *Side-by-side*, *L-shaped*, *Vis-à-vis*, and *Semi-circular*, but also identified four *new formations* not described previously. We named these new formations *wide V-shaped*, *Spooning*,



Figure 1. Known F-formations observed while cooking together

Z-shaped, and *reverse L-shaped*. As these eight formations all play an important part in our further analysis of interactions while cooking together presented in this paper, we will briefly describe them.

Figure 1 illustrates the four known F-formations observed while cooking together: *Side-by-side*, *L-shaped*, *Vis-à-vis*, and *Semi-circular*. The *Side-by-side* formation occurs when two people are standing next to each other abreast and facing in the same direction. As described by Kendon [26] this is a common formation when people are jointly concerned about something in their immediate environment. While cooking together, this formation was typically observed when people were working at the same kitchen bench. The *L-shaped* formation occurs when people are standing or sitting with their torsos orthogonal to each other and facing towards a shared transactional space in front of them. Kendon relates this formation to conversations between two people where the topic is disembodied. This formation was often observed when working on a shared task at the corner of a bench while cooking together. The *Vis-à-vis* formation occurs when people are standing opposite and facing each other, creating a shared transactional space between them. Unlike the L-shaped formation, Kendon relates this formation to conversations between two people where the topic is related to their relationship. While cooking together this formation was often observed when working on two sides of a freestanding kitchen bench. The *Semi-circular* formation

occurs when more than two people are standing in a circular shape with one side of the circle open. This formation is common when more than two people are engaged in a shared conversation while either watching the same thing or being open to other people joining in. This formation was often observed when gathered around a stove or similar shared point of focus.

While it is in itself interesting to observe how known F-formations appear slightly differently from purely conversational situations in the context of human-food interaction, the more remarkable finding was that the activity of shared cooking clearly made way for new formations notably different from those normally found in interpersonal communication, illustrated in Figure 2: *wide V-shaped*, *Spooning*, *Z-shaped*, and *reverse L-shaped*. The *wide V-shaped* formation occurs when two people are standing next to each other, but rather than facing in the same direction are turning their bodies slightly towards each other so that they are facing the same shared space in front of them. While cooking together, this formation occurs frequently when people are working along the same kitchen bench but also engaged in conversation or collaboration. The *Spooning* formation describes a rather intimate interaction occurring when one person approaches another from behind and enters closely into facing the same direction and sharing the same immediate space in front of them both. While cooking together this formation was often observed when the physical layout prevented other ways of



Figure 2. New formations observed while cooking together

sharing the same transactional space, and where the people cooking together had a close personal relationship (e.g. partners or parents with their children). The *Z-shaped* formation occurs when two interacting people are standing side-by-side but facing in opposite directions. This would seldom happen in a purely conversational context, but appears to be quite a comfortable way of interacting while cooking together. Here it is often observed when people are working at different kitchen benches, or when passing each other, while also engaged in conversation or collaboration. Finally, the *reverse L-shaped* formation occurs when two people are placed orthogonal to each other but with their torsos facing *away* from each other. This would seldom happen during normal conversation, but was observed often when people are working on orthogonal kitchen benches.

COOPERATION IN COOKING TOGETHER

The findings presented above and further below are based on an observational study of people cooking together in their own kitchens. However, in order to gain access to such situations we have had to deal with some methodological challenges. As discussed several times in CSCW and HCI research (e.g. [9, 36]), and charmingly illustrated in Hamer's movie "Kitchen Stories" [18], although the home is relatively easy to access, it is not always possible or desirable to meaningfully do traditional direct observation in people's homes. Not only is it a well-known social phenomenon that people alter their behavior when being watched (The Hawthorne Effect), but researcher presence in people's *private homes* is also likely to disrupt the flow of household activities as they would normally unfold, and raises new challenges of ethics and preservation of privacy. Consequently, conducting research in a home presents its own set of particular challenges, and although anthropologists have been faced with these for decades, and valuable insights on the practical aspects of such research exists (e.g. [1, 28, 39]), these challenges are still considered difficult to overcome with traditional methodological tools.

In response to some of these challenges, research into the domestic domain has explored a range of data collection approaches, including installing video cameras in people's homes [9], and has to a large extent turned to methods of self-reporting for obtaining insight into people's daily lives with minimal interference. This has been done through various techniques and data collection media, such as cultural probes [15], technology probes [24], experience sampling [36], and diaries [45]. Taking a step further, Masten and Plowman [32] suggest that if we really want to understand people and their everyday lives then we need to rethink and reconfigure the principles and techniques of traditional ethnography to include *digital* realms to a much greater extent. As many aspects of our everyday lives today leave an increasing digital footprint, we should use this as a key source of data. Masten and Plowman propose the notion of "digital ethnography" as a form of ethnography where the source of information includes all sorts of digital

technology capable of sensing aspects of their surroundings. In using these data sources for digital ethnography, researchers make use of digital data, such as written words, images, audio files, video and online communities, about the phenomenon of interest, which is often publically available in vast amounts.

Using YouTube as a Data Source

Inspired by the notion of digital ethnography, and other studies making use of YouTube as a data source (e.g. [3, 4, 19, 50]), we have explored an alternative approach to obtaining insights into what happens in people's private homes while cooking together by drawing on user-generated video recordings posted on YouTube. Using YouTube videos as a data source in an ethnographic-type study is not a new idea and has previously been used to good effect in other studies. For example, YouTube video postings were used by Blythe and Cairns [4] to study user engagement with a specific digital musical instrument, by Harley and Fitzpatrick [19] to study social contact between older and younger people, and by Rotman and Preece [50] to study the creation of online communities through video sharing. When chosen carefully, Blythe and Cairns [4] argue that YouTube potentially provides several interesting avenues of inquiry as a digital on-line data source. However, working with YouTube data in a digital ethnography also presents a number of challenges and potential limitations that need to be acknowledged.

While YouTube videos potentially give us access to study life and behavior in real world settings that are difficult to study with traditional ethnographic techniques, it is important to notice that the context in which these videos were produced is very different from the context of an observational study by an ethnographer or anthropologist. Firstly, it is important to acknowledge that ethnographic engagements are rarely matters of detached objective observation but rather built upon researcher participation, or membership, where the observer can inquire into the details of a situation, and move beyond their own "local" reasoning into the reasoning of the observed parties. In the case of using YouTube videos, or similar video material detached from the actual situation, this participation is obviously restricted, limiting the extent of possible insight. In relation to this, it is also important to recognize that some contextual cues are bound to be missing from the relatively low-fidelity recordings posted on YouTube (compared to in-person observation), and that context, which might have been recognized as important by a researcher, may not have been attended to by the camera-person due to the different purpose of the recording. This means that important things may well be happening off-screen and go unnoticed by the digital ethnographer. Consequently, YouTube videos cannot be treated as documentation of particular situations to the same degree that researcher video recordings might be. These limitations need to be considered when analyzing YouTube video data.

Another challenge in the use of YouTube video material is their potentially “performative” nature. Here it is important to recognize that YouTube videos are particular media artifacts in their own right, and not necessarily neutral recordings of the world around us. YouTube videos are often filmed and published for a particular purpose, and may carry an implicit view or perspective of the person who created and uploaded it. This means that YouTube videos are not always pure captures of home life, but are sometimes “staged” performances for the camera and for the YouTube medium. This clearly introduces a potential Hawthorne Effect (possibly different from the one introduced by researcher presence), which needs to be taken into consideration when analyzing such videos.

Despite these challenges and limitations, however, we find that there is some merit in exploring the use of YouTube videos as a source of data for digital ethnography in CSCW and HCI research, provided that they are acknowledged for what they are, and what they are not, in relation to the observational power that we assign to them. In fact we feel that the vast amount of video recordings from ordinary people’s lives is a potentially rich source of additional data for ethnographic research in CSCW and HCI. But as with any new source of data, of course, we need to learn how to best analyze it, and understand its limitations.

Method

As people voluntarily and enthusiastically share a variety of different cooking situations from their homes on YouTube, we chose this as our data source for investigating the physical interactions of people cooking together at home. Through “purposeful sampling” [47] of YouTube material we were able to rapidly gain access to a breadth of data about a variety of people’s social cooking activities across different generations, cultures and situations.

Using the keyword phrase “cooking together”, 169 videos were found on YouTube through a search on 15th November 2010. After discarding duplicates and unrelated videos (e.g. a song with the words “cooking together” in the lyrics), a final set of 61 videos of people cooking together remained. Three trained researchers all analyzed these 61 videos from three different conceptual perspectives (using different analytical lenses): 1. *Spatial Arrangements*; 2. *Social Interactions*; and 3. *Cooking Situations*. Early findings from our analysis of spatial arrangements are described in a previous paper [44], while in this paper we are looking across all three perspectives.

As the first step of our analysis, inspired by Blythe and Cairns [3, 4], we conducted a qualitative content analysis of the 61 videos in order to establish their purpose, what they were about, who was in them, etc. This analysis showed that the videos spanned a range of genres, people and social settings. The majority of videos were made by families and friends in their home kitchens. The largest group were home videos showing amateurs cooking together as part of

daily life (28 videos) or for sharing cooking advice with others (13 videos). Others were “mini documentaries” with both professional and amateur cooks (10 videos). A small number were educational to teach children how to cook (5), or professional cooking shows and ads for such shows (5).

As the second step, we used Kendon’s diagramming practice for recording F-formations at a birthday party [25] to “transcribe” the videos, recording patterns of behavior by creating a series of maps of the physical kitchen layouts (figure 3), showing the arrangements and movements of people and cameras within that space.

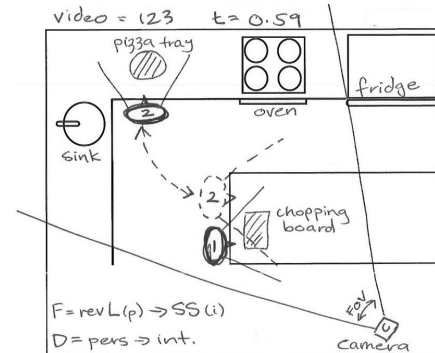


Figure 3. Map transcription from video 123 at t=0:59

Each map is time stamped (t) and records a newly established position of individuals. People are shown as numbered ovals with two lines extending to show their transactional segment, and hence where these intersect to create an F-formation. Previous positions and paths of movement through spaces are represented, as are kitchen artifacts currently being used. We also documented the virtual position of the camera and its field of view. Finally, we noted each person’s current activity, the topic of conversation (from the audio track), and the focus of the current camera shot.

As the third step, we examined the map transcriptions of the 61 videos through applied thematic analysis [14]. As our guiding themes of analysis we used Kendon’s framework of F-formations [25] to identify the interpersonal spacing and orientation between the cooks, and Hall’s framework of interpersonal distances [16] to identify their proximity. Subtle differences in F-formations, such as the side-by-side versus the wide-V formation, were established by looking at the overlap between people’s transactional space, and the angle between their upper bodies (less than 180°). While the map transcriptions mostly provided enough detail for the applied thematic analysis, such subtleties sometimes required going back to the video recordings.

As the final step, we revisited the identified formations in the 61 videos, using timestamps from the map transcription. Looking at these formations we were able to further identify three distinguishable task-related configurations, and four related collaboration-oriented actions. This gave us an awareness of different interdependencies between



Figure 4. Working on related tasks with: side-by-side (top row) and L-shaped (bottom row)

cooking tasks, interpersonal interactions, and spatial patterns of cooks in respect to the kitchen layout.

FINDINGS

In the following sections, we describe and illustrate the different cooking situations identified in our analysis, and the spatial formations that characterize them. These are three distinguishable task-related configurations: working on *related tasks*, working on *independent tasks* and working on *shared tasks*; and four related collaboration-oriented actions: *observing*, *checking*, *helping* and *showing*.

Task-Related Configurations

Working on Related Tasks

When cooking together people often work on related tasks, e.g. chopping up vegetables for a meal. In such situations, people will distribute sub-tasks between them, and they will often be working in parallel on tasks that are related to each other, but can advantageously be done at the same time. This often happens, for example, during the early stages of producing a meal where various food materials and ingredients need to be prepared for later stages, including cleaning and chopping up vegetables, seasoning meats, and mixing ingredients. In these situations, we found that people often work in either Side-by-side or L-shaped formation, depending on the layout of the specific kitchen (see figure 4). In kitchens with long bench tops and bench tops placed along a wall, the Side-by-side formation is prevalent. In kitchens with freestanding benches accessible from several sides, the L-shaped formation is also characteristic of this situation. The reason for these formations happening while working on related tasks is that these tasks often naturally take place in the same area of the kitchen, requiring the collaborating cooks to negotiate this space between them. This can be done by standing abreast or orthogonal to each other, depending on the layout and other affordances of the physical cooking space.



Figure 5. Working on independent tasks with: Z-shaped (top row) and reverse L-shaped (bottom row)

Working on Independent Tasks

The distribution of responsibilities while cooking together also happens in relation to tasks that are not closely related, e.g. one mixing up ingredients while another is washing up dishes. This means that at some point of the joint production of a meal, people start working individually on independent tasks. These independent tasks are typically coordinated, but do not need close collaboration and may be quite different. This often happens at the intermediate and late stages of meal production, and involves working with different food materials, working on different dishes, and using different types of tools. In these situations we found that people tend to spread out more in the shared kitchen space, and enter into formations that are quite different from those found in relation to interpersonal communication. We especially found that while working on independent tasks, people often enter into Z-shaped or reverse L-shaped formations, subject to the layout of the kitchen and location of particular artifacts such as ovens, stoves, sinks, etc. (see figure 5).

In kitchens with two parallel benches, the Z-shaped formation is prevalent. In kitchens with orthogonal benches, formed in a corner, the reverse L-shaped formation is prevalent. In kitchens with freestanding cooking islands, both Z-shaped and reverse L-shaped formations occur. The reason for these two formations happening while working on independent tasks is that they typically involve working in different places in the kitchen, resulting in the collaborating cooks facing away from each other, while still talking to each other.

Working on Shared Tasks

Contrasting distributed work on independent tasks, cooking together also involves situations of closer collaboration where people work together on the same shared task. This happens at all stages of the meal production, from preparation of ingredients, handing over or exchanging food materials at intermediate stages, to putting together and



Figure 6. Working on shared tasks with: wide V-shaped (top row), semi-circular (middle row), and vis-à-vis (bottom row)

arranging the final meal, and typically involves working together on the same food materials or with the same tools. In these situations we found that people often work in wide V-shaped, Semi-circular or Vis-à-vis formations depending on group size and the layout of the kitchen (see figure 6).

In groups of two the wide V-shaped formation is very common, and typically occurs when working along the same kitchen bench. Here it is often observed that two collaborating cooks will shift between working Side-by-side on related tasks to a wide V-shaped formation when exchanging or mixing the materials they were each working on independently. In groups of more than two, the wide V-shaped formation is often extended into a Semi-circular formation with people standing around a common focus point, such as a bowl or tray on the bench, or a pot or pan on the stove. In kitchens with a freestanding bench, working on a shared task is sometimes done across the bench from opposite sides in a Vis-à-vis formation.

Collaboration-oriented actions

Observing

While cooking together is mostly an activity where all people involved are active, we further identified situations where only one person would be working or cooking while others just passively observed this person. This often happens when one person is demonstrating a cooking technique, or explaining something related to a particular cooking task or food product. It also happens when a person has finished what they were doing but the other still has something to do. In these situations people tend to enter



Figure 7. Observing with: L-shaped - end of bench (top row), L-shaped – along bench (middle row), and spooning (bottom row)

into the facing formation typically associated with interpersonal communication about a disembodied topic, namely the L-shaped formation (see figure 7). This means that they place themselves orthogonal to the person cooking and sometimes the cooking bench. Depending on the kitchen layout, this can either be along one side of the bench, or, if the bench sticks out into the room, by standing at the end of it. Alternatively, in some situations, people enter into a Spooning formation observing over the shoulder of the person in front of them.

Checking

Somewhat related to the situations where people observe someone else who is currently engaged with a cooking task, people cooking together often check up on what their co-cooks are doing. Sometimes this would happen only briefly while in between doing other tasks and served the general purpose of creating and maintaining a level of common ground throughout the overall joint activity of creating the meal. In other situations where children were involved in the cooking activity, checking would happen when parents (or other adults) checked up on the task of the children. In both situations, people are able to align what they do with the tasks and doings of others, and they are also able to intervene if deemed necessary. The difference between observing and checking is that observation is typically passive and removed, while checking is active and focused with the person knowing what they are looking for. When checking up on what others are doing, people typically enter into formations looking over the other person's shoulder – sometimes with intimate proximity – in order to



Figure 8. Checking with: spooning-intimate (top row), spooning-personal (middle row), and L-shaped (bottom row)

be able to see up close what the other person is seeing and doing. Depending on their relationship, formations are typically the Spooning or L-shaped type (see figure 8). If the collaborating cooks are intimate others (i.e. partners, parents with young children, close friends, etc.) the Spooning formation is often observed. If they are not, then more distance and personal space is usually maintained, which is possible in L-shaped formations.

Showing

When people cook together they sometimes need to intervene with what the other person is doing and show them how to do it, e.g. adding seasoning to a meal. This means that they will sometimes enter into the other person's transactional space and modify their actions or sometimes even take over completely. This often happens during the middle or final stages of meal production where attention to detail is important, and involves specific actions such as adding herbs or spices, stirring or turning, removing food from a heat source, changing the temperature of the oven or stove, etc. In these situations we found that people very often enter into Spooning configurations with partially or completely overlapping transactional spaces in front of them (see figure 9). The reason for this particular formation, which is not usual during, for example, normal interpersonal communication, is that intervention is often something that needs to happen quickly, and will only last for a very short time before control is handed back. Hence, when entering into a Spooning configuration one does not force the other to withdraw from the task they are working on – in fact by standing behind someone in a Spooning formation at a bench or stove it becomes rather difficult for



Figure 9. Showing with spooning-intimate (top row), spooning-personal (middle row), and L-shaped (bottom row)

the other person to withdraw. When showing in a Spooning formation, you are merely providing an extra pair of hands and eyes for a short period of time.

Depending on their level of intimacy, people may show things in a way that is less intrusive of the other person's physical space. Here the L-shaped formation allows for reaching into another person's transactional space from either side without bodies touching or breaching the other person's intimate distance (figure 9, bottom left). The L-shaped formation can also be observed when the intervention is major and will not be brief, for example when taking over a task completely (figure 9, bottom right).

Helping

When producing a meal together, situations sometimes occur that require very close and coordinated collaboration between two or more people acting together as one to help each other. This typically happens when a task simply requires more than one pair of hands, such as operating a particular tool, shifting food between pots or bowls, moving something in or out of the oven, etc. It also happens if a person simply needs help with a particular task because they are inexperienced with it, for example when cooking with children. Situations of helping occur at all stages of meal production, from preparing food materials, cooking different parts of the meal, to putting together and arranging the final meal. Depending on the specific task, the kitchen layout, the number of people, and their relationship, it leads to a variety of different physical configurations of people, most notably Vis-à-vis, L-shaped, Spooning and Semi-circular formations (see figure 10). Depending on the layout



Figure 10. Helping with vis-à-vis (top row), L-shaped (middle row), and spooning and semi-circular (bottom row)

of the kitchen, the Vis-à-vis formation allows two people to help on food and tools in the transactional space between them, such as grating a piece of fruit (figure 10, top left) or shifting something hot or heavy (figure 10, top right). If working along the same side of a kitchen bench, or directly in front of an appliance, helping often takes place by entering into L-shaped or Spooning formations. As described earlier, the L-shaped formation is the less intimate of the two, and allows for the creation of a shared transactional space by standing orthogonal to each other and reaching into the area in front without breaching intimate distance and having bodily contact (figure 10, middle left and right). The Spooning configuration, however, allows two people to completely share a transactional space in front of them both, simply allowing more than one pair of hands to operate in this space, like when scraping food out of a pot (figure 10, bottom left). In Semi-circular formations, people are able to help with food or tools in their shared transactional space, like operating a pasta maker (figure 10, bottom right).

DISCUSSION

The kitchen as a design space

We reviewed CSCW and HCI literature for research on the design of innovative technologies for the home, specifically the kitchen. Of the designs that we found, most aim to provide cooking assistance in the form of digital recipe systems or how-to-cook applications. Notable examples of purpose built kitchens which explore this concept are AR Kitchen [5] which gives real time recipe assistance to cooks by projecting steps, quantities and temperatures onto

kitchen surfaces, and Ambient Kitchen [42] which uses a combination of projectors, cameras, RFID tags and readers, object mounted accelerometers, and under-floor pressure sensing to guide people with dementia through a cooking session. Both involve natural interactions with technology. In the AR Kitchen people use a combination of touch and mid-air gestures to control the advance of recipe instructions. In the Ambient Kitchen people interact with the system by physically moving around the kitchen, by working with augmented cooking utensils and by turning pages of a paper based recipe book.

Natural user interactions in the kitchen make good sense when we think about the other roles that hands must play in the activity of cooking. People are handling ingredients and cooking implements while operating kitchen equipment such as stoves and ovens. Their hands go through various stages of business, wetness and messiness making traditional direct manipulations interfaces, and interaction devices such as keyboards and mice problematic. Some systems have tried to work with this problem. The eyeCOOK recipe system progresses a cook through the recipe using the natural input modalities of eye movement and speech [5]. Cooking Navi [17] uses a foot switch to replace the mouse, and a waterproof touch pen for use on a touch screen for operating their daily cooking assistance system. More recently, Panger [46] explored the use of a Microsoft Kinect in the kitchen to implement a recipe navigator, timer and music player using touch-less gestural control during the activity of cooking with messy hands. Although these designs explore natural interactions in the kitchen, they are aimed at supporting a single cook, which in reality is not always the case. Often the kitchen becomes a meeting place for families and friends, where people come together and share their daily experiences, while participating in the activity of preparing a meal.

Of more interest to us within the context of our study were: 'Living Cookbook' [52] and 'Personal Chef' [33]. The 'Living Cookbook' enhances the social aspects of cooking while learning to cook. It includes a touch screen display mounted on a cupboard in front of the cooks, in a vis-à-vis configuration. Videos of friends/instructors cooking, ingredients and recipes are shown on the screen, and sometimes involve multiple cooks at the same location. Multiple cameras were used to video both the cook and the preparation area. Similarly 'Personal Chef' aims to support the social process of cooking using a multi-display, dual-perspective interactive kitchen counter. Their upright display is also in a vis-à-vis configuration, behind the stove with a second display on the counter top. So, when these designs add the presence of a video tutor or video link, they place it in the vis-à-vis configuration with the cook – just like traditional video conferencing - but in the total number of formations (929) that we identified in the videos of people cooking together, only 7% were vis-à-vis. It was not a typical way for people to relate to each other when

cooking together, nor was it used for the activities of observing, checking, showing or helping.

In designing for the situation of people cooking together, a deeper understanding of the social and physical nuances of cooking becomes extremely important. The systems we reviewed, rather than being built on empirical findings, were technology driven, deployed as technology probes [24], and then evaluated with users to find design ideas and opportunities for redesign. Our approach in exploring human-food interaction technologies as a potential design space [13] was to study multiple cooks coordinating and cooperating in a home kitchen. From this we have provided an empirically devised catalogue of physical interactions in the kitchen based on the configurations and actions people engage in when cooking together. From this, we can then start to think about different design opportunities for new technologies in the kitchen.

Designing for social natural user interaction

Motion sensing natural interaction technologies, such as Microsoft's Kinect, invite a new approach to the interaction design of technologies for the kitchen. Rather than the kinds of direct manipulation interfaces that predominate most cooking support systems surveyed, we can start to think about proxemics and formations of multiple people as inputs to technology. By understanding concurrent activities and movements, within a variety of kitchen layouts, we can imagine a system that aligns with and augments different aspects of the cooking experience for people sharing the activity of cooking. With multiple cooks it becomes important to consider both the physical aspects of food preparation and the social aspects of cooking together. Our findings give a useful starting point from which to consider the kinds of opportunities available for touch-less gestural control of technology in the kitchen. For example, one could imagine gestural controls designed specifically to be available and functioning from a range of different interpersonal spatial arrangements and distances. Studies that propose informed designs for natural user interfaces for emerging technologies, such as Microsoft's Kinect, Nintendo Wii, and wearable sensing devices in non-gaming areas are starting to appear in CSCW and HCI literature. This is especially pertinent in situations where users are trying to achieve important and timely tasks, and where touch-less interaction and gestural input are necessary for reasons of hygiene (e.g. surgery [40]), messiness (e.g. cooking [46]) or movement (e.g. physical training [11], learning physics theory [29]).

In one of our related studies on collocated intergenerational cooking [37] we watched different combinations of grandchildren, parents and children cooking together in their home kitchens. In our findings, we note that the hands of cooks are often busy or fully engaged, using tools, picking up ingredients, or communicating to each other through gestures and even physical touch. So when considering the usefulness of natural interaction

technologies in the home kitchens we need to be mindful that the gross physical movements required by contemporary gestural sensors, such as Microsoft's Kinect, and input from multiple closely positioned users could be problematic in a cluttered cooking space. The amount of movement in the kitchen, comprising body movements, as well as hands moving to communicate and undertake cooking tasks, has implications for technology attempting to capture specific inputs.

It is therefore valuable to be able to assign meaning to different configurations of spatial positioning within the cooking context. Design approaches should not intervene directly in cooking activities to the extent that they distract from the experience of cooking but rather that they complement it [37]. In our user-centered approach to designing technology for the kitchen, it is the knowledge and understandings of current practice generated through studies of people cooking together that should be used to inspire creative interventions that add to the experience.

We suggest that in designing technology to augment the cooking together experience, displays, touch screens and cameras installed at a single location in the kitchen do not cater for the varied physical movements and social interactions that take place around food preparation and make it a meaningful and positive experience. The kinds of designs we envisage use multiple sensing devices, for example, several Kinect camera's taking in people's positions from various angles, not just face-to-face. Or using ceiling mounted Kinect camera's to recognize and relate our task-related configurations to kitchen architecture and cooking artifacts, or to register proxemic distances between cooks. When combining this with gestures as input to the system it should be possible to provide both ambient and instrumental additions to the cooking experience that align with human actions and desires.

Digital Ethnography with YouTube Videos

Our study is based on a data set of YouTube video postings, which we have analyzed by looking at people's spatial arrangements and interpersonal distances. As we have discussed earlier, although YouTube offers an appealing amount of rich video data, and in some sense a window into people's lives, using video postings as part of an ethnographic study is not without challenges and potential limitations. Firstly, we were not able to enquire further into what was happening in the kitchens, and outside the scope of the camera, as one would during participating observation. In our particular study, however, we did not find this to be a significant problem as the spatial arrangements and interpersonal distances could be observed visually, and did not rely on the reasoning of the observed participants. To investigate further *why* people arrange themselves in the way observed, and *what it means* that they arrange themselves in these ways, one would have to include additional techniques, such as cued recall, or direct observation and interviews as we have done in a follow-up

study [37]. Secondly, because YouTube videos are not just neutral video recordings, but particular media artifacts produced and posted with a purpose and audience in mind, one has to be sensitive to how this may potentially impact what is seen. In our study, we were particularly conscious and attentive to the fact that some videos might be staged performances, and perhaps take place in studio settings, which could influence the way people arrange themselves spatially. In looking at the videos, however, we were surprised to see that most take place in home kitchens and the camera typically appears to recede into the background, with people clearly focusing on the social cooking activity, with very little performing for the camera (see figures 4-10). Hence, the presence of the camera, and the fact that the video might end up on YouTube, did not appear to influence the way people were cooking together and arranging themselves spatially in the kitchens.

CONCLUSIONS

We have taken an empirical approach to understanding what it is that people do when they are cooking together in their home kitchens. We have analyzed 61 YouTube videos in a digital ethnography and identified formations made while people share the activity of food preparation. We did this to understand the spatial patterns that people form when cooking together and relate this to cooking activities and kitchen architecture.

Our contribution includes 8 identified spatial formations of people cooking together, which are very different from formations characterizing purely conversational situations. We have also identified three distinct task-related configurations: working on *related tasks*, working on *independent tasks* and working on *shared tasks*; and four related collaboration-oriented actions: *observing*, *checking*, *showing* and *helping*. This catalogue of the relationships between formations, cooking situations and kitchen architecture is a way to understand and represent the physical interactions of people cooking together.

Our research has shown that cooking together is a complex activity and based on our findings, we have reflected on possible ideas and implications for technology designed to enhance the cooking experience that accounts for the fact that people coordinating in a kitchen do so by positioning and interweaving their bodies within the available space to connect, converse, coordinate, cooperate and cook.

REFERENCES

1. Arnold, J.E., Graesch, A.P., Ragazzini, E. and Ochs, E. *Life at Home in the Twenty-First Century: 32 Families Open Their Doors*. The Cotsen Institute of Archaeology Press. 2012.
2. Bell, G., Blythe, M. and Sengers, P. Making by Making Strange: Defamiliarization and the Design of Domestic Technologies. *TOCHI*, 12, 2 (2005), 149-173.
3. Blythe M. and Cairns P. Tenori-On Stage: YouTube As Performance Space. In *Proc. NordiCHI 2010*, ACM (2010), 72-81.
4. Blythe, M. and Cairns, P. Critical Methods and User Generated Content: the iPhone on YouTube. In *Proc. CHI 2009*, ACM (2009), 1467-1476.
5. Bonanni, L., Lee, C. and Selker, T. Attention-Based Design of Augmented Reality Interfaces. In *Proc. CHI 2005 EA*, ACM (2005), 1228-1231.
6. Bradbury, J., Shell, J. and Knowles, C. Hands on Cooking: Towards an Attentive Kitchen. In *Proc. CHI 2003 EA*, ACM (2003), 996-997.
7. Chamberlain, A. and Griffiths, C. Wild Food Practices: Understanding the Wilder Implications for Design and HCI. In *Proc. UbiComp 2013*, ACM (2013), 575-584.
8. Clear, A., Hazas, M., Morley, J., Friday, A. and Bates, O. Domestic Food and Sustainable Design: A Study of University Student Cooking and its Impacts. In *Proc. CHI 2013*, ACM (2013), 2447-2456.
9. Crabtree, A. and Rodden, T. Domestic Routines and Design for the Home. *Computer Supported Cooperative Work*, 13, 2 (2004), 191-220.
10. Crabtree, A., Rouncefield, M. and Tolmie, P. Cooking for pleasure. In P. Tolmie and M. Rouncefield (Eds.) *Ethnomethodology at Play*. Ashgate, 2013, pp. 21-52.
11. Fritz, T., Huang, E., Murphy, G. and Zimmerman, T. Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness. In *Proc. CHI 2014*, ACM (2014), 487-496.
12. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R. and Wang, M. Proxemic interactions: the new ubicomp?. *Interactions*, 18, 1 (2011), 42-50.
13. Grimes, A. and Harper, R. Celebratory Technology: New Directions for Food Research in HCI. In *Proc. CHI 2008*, ACM (2008), 467-476.
14. Guest, S., MacQueen, K. and Namey, E. *Applied Thematic Analysis*. SAGE, 2012.
15. Haines, V., Mitchell, V., Cooper, C. and Maguire, M. Probing user values in the home environment within a technology driven Smart Home project. *Personal and Ubiquitous Computing*, 11, 5 (2007), 349-359.
16. Hall, E.T. *The Hidden Dimension*. Anchor Books, 1966.
17. Hamada, R., Satoh, S., Okabe, J., Sakai, S., Ide, I. and Tanaka, H. Cooking Navi: Assistant for Daily Cooking in Kitchen. In *Proc. MM 2005*, ACM (2005), 371-374.
18. Hamer, B. *Kitchen Stories* (Salmer fra kjøkkenet) [Motion picture]. Norway, IFC Films, 2003.
19. Harley, S. and Fitzpatrick, G. Creating a conversational context through video blogging. *Computers in Human Behaviour*, 25, 3 (2009), 679-689.

20. Heaton, L. Preserving Communication Context: Virtual Workspace And Interpersonal Space In Japanese CSCW. In *Proc. of Cultural Attitudes Towards Com. and Tech.* 1998, U Syd (1998), 207-230.
21. Huettenrauch, H., Eklundh, K., Green, A., Topp, A., 2006. Investigating Spatial Relationships in Human-Robot Interaction. In *Proc. IROS 2006* (2006), 5052-59.
22. Hung, H. and Kröse. Detecting F-formations as Dominant Sets. In *Proc. ICMI 2011* (2011), 231-238.
23. Hupfeld, A. and Rodden, T. Laying the Table for HCI: Uncovering Ecologies of Domestic Food Consumption. In *Proc. CHI 2012*, ACM (2012), 119-128.
24. Hutchinson, H., Mackay, W., Westerlund, B. et al. Technology probes: inspiring design for and with families. In *Proc. CHI 2003*, ACM (2003), 17-24.
25. Kendon, A. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University, 1990.
26. Kendon, A. Spacing and orientation in co-present interaction. In *Proc. COST 2009*, Springer (2010), 1-15.
27. Kerr, S., Tan, O. and Chua, J. Cooking personas: Goal-directed design requirements in the kitchen. *Int. J. Human-Computer Studies*, 72 (2014), 255-274.
28. Lareau, A. *Unequal Childhoods: Class, Race, and Family Life*. University of California Press. 2011.
29. Lewis, S., Acholonu, U. and Ju, W. Using Low Cost Game Controllers to Capture Data for 6th Grade Science Labs. In *Proc. CSCW 2012*, ACM (2012), 1115-1124.
30. Marquardt, N., Hinckley, K. and Saul Greenberg, S. Cross-device interaction via micro-mobility and f-formations. In *Proc. UIST 2012*, ACM (2012), 13-22.
31. Marshall, P., Rogers, Y. and Pantidi, N. Using F-formations to analyse spatial patterns of interaction in physical environments. In *Proc. CSCW 2011*, ACM (2011), 445-454.
32. Masten D. and Plowman, T. Digital Ethnography: The next wave in understanding consumer experience. *Design Management Journal*. 14, 2 (2003), 75-81.
33. Mennicken, S., Karrer, T., Russell, P. and Borchers, J. First-Person Cooking: A Dual-Perspective Interactive Kitchen Counter. In *Proc. CHI 2010 EA*, ACM (2010), 3403-3408.
34. Mentis, H., O'Hara, K., Sellen, A. and Trivedi, R. Interaction proxemics and image use in neurosurgery. In *Proc. CHI 2012*, ACM (2012), 927-936.
35. Morrison, C., Jones, M., Blackwell, A. and Vuylsteke, A. Electronic patient record use during ward rounds: a qualitative study of interaction between medical staff. *Critical Care*, 12, 6 (2008), R148.
36. Nagel, K., Hudson, J. and Abowd, G. Predictors of Availability in Home Life Context-Mediated Communication. In *Proc. CSCW 2004*, ACM (2004), 497-506.
37. Nansen, B., Davis, H., Vetere, F., Paay, J., Skov, M. and Kjeldskov J. Kitchen Kinesics: Gestural Communication within Familial Cooking Interactions. In *Proc. OZCHI 2014, ACM and SIGCHI* (2014)
38. Obata, A. and Sasaki, K. Office Walker: a virtual visiting system based on proxemics. In *Proc. CSCW 1998*, ACM (1998), 1-10.
39. Ochs, E. and Kremer-Sadlik, T. (Eds.) *Fast Forward Family*. University of California Press. 2013.
40. O'Hara, K., Harper, R., Mentis, H., Sellen, A. and Taylor, A. On The Naturalness Of Touchless: Putting The "Interaction" Back Into Nui. *TOCHI*, 20 (2013).
41. O'Hara, K., Kjeldskov, J. and Paay, J. Blended Interaction Spaces for Distributed Team Collaboration. *TOCHI*, 18, 1 (2011), Article 3.
42. Olivier, P., Xu, G., Monk, A. and Hoey, J. Ambient Kitchen: designing situated services using a high fidelity prototyping environment. In *Proc. PETRA 2009*, ACM (2009), Article 47.
43. Paay, J., Kjeldskov, J., Skov, M. and O'Hara, K. Cooking Together: A Digital Ethnography. In *Proc. CHI 2012 EA*, ACM (2012), 1883-1888.
44. Paay, J., Kjeldskov, J., Skov, M. and O'Hara, K. F-formations in Cooking Together: A Digital Ethnography using YouTube. In *Proc. Interact'13* (2013), 37-54.
45. Palen, L. and Hughes, A. When home base is not a place: parents' use of mobile phones. *Personal and Ubiqu. Comp.* 11, 5 (2007), 339-348.
46. Panger, G. Kinect in the Kitchen: Testing Depth Camera Interactions in Practical Home Environments. In *Proc. CHI 2012 EA*, ACM (2012), 1985-1990.
47. Patton, M. *Qualitative Research & Evaluation Methods*. Sage Publications, Thousand Oaks, CA, 2002.
48. Reitmaier, T., Benz, P. and Marsden, G. Designing and theorizing co-located interactions. In *Proc. CHI 2013*, ACM (2013), 381-390.
49. Rogers, Y. *HCI Theory: Classical, Modern, and Contemporary*. Morgan & Claypool, 2012.
50. Rotman, D. and Preece, J. The 'WeTube' in YouTube – creating an online community through video sharing. *International Journal of Web Based Communities*, 6, 3 (2010), 317-333.
51. Sellen, A., Harper, R., Eardley, R., Izadi, S., Regan, T., Taylor, A. and Wood, K. HomeNote: Supporting Situated Messaging in the Home. In *Proc. CSCW 2006*, ACM (2006), 383-392.
52. Terrenghi, L., Hilliges, O. and Butz, A. Kitchen Stories: sharing recipes with the Living Cookbook. *Personal and Ubiqu. Comp.*, 11, 5 (2007), 409-414..