ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/43465830

Work at hand: an exploration of gesture in the context of work and everyday life to inform the design of gestural input devices

Article · January 2003

Source: OAI

CITATIONS
12

READS
172

5 authors, including:

Margot Brereton
Queensland University of Technology
228 PUBLICATIONS **2,119** CITATIONS

SEE PROFILE

Nicola J Bidwell
University of Namibia
109 PUBLICATIONS **1,714** CITATIONS

SEE PROFILE

Jared Donovan
Queensland University of Technology
34 PUBLICATIONS **253** CITATIONS

SEE PROFILE

Brett Campbell
Utah Department of Education
10 PUBLICATIONS **45** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

OzCHI 2016 Short Paper View project

Digital Matatus View project

All content following this page was uploaded by Margot Brereton on 27 May 2014.

The user has requested enhancement of the downloaded file.

Work at Hand: An exploration of gesture in the context of work and everyday life to inform the design of gestural input devices

Margot Brereton, Nicola Bidwell, Jared Donovan, Brett Campbell, Jacob Buur

School of Information Technology and Electrical Engineering
The University of Queensland
St Lucia QLD 4072 Australia,
Mads Clausen Institute for Product Innovation
University of Southern Denmark
Grundtvigs Allé 150 – DK 6400 Sønderborg

{margot, bidwell, jared, brettc}@itee.uq.edu.au, buur@mci.sdu.dk

Abstract

This research explores gestures used in the context of activities in the workplace and in everyday life in order to understand requirements and devise concepts for the design of gestural information appliances. A collaborative method of video interaction analysis devised to suit design explorations, the Video Card Game, was used to capture and analyse how gesture is used in the context of six different domains: the dentist's office; PDA and mobile phone use; the experimental biologist's laboratory; a city ferry service; a video cassette player repair shop; and a factory flowmeter assembly station. Findings are presented in the form of gestural themes, derived from the tradition of qualitative analysis but bearing some similarity to Alexandrian patterns. Implications for the design of gestural devices are discussed.

Keywords: gestures, information appliances, video card game, natural interfaces, video interaction analysis, patterns.

1 Motivation

Our research explores gestures used in the context of activities in the workplace and in everyday life in order to understand requirements and devise concepts for the design of gestural information appliances, in particular, gestural input devices. While there is a large corpus of documented research focusing predominantly on how they arise in conversation (e.g. MacNeill 1985), there has been little exploration of gestures in the context of activities, object manipulations, information work and interactions with the physical environment. As a result there is little understanding of how to apply research on gestures to specific interface and information environment design problems. Our research examines and interprets gestures in the context of activity through examining videotape in real time, rather than frame by frame. This approach leads us to understand the quality, character, use and function of gestures in episodes of activity. Based on this understanding we propose to identify how gestures could be designed for use with gestural appliances.

2 Designing Gestural Interaction

Gestures are attractive as an input vehicle because they come naturally to the human and exercise the human body appropriately, rather than constraining the human to the postures and patterned actions of keyboard and mouse type technologies. Using gesture may help to avoid common problems of hunched shoulders, repetitive motion injury and eye strain. However as with any technological interventions, gestural devices must be designed with an understanding of the activities and work practice that they seek to support.[1] There is substantial evidence that the origin of gestures used in communication lies in tool use and object manipulations (see; Spinney 2000 for review). This is further inspiration for exploring how we might derive inspiration for the design of new kinds of tools - gestural input devices - from gestures used in the workplace.

Gestures may arise spontaneously in conversation or they may belong to discrete gesture sets which have been designed or evolved for the purposes of communication in specific contexts, e.g. diving signals, semaphore, signing language for the hearing and speech impaired, etc. Where gestures are designed for activating information appliances, they should demand little cognitive attention, be easily remembered, repeated, recognized and unobtrusive, yet sufficiently distinct from natural gestures, such that a natural gesture does not effect an unintended operation. This will allow the human to maintain naturalistic social interactions in physical space. Whether the design of such distinct gestures is possible such that they can merge seamlessly into natural human interaction is an ongoing research question.

Our exploration of gestures was motivated in particular by our interest in designing wireless gestural input appliances. The appliances of interest are self contained, not relying on external cameras, screens or devices in order to effect gesture detection, as do many existing gesture recognition systems which are designed primarily

Copyright © 2002, Australian Computer Society, Inc. This paper appeared at the *Fourth Australasian User Interface Conference (AUIC2003),* Adelaide, Australia. Conferences in Research and Practice in Information Technology, Vol. 18. R. Biddle and B. Thomas, Eds. Reproduction for academic, not-for profit purposes permitted provided this text is included.

[1] Since interventions will inevitably change work practice, the goal is to understand and design in such a way that work practice is improved rather than obstructed or requiring excessive work-arounds.

for instrumented rooms, CAD systems or Virtual Reality. They are small, package-able into devices such as rings, pens and watches. Their likely limitations are in the amount of gestural data that they can capture and transmit when subject to the constraints of battery power usage. However, even with a limited gesture recognition capability, they have the potential to be powerful and flexible in configuration by working in concert with other devices such as speech and location detection devices.

Our research has proceeded along two paths in collaboration with a team of electrical engineers. While the engineering team is determining the capabilities of wireless gesture detection through building prototype gestural input appliances, our research investigations are aimed at understanding how gestures are used in context and in determining whether promising domains exist for the application of such appliances. This paper focuses on the latter investigation.

3 Defining Gestures

Our research corresponds with Keates and Perricos (1996) definition of gestures:

"*an actual, time-varying physical motion, or measurable intention to produce the same with specific intent and meaning on the part of the user*"

This definition enables us to explore, within the richness of all body movements, those gestures where the purpose does not seem to be primarily communicative while accepting the increasingly common sentiment of communication theorists and semioticians that all body movements communicate. Our definition places gestures for communication along a continuum that extends from body movements to manipulate objects to those that are intended to express meaning to another or to oneself. This perspective represents a broader definition of gesture than that which pervades HCI design and which focuses on "instances where an individual engages in movements whose communicative intent is paramount, manifest, and openly acknowledged" (Nespoulous, et al 1986).

The majority of endeavours seeking to migrate "natural" human-to-human communication to HCI explore gesture within a speech related communication arc (e.g. CHI'95 gesture Workshop; Coutaz and Crowley, 1995). Gestures are seen to complement or supplement verbal communication and/or reflect or facilitate the cognitive processes underlying speaking (e.g. McNeill 1992; Krauss et al. 1991, 1995, 1996; Rauscher et al. 1996). By exploring gestures within a speech related communication arc researchers have drawn upon classifications of gestures based on linguistic, paralinguistic functional and rhetorical communication models (e.g. Jakobson, 1990). This approach provides a symbolic framework for understanding gestures and allows them to be explored with reference to the functions of spoken language.

While the tradition of loosely defining gestures has allowed some freedom in visualizing and classifying gestures, many gesture taxonomies have been influenced by the communication models based on functionalist linguistics. Some classifications mirror the treatment of words, such as categorising gestures as "nouns", "adjectives", "verbs", "adverbs", etc., while others attempt to classify gestures by meaning. Pelachaud and Poggi (1997) suggest distinguishing between at least three classes of meanings in non-verbal communication relating to information on the world (deictic), the speaker's identity (iconic) and the speaker's mind (symbolic). Other classifications decompose the body into parts/subparts each bearing its own specific communication system and describe different mode-specific body signals and sets of rules to link signals and create new signals (e.g. Pelachaud and Poggi, 1997).

Gestures may perform many functions in addition to those involving the generation or supplementation of language. For example they may help us to think about and explore the physical world (Brereton and McGarry 2000, Brereton 2001), and they possibly reflect back environmental or social constraints. In this way they provide an avenue for designing information appliances that exploit the natural coupling of human perception and action. Accumulating developmental and neurophysiological evidence suggests that there may be a more direct link between body movement and meaning than via spoken language (see; Spinney 2000 for review). The neuroarchitectures supporting motricity appear to interface with those facilitating language and the construction of meaning. For example, Rizzolati et al. (2001) have discovered that neurons controlling motricity fire both when a monkey makes a particular movement and when it watches another making the same movement. This might indicate that during evolution, human language was built upon action-oriented representations (e.g. Clark 1997) by usurping neural systems that were responsible for more basic sensory and motor functions (Bates 1994). By exploring gestures in natural environments along a continuum that includes body movements around objects and with tools as well as those for communication we will be better informed in designing information objects for use in those environments.

4 Method

This paper describes a method for investigating gestures in context using the Video Card Game devised by Buur and Soendergaard (2000). In the Video Card Game, participants create new understandings of how people use artefacts through examining and discussing interactions recorded in the video media. The Video Card Game was chosen as a method over more established methods in the tradition of ethnomethodology and grounded theory such as Interaction Analysis, because it treats video material as a media for interpretation, participation and emphatic engagement after McLuhan 1964). "In this case video recordings are no longer seen as hard data but the first attempts to create stories that frame the design problem and impose order on the complexity of everyday life." (Buur, Binder and Brandt 2000). We decided against the detail of protocol analysis and coding schemes for analysing the video material, as this was an explorative study aimed at opening a broad discussion in the team. Our goal was to identify ways to understand, view and interpret gesture in the wild that could help inspire

gestural design concepts. Our epistemological stance was underpinned by the fundamental assumptions of Interaction Analysis (Jordan and Henderson, 1994) that:

• Knowledge and action are fundamentally social in origin

• Theories of knowledge and action should be grounded in verifiable observable empirical evidence

• Theorizing should be responsive to the phenomenon itself rather than to the characteristics of the representational systems that reconstruct it – analysis is done directly on videotape

However, our approach departed from that of Interaction Analysis in two respects. First, our ultimate aim is not to produce a theory but a set of ways of seeing gesture that we call "gestural themes" in order to inform design, with an acknowledgement that these views are partial and overlapping. Second, although our analysis of the tape sought to let the data speak for itself and involved playing and replaying clips to check that we had identified and understood typical ways that gesture are used, we also admitted our views as gestural beings into the discussion, rather than taking an analytic stand that only allows what is seen and heard on the tape to inform our understanding.

One can imagine a continuum of approaches that use video. One end is represented by Interaction Analysis and ethnomethodological techniques that seek as much as is possible to let the data speak for itself in order to form a theory that is as grounded to the data as possible. The other end represented by design approaches that use the video as a form of inspiration. In this case video facilitates generation of ideas by revealing specific interactions in specific contexts. This brings the interpretive role and intentions of the video recorders, editors and viewers to the foreground.

Our approach struck a middle ground that sought to ground our understandings in both our experiences as gestural beings and in the video data that we viewed as a team.

4.1 Domains of Study

The domains chosen for study involved people working or undertaking daily activities in physical environments with objects. They either worked collaboratively or alone. In the cases where they worked alone, they described what they were doing as they worked (situated interviews).

The video sequences contain:

- An experimental biologist describing experiments and apparatus in a laboratory setting
- A dentist performing a check up on a patient in a dental surgery
- Two PDA and mobile phone users sending messages to one another in an office building environment
- Deckhands, drivers, till operators and passengers travelling on a city ferry
- A repair technician fixing a CD-player and television at his workbench in a repair shop

• An assembly line worker assembling a flowmeter and instructing a member of the design team at her workbench

4.2 Video Card Game

The Video Card Game is a technique which can be applied to video interaction analysis allowing video segments to be turned into arguments to support design work (Buur, Sondergaard 2000). It provides an environment which promotes open discussion allowing the exploration of themes, patterns, and trends embedded within video data. One of the primary benefits of this technique is that it affords individuals whom are not experts at video analysis and have limited training, the ability to draw meaningful themes from raw video footage.

Video analysis is a task which is mentally intensive. As a result it can be difficult to maintain the required level of concentration and focus long enough to obtain meaningful data from an entire video. A significant strength of the Video Card Game approach to video analysis is that it commands intense focus for short periods of time as opposed to a more traditional approach where the focus is broadened to encapsulate the entire video.

Instead of watching the footage sequentially, the video is divided into smaller subsections, averaging around a minute in length. These don't necessarily have to be viewed in a strict order. This process ensures that all of the interactions contained within a short video clip can be examined in detail.

The technique is derived from a children's card game, "Happy Families", where the object is to collect families of cards. By putting frames from the video clips onto small pieces of cardboard a physical representation of a digital resource is created. An advantage of this approach is that it affords the ability to manipulate digital information in a physical space. This is an effective and efficient approach as it's believed it provides a more intuitive means of integrating information (Buur & Sondergaard 2000).

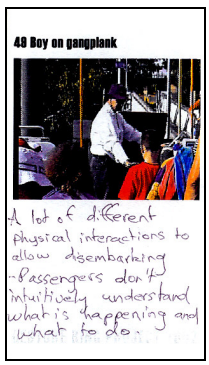

Figure 1: Sample card for video card game

The Video Card Game technique requires some preparation before the game can be played. The raw video data must be separated into short, 30 to 90 second video

clips. A single frame from each clip is placed on a small piece of cardboard, leaving enough room for some notes to be hand written underneath. Each card should be duplicated. A sample card is illustrated in Figure 1.

For the purposes of our study of gestures in context, the game was played with a design group comprised of 12 participants. Our team brings with it a diversity of backgrounds and experience encapsulating fields of study such as Information Technology, Electrical Engineering, Mechanical Engineering, Architecture, Information Environments, Work Studies, Film Production, Dance, Visual Arts, and Neurophysiology.

Our process is best explained by describing how we stepped through the video card game.

Step 1: The group divided into 6 pairs, with each pair being dealt a duplicate set of 10 cards. The cards were then divided within the pair so that each person had a copy of all 10 cards. The rules for the game were also explained.

Step 2: The pairs then watched their 10 video sequences which had been transferred to a PC (running a Windows operating system) and were viewed using Windows Media Player. Individual notes were made on the appropriate cards. This step took approximately 2 hours.

Step 3: The next 2 hours were spent with the entire team gathered together. Each player's cards were arranged and grouped in front of them on a table. This was accompanied by each player describing their groupings to the rest of the team.

Step 4: Each player selected a favourite grouping to develop further into a theme. After carefully articulating their theme to the rest of the team, cards were offered from other participants which were related to the theme. This process was repeated around the table. As themes were built up the participating video cards for each theme were attached to a large sheet of paper with the theme's title prominently displayed. We found that in some cases a video card was actually appropriate for inclusion in more than one theme. This step took approximately 1 hour.

Step 5: To conclude we spent an hour discussing the themes and what they meant for the design of gestural input devices.

Buur and Soendergaard, and collaborators (2000) have used the Video Card Game within the rubric of participatory design to create user-centred design discussions. For example, in the domain of pumps for domestic heating systems, they engaged six heating system installers and six developers with their user-centred design group to explore designs using the video card game.

In this research, rather than focus on participatory design for a particular domain, which may reveal gestural devices to be inapplicable in that domain, we (as a research team) explore gestural interaction across several domains. This is done with a view to understanding how gestures are used in a variety of contexts, to explore possible design applications and to identify domains worthy of further investigation.

5 Gesture Themes

We organized the results of our video card game in to a series of summaries that we call Gesture Themes. The intention is that they should serve as an entry point into the information and understandings that we arrived at through the process of the Video Card Game.

The Gesture Themes all follow a standard format. Each has a title. A memorable but imprecise title is preferable to a longer one that tries to exhaustively define the theme. The body of each Gesture Theme begins with one or more introductory paragraphs that describe how the gestures in the theme are used and the contexts in which they are found. Following this, one or more specific examples from the source video material are described.

These examples consist of a textual description of the instance of the theme along with a still image from the video clip. Next, a list of 'Descriptors' is given. Descriptors are words or short phrases that describe observable qualities of the theme that we consider important. Typically, they describe qualities of the gestures, the type of activity being undertaken, the nature of the work areas, and so on. The intention is that the Descriptors act as keywords that a person could use to help them recognize an instance of the particular theme.

Following on from the preceding sections is one called 'Implications' that presents speculations on how the theme might inform design. The purpose of this is to provoke both author and the reader to ask, "How could this information be useful for design of gestural information appliances?" The speculations are not intended as a definitive answer to this question.

Finally a list of clips to watch for the theme is presented. This list is basically the same as the groupings of Video Cards for the theme from the Video Card Game. However following completion of the theme description, the clips were viewed again to determine whether the phenomena that were identified in the Video Card Game, still seemed to present themselves in the data when viewed afresh after several days or weeks. We also looked through the rest of the clips for any others that were missed in the video card game.

5.1 Themes and Patterns

Initially, our themes appeared to bear some resemblance to patterns as originally conceived of by Alexander (1977) and since explored by several researchers in different domains (e.g. Erickson 2000, Crabtree 2002). However upon reflection we determined that our themes were quite distinct from patterns in the Alexandrian tradition. The themes were derived from the Video Card Game and related approaches in qualitative analysis, such as those of Strauss and Corbin 1998) which attempt to distill structure and meaning from qualitative data.

Alexandrian patterns and related approaches were conceived to succinctly describe design approaches, by presenting in a pattern a specific example, the problem it

sought to solve, a diagram of the solution, a description of the solution, the larger context etc.

Although Alexander did not intend his patterns to be used to inform design, researchers and practitioners in fields such as software engineering have adopted patterns as a succinct means of describing and sharing design approaches.

Our Gestural Themes do not identify design problems and gestural appliance solutions, and so do not have the qualities of a design repertoire. Rather they identify typical roles that gestures play in activity through qualitative analysis. Hence we choose to present our findings as descriptive themes. These include only speculations on possibilities for gestural devices.

5.2 Theme Summaries

In the following sections we present five of our Gesture Themes, Commanding Gestures, Preparatory Gestures, Gestures as Placeholders, Shared Tools-Shared Workspace and Mirroring Gestures.

5.2.1 Theme: Commanding Gestures

Commanding gestures are used to direct another person to do something. These gestures are typically simple, short, recognisable, and relatively independent of speech. They tend not to be used for every-day interactions; rather, they are used in situations where there is some extenuating factor, such as a noise level which impedes effective verbal communication, a need for urgent communication of danger, a need for discretion when a dentist wants to direct an assistant without alarming a patient etc.

Figure 2: Demonstrating screwdriver grip

Specific Examples

A designer is being supervised as she learns how to build a flowmeter for the first time. The supervisor relied upon simple, but recognisable gestures to guide the trainee in using a tool for the first time when the surrounding industrial noise prevented clear vocal instructions from being received. Figure 2 shows how the supervisor used her hands to demonstrate how to grip the tool.

Figure 3: Deckhand motioning to halt

On the ferry a young boy attempted to disembark across the boarding plank before it was completely secured, (Figure 3). The deckhand held his palm out flat signalling to stop and wait until safe to cross. The boy did not receive any further verbal guidance – it is clear that the gesture caused him to wait. In this situation the extenuating circumstance was that the deckhand needed the boy to stop immediately for his own safety. There was also the possibility that verbal instructions could be lost with the wind.

Descriptors

Commanding; Directing; Atomic gestures; Requesting; Succinct; Gestural communication; Uni-directional.

Implications

Since commanding gestures may be considered impolite when directed at people, they are only used in particular circumstances. However, a gesture directed at an inanimate object, such as a printer, fax machine or coffee machine is less likely to be considered impolite though this may vary with context and culture. The succinctness and legibility of commanding gestures would seem to make them suitable for application to the control of electronic devices. One research question is whether the ease with which humans seem to correctly interpret these gestures in context can be duplicated by electronic detection.

One particular context in which commanding gestures would be useful are those with sterility constraints, such as dental surgeries, which render some objects out of bounds to some people. The dentist cannot operate a keyboard and mouse to modify a patient record, without taking a number of sterility precautions such as sterile sleeves on computer equipment. Gestural navigation of the patient record offers a means to obviate sterility constraints.

5.2.2 Theme: Preparatory Gestures

Gestures are sometimes used as a way of preparing ourselves to undertake a task. These tasks are either ones that require being done right first time, and or ones that require a high degree of motor coordination. Preparatory gestures may rehearse the action to come, such as when a golfer makes a practice swing, or when we are using a new piece of equipment for the first time. Alternatively they may be typical gestures that precede an activity.

Figure 4: Trainee practicing with the tool

Specific Examples

As illustrated in Figure 4 the trainee adjusted her grip on the automatic screwdriver and then practiced pressing the

power switch. Only when she had completed these initial gestural interactions that gave her a feel for the tool, did she shift her focus to the task at hand, screwing the flowmeter case together.

The process of securing the boarding plank of the ferry to the pier required the deckhand to have a free range of arm movement. Figure 5a and b show how the deckhand prepared for docking by shortening his sleeves, freeing his elbows to move and creating a loop in the rope he used to dock with.


Figure 5a & 5b: Deckhand prepares to dock

Descriptors

Preparation, Internal, Reflective, Communicating to yourself.

Implications

The act of making a preparatory gesture could be used to signal that a task is about to happen. This information could be used by ubiquitous devices to decide to switch modes, retrieve data, or suppress distractions. It may be necessary for gestural appliances to have a means of determining when preparatory gestures and actions stop and execution of the task proper begins.

5.2.3 Theme: Gestures as Placeholders

Gestures may be physical, temporal and/or mental placeholders. Such gestures are characterised by being relatively temporally extended rather than of short duration. While, these gestures might arise in communication, most of the instances we observed and described involve only one person. It is interesting to observe that such gestures occur around information of very different granularities. They range from pointing to identify a single object to using several fingers to hold attention to and compare several objects. When gestures are engaging in place-holding fingers are seen to spatially "home-in" on or seek an information object.


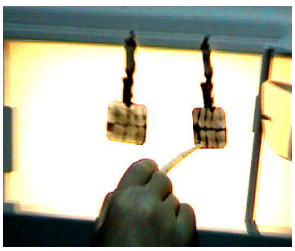

Figure 6a & b: Overt gestural communication

Specific Examples

Gestures as placeholders arise in overt gestural communication, for example when the PDA user tapped on the screen of the PDA while demonstrating its use (Figure 6a) or when the dentist pointed to the X-ray with a dental tool to indicate which tooth he was speaking about (Figure 6b).

Gestures as placeholders also arise when people focus on one piece of data in a densely populated information source. For example, the passenger held his finger to the ferry time table while he attempted to find the time of the next ferry (Figure 8a) and the dentist used his pointing instrument and mirror inside the patient's mouth while he undertook a routine "tooth-by-tooth" examination (Figure 7a).

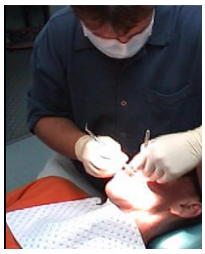


Figure 7: Focussing on one piece of information

Evidence of the use of gestures as placeholders is apparent in the process of spatially "homing-in" on or seeking an information object. For example, the PDA user and the repair technician made small, sweeping movements above a PDA's menu (Figure 6a) and over the panel of buttons on the scope (Figure 8) respectively, prior to selecting from them.


Figure 8a & b: Comparing pieces of data

Gestures as place holders also arise when comparing or integrating pieces of information that are synchronously displayed but spatially separated or asynchronously displayed but positioned in the same place. For example, the ferry passenger touched data in both the ferry time table and the ferry map at the same time to integrate time and space information (**Figure 8**a). Similarly, the repair technician touched the scope in a way that suggested he was marking a point when adjusting the scope from continuous to single sweep mode (**Figure 8**b).

As placeholders, gestures can act as reminders of intention. For example, the passenger held the front panel of the ferry time-table while he checked whether the reverse side also contained information (**Figure 9**a), he also counted off his fingers to mentally "bookmark" the information he intended to communicate when explaining

to other people what he had found out about the times of the ferry (**Figure 9**b).

Figure 9a & b: Gestures as mental placeholders

Descriptors

Waiting and preparing; Punctuating; Focussing and finding; Decision making; Interpreting and comparing; Remembering.

Implications

The natural use of gestures as placeholders has important implications for the design of ubiquitous computing devices in particular with respect to navigation in complex information spaces. In particular it implies that there could be many applications that need to detect position either relatively or absolutely, in addition to velocity and/or acceleration profiles. Most navigation methods in current computing systems use sequences of menus, pointing and clicking, rather than the more complex place holding methods indicated by multiple-finger place-holding in the natural environment. Multifinger selection methods offer the possibility of reducing the number of procedural selection operations and increasing the use of spatial visualisation and complex selection. A ubiquitous computing environment that presents active data of many forms on many different surfaces or objects may be able to use more complex pointing and place holding methods. At the dentist, the selection of a pair of instruments might be sufficient indication of context (particularly if coupled with speech input) that the patient record would automatically open at the appropriate place for modification by the dentist.

5.2.4 Shared Tools, Shared Workspace

The theme of "Shared Tools, Shared Workspace" arises when two or more people undertake different tasks but do so within the same space or using the same objects. In all instances of this theme each person is familiar with the other's task and intentions to some extent. It reminds us of the interweaving actions of dance partners or tennis players without the elements of shared movement or competition.

Specific Examples

"Shared Tools, Shared Workspace" arises in helping a co-worker perform their task, for example, when the assembly line worker reached for the barcode scanner, her co-worker anticipated that she would also need the mouse and passed it over without prompting; the dentist used his descaling pointing instrument to increase his assistant's access to the patient's mouth; the PDA user moved closer to his friend when receiving an SMS message; the ferry captain reversed slightly so that the deckhand could more easily tie off the rope.

In each case people were seen collaborating dynamically in order to utilise their working space and objects within it most effectively, for example, the assembly line worker and her co-worker shared bench space and equipment; the dentist and assistant restricted their movements to accommodate each other.

Figure 10: Using space efficiently

People initiated an action when prompted by the large scale body movements and cues of others, such as when the flowmeter assembly worker's co-worker responded to her movement in space as they used the workbench collaboratively; the dentist and assistant used each other's body movements as cues for an activity; the ferry captain reversed when he saw that the deckhand had moved the gang-plank.

People require their own place in the space, regardless of how small, and use it as a spatial reference in a task. For example, the assembly line worker put the equipment she needed in her "own" place before using it; the dentist and assistant restricted their movements to respect and accommodate each other's boundaries.

Descriptors

Anticipatory; Compensatory; Collaboratory; Waiting and turn-taking; Separate but familiar tasks.

Implications

The beautiful way in which the people we see in our clips cooperate in their work spaces makes us regret that the same level of cooperation seems so difficult to attain in computerised systems. What at first seems a natural and easy process of coordination between work colleagues turns out on closer inspection to rely on constant subtle maintenance of the workspace and one's orientation within it. Understanding in detail how physical space affords this maintenance seems integral to improving the quality of cooperation in computer mediated systems.

5.2.5 Theme: Mirroring Gestures

When we watch two people communicating we often see them mirror each other's body posture and gestures. People are seen using mirroring gestures to confirm what has just been communicated, sometimes slightly adapting the gesture to modify the meaning. They are used as a substitute for or supplement to spoken language when words can't be found or do not seem sufficient to convey meaning. In this situation gestures act as concept-tokens that other people in the conversation can use to attach

meaning to, expand on, or change. The ability to mirror the original gesture is central to this process.

Specific Example

In the dental surgery, the dentist used his hands to describe the sequence of a complicated medical procedure to the patient, (Figure 11). First he indicated with both hands where the gums would be lifted and the bone cut. Then with one hand he described how the upper palate would be free to move around. At this point the patient used a similar gesture and indicated both verbally and with a nod of the head that he understood. The patient's use of the same gesture made it clear what it was that he understood.

Figure 11: The dentist explains the surgery

Descriptors

Imitation, mimicry and adaptation, mutual agreement, prominent gestures and postures, establishing shared understanding.

Implications

We gain an understanding of what other people are doing through our ability to interpret their actions. If electronic devices use the language of gesture and posture in a way that people can understand and interpret then there is the possibility for a different type of communication between people and their devices.

Mirroring gestures have important implications for devices that support communication between people. Adding the ability to transmit and mirror gestures adds to the depth of the communication between the people. There is potential for couples or groups using gestural devices to easily develop a set of gestures that they agree upon that could be learned by gestural input devices. There is potential for gestural confirmation mechanisms in CSCW.

When we see a gesture, we can mirror it without paying much visual attention, because we are attuned to the positioning of our bodies through feel. This implies that gestures can be used as information input mechanisms without need for visual feedback.

6 Discussion

This research has attempted to explore how gestures are used in the context of everyday work through use of the Video Card Game method. This led us to present our findings as a number of gestural themes that might inspire design solutions.

It must first be acknowledged that these themes are overlapping. The deckhand rhythmically rolling up his sleeve was seen to fit both in the theme of "preparatory gestures" and in the theme of "gestures as placeholders" (in this case temporal placeholder).

Furthermore, the themes presented are not exhaustive. In the Video Card Game, approximately forty themes were proposed for exploration in initial discussions. Each player then chose their favourite theme leading to development of ten themes through the process of viewing videotape, annotating, sharing, discussing and gathering video cards. The themes presented are illustrative rather than exhaustive.

Some themes identified gestures that appear to fulfil some sort of function, e.g. "gestures as placeholders", "commanding gestures", "mirroring gestures". Other themes related to the character and quality of the activity in which the gestures played a part e.g. "competence and confidence". The theme "being still and thinking" identified activities that appeared to be so demanding of cognitive and visual attention that there were no gross motor movements and few fine ones. Other themes related to the character and quality of the gestures, such as "hands on but imprecise". Other themes such as "shared tools, shared workspace" pertained to particular social and environmental contexts in which the role of gestures would appear to be important.

The themes present views on gesture and are not intended as a preliminary attempt at gesture classification. There has been much work in the area of gesture classification, particularly from the perspective of semiotics. This work usually begins with the problem of spatial and temporal boundaries, considering where a gesture starts and ends, and then how to circumscribe the space that it delineates. The lexicon approach, favoured by many semioticians uses nomenclatures e.g. shrugging, nodding, waving, pointing" as a primary resource for the segmentation of the flow of movements. This work has led to dictionaries of gesture and has been applied in the computer representation of avatars. These approaches have generally removed the gesture from its context and taken an approach of segmenting and reducing the gesture to constituent parts. While the results of the semiotic approach to gestural research are demonstrably useful, from an epistemological point of view many questions remain unanswered. (See Bouissac 2002 for a critique). For example, rather than assuming that gestures have boundaries and asking where they start and stop, Bouissac contends one must ask whether gestures have boundaries at all and if so what is their nature? The work described in this paper does not attempt to address such epistemological questions, instead having its own pragmatic perspective. We have taken a broader view of gestures in the context of activities, incorporating physical, environmental and social context, with a view to understanding how gestures can be used in design. While the epistemological questions are of interest, they are beyond the scope of this inquiry and a subject of future work.

It is important to consider the nature of our findings in relation to and as a consequence of the Video Card Game methodology.

We shot, edited and viewed video material gathered in domains of which we had only layman's knowledge or experience. Thus, our perceptions and interpretations are preconditioned by our own wide range of lived experiences. The video data simply renders observable specific activity in a particular domain that occurred in a particular timeframe at a particular place.

Our interpretations are thus a blend of our own life experiences and our videotape observations, developed through discussions. This raises the question of what causes a group of observers to pick out certain gestures from the complexity of human activity?

The Video Card Game allows us to explore gestural instances using a "human sized" scale of analysis and makes no attempt to physically increase this threshold using artificial devices such as frame-by-frame analysis. By videotaping in several different environments for approximately 30 minutes, and selecting clips of approximately 90 seconds for further scrutiny, we are somewhat predisposed to identifying gestures that stand out from the general pace of action. If instead, we spent a full day observing interactions with the ferry timetable, we might notice for example that half of the people look without pointing and that this is a more typical interaction than the place holding activity that we identified.

In order to identify gestures that are typical in a domain, rather than those that stand out, it would be necessary to follow up with more detailed observational studies such as ethnography or interaction analysis in the tradition of ethnomethodology, or protocol analysis.

As analysers it must be recognized that we possess tacit cultural knowledge that biases our observations. We also have some "primal mechanisms that appear to have evolved for the vital management of [gestural] information" which are bound to bias our observations, (Bouissac 2002).

Further, from the analyser point of view, the distinction between gestural events that we can and cannot notice and articulate to each other using the video-card game is complicated by virtue of our conscious processing thresholds and our natural spoken language. The semantic categorisations of natural language sometimes presented problems in sharing themes. We often found ourselves sketching alternative easily accessible mini scenarios for each other as a way to express what we meant by a theme. Often we enriched these little instances by using gestures to express what we meant. That is, as investigators we impact upon the production of the data, since direct investisations of gestures create gestural interactions of their own.

The many issues described above pose epistemological difficulties for gestural research, in particular attempts to understand and classify gestures definitively. However, seen from the point of view of developing better contextual understanding and inspiration to guide design, the Video Card Game methodology was very effective in developing a preliminary and shared understanding of the ways that gestures are used in a number of contexts, allowing comparisons across contexts. Designs must respond to particular needs, opportunities and problems and the Video Card Game was, in our experience, an effective way to elucidate some of these needs, opportunities and problems with respect to gestural interaction.

Our discussion now turns from the analysis of gestures to the use of our observations in design.

It is important to note that while design ideas arise from examining themes, objects and environments, there is no simple route from analysis of activity to design of devices. At most, analysis of activity reveals design problems, opportunities and possibilities, but it does not prescribe design solutions. Design solutions arise from synthesis rather than analysis activity.

Design ideas arose from identifying particular aspects of environments, objects and themes which rendered hand and body use either clumsy or elegant. For example, sterility constraints forced the dentist to dictate to his assistant who typed patient information into the patient record using a keyboard because his hands were sterile. This clumsy interaction revealed opportunities for obviating sterility constraints through use of gestural commands to navigate and modify the patient record. By coupling say seven distinct gestural commands in a gestural input device such as a wireless ring and watch, in concert with a microphone/speech chip that recognises only dental terms, it would be possible to design a very lightweight naturalistic means of modifying the patient record that did not render the dentist feeling excessively instrumented or concerned about privacy. Alternatively the dentist could navigate and write notes using a gestural pen device that could be sterilised just like other dental instruments.

The elegant way in which people use gestures as placeholders in complex information spaces and the elegant way in which people adjust their body postures and gestures to effectively shares workspaces, suggests that we should endeavour to take advantage of these human abilities in interface design, in order to design computer systems that can be used collaboratively.

We see the Video Card Game as a useful, engaging, participatory method for exploring contexts for design intervention. In particular the method is helpful for generating a wide range of views of one or more domains quickly, grounded in specific videotape data. It is possible to engage participants from the domain in the Video Card Game, enabling the participants and designers/video makers to construct understandings and designs together. The Video Card Game thus plays a unique and useful role in design work. Clearly it must be used in conjunction with other methods. Based on the Video Card Game study, we selected the domain of dentistry to examine the possibilities for gestural control in sterile environments. Our work is proceeding with observational studies grounded in ethnomethodology, a second video card game, prototype development, scenario development, and informance design among other methods that we are

evolving in the course of the design project to meet the design questions that arise.

7 Acknowledgements

We would like to thank all of the game players and the videotaped participants that made this research possible. This work was funded by ARC Discovery Grant DP0210470

8 References

ALEXANDER, C., ISHIKAWA, S., SILVERSTEIN, M., JACOBSON, M., FIKSDAHL-KING, I. and ANGEL, S. (1977) *A pattern language: towns, buildings, construction,* New York, Oxford University Press.

BATES, E., (1994) Modularity, domain specificity and the development of language. In *Evolution and Neurology of Language. Discussions in Neuroscience* **10**(1-2) 136-149.

BOUISSAC, P. (accessed 2002) Describing gestures: boundaries, scales and perspectives, *Virtual Symposium on the Multimodality of Human Communication: Theories, Problems and Applications,* University of Toronto, May 3-5 2002.

BRERETON, M. (2001) Drawing Lessons in the Design of Tangible Media from a Study of Interactions with Mechanical Products, *Proceedings of the Australian User Interface Conference*, pp 92-100, 24th Australasian Computer Science Conference ACSC 2001 Held: Gold Coast, Australia, Jan 2001.

BRERETON, M.F. and MCGARRY, B. (2000) An Observational Study of How Objects Support Engineering Design Thinking and Communication: Implications for the design of tangible media. *Proceedings of the ACM SIGCHI Conference on Computer Human Interaction*, CHI 2000, pp 217-224, The Hague April 1-6th 2000. ACM Press.

BUUR, J., BINDER, T. and BRANDT, E. (2000), Taking Video beyond Hard Data in User Centered Design. *Proceedings of the Participatory Design Conference,* Nov. 00. New York

BUUR, J. and SOENDERGAARD, A. (2000) Video Card Game: An augmented environment for User Centred Design discussions. In *Designing Augmented Reality Systems Conference,* ACM Press, Elsinore

CHI'95 gesture Workshop: http://wex.www.media.mit.edu/people/wex/CHI95-workshop-writeup.html

CLARK, A. (1997) *Being There: Putting Brain, Body and World Together Again,* Cambridge, Mass, MIT Press

COUTAZ, D.J. and CROWLEY, J.L., (1995) Interpreting Human Gesture with Computer Vision. *Position paper for the workshop "Gesture at the User Interface", CHI'95,*

CRABTREE, A., HEMMINGS, T. and RODDEN, T. (2002) Pattern-based support for interactive design in domestic settings. *Proc. Designing Interactive Systems Conference,* 265-27, ACM Press.

ERICKSON, T (2000) Lingua Francas for design: sacred places and pattern languages, In *Designing interactive systems* Conference. New York

JAKOBSON, R (1990) *On Language*. eds WAUGH L.R and MONVILLE-BURSTON M. Cambridge, Mass.: Harvard University Press.

JORDAN, B. and HENDERSON, A. (1994). Interaction analysis: foundations and practice. The journal of the Learning Sciences **4** (1)

KEATES S, and PERRICOS C (1996) Gesture as a means of computer access. *Communication Matters*. **10**(1)17-19

KRAUSS, R.M., CHEN, Y. and CHAWLA, P. (1996) Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In *Advances in experimental social psychology* 389-450. M. Zanna (ed.) Tampa: Academic Press.

KRAUSS, R.M., DUSHAY, R., CHEN, Y., and RAUSCHER, F. (1995). The communicative value of conversational hand gestures. Journal *of Experimental Social Psychology* **31**, 533-552.

KRAUSS, R.M., MORREL-SAMUELS, P. and COLASANTE, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology* **61**, 743-754.

MCLUHAN, M (1964) *Understanding Media: The Extensions of Man*. New York: McGraw-Hill

MCNEILL, D. (1992) *Hand and Mind: What Gestures Reveal about Thought*." Chicago: University of Chicago Press

MCNEILL, D. (1985) So you think gestures are nonverbal? *Psychological Review* **92**:350-371.

NESPOULOUS, J.L. PERRON, P and LECOURS A.R., (1986) The Biological Foundations of Gestures: Motor and Semiotic Aspects. Lawrence Erlbaum Associates, Hillsdale, MJ,

PELACHAUD, C. and POGGI, I. (1997) Multimodal Embodied Agents ACM 0-89791-88-6/97/05

RAUSCHER, F.H., KRAUSS, R.M., and CHEN, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science*, **7** (4), 226-231.

RIZZOLATTI, G. LEONARDO F and VITTORIO G (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience **2*** : 661-670

SPINNEY, L., (2000) BodyTalk *New Scientist* **166**(2233) 30-31

STRAUSS, A. L., CORBIN, J. (1998) *Basics of qualitative research: techniques and procedures for developing grounded theory.* Thousand Oaks, Sage Publications

View publication stats