
PsychNology Journal
Vol 5 n.2
ISSN 1720-7525
5.2

PSYCHNOLOGY JOURNAL
The Other Side of Technology

PSYCHNOLOGY JOURNAL, PNJ
Published online since Summer 2002
Web Site: http://www.psychnology.org
Submissions: articles@psychnology.org

TABLE OF CONTENTS

PsychNology Journal, 2007
Volume 5, Number 2, pp. 105 – 132

Extended Television: A Study of How Investigations of Use Can Inform Design Processes in Nursing Homes

Peter Abdelmassih Waller ♠♣

♣ Rehabilitation Engineering Research, Department of Design Sciences, Lund University, Sweden

ABSTRACT

This paper describes the shortcomings in the support that replaces the lost distributed cognition in older people who move to nursing homes and how artifacts can improve this by functioning as distributed cognition. The 30 older persons in the study were men and women of different backgrounds and ages (between 60 and 100 years), all of whom had some kind of functional limitation. The observations and analysis were carried out as a part of the iterative design phase of TV functions for the older people, and the analysis was based on distributed cognition theory and the FACE conceptual design tool. Poorly designed artifacts resulted in the older person's loss of control, and hindered the creation of distributed cognition. However, these aspects improved in the older persons' TV watching when individually adapted assistive technology was used.

Keywords: *Distributed cognition, older people, assistive devices, design, television.*

Paper received 29/03/2007; received in revised format 20/07/2007; accepted 03/08/2007.

1. Introduction

Distributed cognition describes how human acts or artifacts can mediate knowledge and memory (Hutchins, 1995a, 1995b; Norman, 1993), without people necessarily being aware of it (Norman, 1993, p. 143). Systems with distributed cognition can not be reduced to cognitive properties of individuals (Hutchins, 1995b, p. 355). Empirical investigations of the role of technology in collaboration and in distributed cognition in nursing homes are scarce. This is especially true for investigations with a focus on the older individuals.

♠ Corresponding Author:
Peter A. Waller
Certec, LTH,
P.O. Box 118, 221 00 Lund, Sweden.
E-mail: peter.waller@certec.lth.se

This paper discusses the memory aspects of distributed cognition. This means other aspects of distributed cognition, such as decision making, inference, reasoning and learning (Hutchins, 2001), are not analyzed. The term *distributed memory* is used to denote the segment of distributed cognition that concerns memory. Distributed memory processes are disrupted when older people move into nursing homes and need to be rebuilt there. The relocation results in new routines, changes in relations with relatives, new people to interact with and loss of old artifacts. The old artifacts that are kept in the one-room apartment in the nursing home can originate from several different rooms in the former home and from different places where the older person used to stay before the relocation, for instance, from an apartment or a summer cottage. New artifacts are introduced and new arrangements of these and old artifacts are made. Furthermore, this happens at a time in life where memory problems can arise without a disruption in the distributed memory. In the relocation, the older person's self-image can change from one of self reliance to one of becoming a burden (Svidén, Wikstrom, & Hjortsjö-Norberg, 2002), and older people with dementia can become confused and experience discomfort (Son, Therrien, & Whall, 2002). However, there are also older people who experience relief and security after the relocation to a nursing home (Lee, Woo, & Mackenzie, 2002).

The observations in this study were made as a part of the design of a suite of TV functions in a nursing home that are referred to as *Extended Television* (see section 4 for a description of the functionality). The overall research question is how the conceptual design tool FACE [Function, Attitude, Control, Enabling] (Anderberg, 2006) and distributed cognition analysis of the usage can contribute to the design process. Furthermore, the purpose of the article is to describe the shortcomings in the support that replaces the lost distributed memory of older people in nursing homes, and to understand how artifacts can strengthen the older person's actions and control, when they are a part of his or her distributed memory (meaning that they act as representations in distributed memory processes). The study also seeks to understand how this knowledge can be used in the design process. The focus is on the older persons, although care workers and relatives are included in the analysis.

The use of extended television is analyzed using distributed memory and FACE (Anderberg, 2006), which is a conceptual design tool to analyze perceived attitudes, the individuals' control, and enabling of the analyzed function, such as watching TV. The main contribution of this article is in the qualitative empirical study of older people's use of extended television in a nursing home, and in the use of a combined FACE and distributed memory analysis in this context.

The design phase (which involved older persons, relatives and care workers) and usage of extended TV is described by Abdelmassih Waller, Östlund, and Jönsson (2007).

It is important to acknowledge that older people watch more television than other age groups, that the activity seems to increase over the adult life span and that this increase has been constant through time (Nordicom, 2006; Robinson & Skill, 1995). Television is poorly adapted to older people's social inclusion in their immediate surroundings, but is used by older people to support their own reflections (Östlund, 1995, p. 137, 147). These factors were an early inspiration in the design phase of extended television.

Some of the prerequisites of designing for older people in a nursing home are described in section 2, and in section 3 FACE and distributed memory are discussed in greater detail. Section 4 describes the settings and the participants, while section 5 is devoted to presenting the method used. Section 6 describes the use of extended television, as well as an analysis thereof. A concluding discussion is presented in section 7 and conclusions in section 8.

2. Design for Older People

In this section, general considerations on how to design for older people are discussed first, then how photos can strengthen communication with older people, and finally how older people's memory changes with age.

Familiarity with visible parts of technology increases the probability for acceptance of a "new" artifact by older people (Östlund, 1995, pp. 234-235). In addition to this, poor health, unsatisfactory contact with others, and high education can contribute to older people's acceptance of new technology that enhances the quality of life (Zimmer & Chapell, 1999). The need to minimize physical and mental efforts probably influences "their choice of technology and their willingness to learn." Furthermore, older people "tend to accept technology that saves energy and makes life easier" (Östlund, 2005, p. 38), and it is important that new devices are constructed so that the older person is in control (Anderberg, 2006; Norman, 1999, p. 174; Orpwood, Bjørneby, Hagen, Mäki, Faulkner, & Topo, 2004). There is also a design recommendation to change as little as possible, and to design for the older person's own actions (Jönsson, 2003, p. 11). This can mean that the older person would like to keep a familiar artifact, but also that there is a familiar function that the older person wishes to keep. Furthermore, to design for the older person's own actions does not necessarily mean that the older person should be physically active; it could also be a design facilitating the older person's reflections on the past.

For people with dementia it can be advantageous if new features are included imperceptibly, so that the new device (that looks like equipment that existed in their home) operates just like the ones they are used to (Orpwood et al., 2004; Orpwood, Gibbs, Adlam, Faulkner, & Meegahawatte, 2005). Such devices are not necessarily a part of compensatory memory aids that are consciously used by older persons, such as those discussed by Caprani, Greaney, and Porter (2006). The importance of adapting the artifact to past experiences and utilizing external cues as well as environmental ones is also stressed by Gamberini, Alcaniz, Barresi, Fabregat, Ibanez, and Prontu (2006).

Human support can be valuable, among other things, for its assistance in adapting to a changing environment; however, the supported person's loss of control is always a problem. Well-adapted technology, on the other hand, can be an extension of the individual's own wishes (Anderberg, 2006). However, a study by Drageset (2004) indicates that help from care workers in performing activities of daily living can counteract feelings of loneliness.

In order for care workers to provide quality care, their actions should preferably be based on a mental image of the older person that includes the time before living at a nursing home. One possible consequence of lack of knowledge is a spiral of increasing dissociation of the older person and the care worker. In this situation the care worker can feel stressed, and come to believe that he or she carries the entire responsibility (Rahm Hallberg, 2002, p. 48). Beck-Friis states that people with dementia need a trail blazed down to their memory, such as a scent, a melody or a photo. She states that this is important to consider in the care of people with dementia, and that support is to be given in order to *"help to bring forth the mental images that are still there/that still exist."* (Beck-Friis, 2000, p. 41)

2.1 Using Photos for Communication

Technical support in the form of digital and printed photos has proven to be a valuable communication and memory tool. An example of this is the extensive use of personal photos by adults with developmental disabilities as described by Plato and Jönsson (2001). They report that photos

> "support and confirm memory as well as working as a written language. In addition, they are a source of inspiration for conversation, that is, they work as a spoken language or support to such as a written language" (Plato & Jönsson, 2001, p. 7).

The personal photos were much better suited for these purposes than non-personal, symbolic, photos (Jönsson, Philipson, & Svensk, 1998). A photo can be experienced as personal if the concerned individuals have memories associated with it. In this way, photos can act as representations of memories to individuals in processes with distributed memory. A photo that only one person has memories tied to can be used for communication; however, a photo works better as a conversation support if both persons have memories tied to it. Furthermore, memory books with sentences described by either personal photos or general sketches have been found to augment the communication between care workers and older people with dementia (Allen-Burge, Burgio, Bourgeois, Sims, & Nunnikhoven, 2001; Bourgeois, Dijkstra, Burgio, & Allen-Burge, 2001). Multimedia solutions for older people with dementia using photos and music have also shown positive results (Cohene, Baecker, Marziali, & Mindy, 2007; Topo et al., 2004).

2.2 Memory and Familiarity

Human implicit memory categorizes the effects of prior experience without conscious recollection thereof. This means that unconscious procedural skills such as bicycling and combing hair are included in implicit memory. Furthermore, the usage of implicit memory can give a feeling of familiarity (Son et al., 2002). The performance of explicit memory tasks, such as conscious recall, often becomes poorer with age, but implicit memory is often unaffected (Fisher, 1998, and as summarized in Caprani et al., 2006) or is only slightly affected (Woodruff-Pak & Lemieux, 2001). This deficit in explicit memory is observable among older people with dementia, and it is recommended that new interventions should use aspects of prior familiar environments including objects and pictures "to maximize functional ability in elders with dementia" (Son et al., 2002, p. 266).

Investigation of how older people with dementia, who have a working implicit memory but poor explicit memory, use technology could shed light on design issues for older people in general as well as older people with dementia. Their poor explicit memory means that they have difficulties with conscious learning. Consequently, artifacts that the older person with dementia does not use but wants to use become indicators of what is difficult to learn. People with a working explicit memory can consciously compensate for poorly designed artifacts, while older people with dementia are unable to do so. Strengths in the implicit memory of people with dementia can be used to design artifacts that give a feeling of familiarity and may also be appreciated by a wider range of older people.

3. Theory

This section provides a theoretical background for the distributed cognition and FACE analysis. Distributed cognition is described first, and discussed from the viewpoint of activity theory and situated action. Thereafter, distributed cognition is discussed from the perspective of aging, and finally FACE is described.

3.1 Distributed Cognition

Distributed cognition is one of the conceptual frameworks that can be used to analyze the use of extended television in its context. Examples of other possible frameworks are activity theory and situated action (Johansson & Gärdenfors 2005; Kaptelinin et al., 2003). These frameworks all focus on both users and context. As stated earlier, this paper uses the memory aspects of the distributed cognition framework to analyze the use of extended television. This non-reductionist analysis treats complete processes, such as watching TV. The analysis shows how memory in a process is represented in the different steps needed to carry out the process. For instance, the idea of watching a certain TV program can be presented orally by a relative, and later the presence of the TV acts as a reminder to watch the program. In this example, the idea of watching TV was first represented by the relative and later by the TV. This transformation of different representations of memory is also treated by Hutchins (1995a).

A disrupted distributed memory can result in the loss of representations needed to carry out one or several of the steps in a process, thus making the activity difficult or impossible to perform. Consequently, a distributed memory analysis can find the missing representations needed for the older people and discuss how to assure the existence of such representations.

The theory of distributed cognition (Hutchins, 1995b) is symmetric, meaning that both artifacts and humans can be thought of as cognitive representations. This is different from activity theory, which distinguishes between the user (subject) and the artifact (object) (Nardi, 1996, p. 73). Distributed cognition can be used to analyze processes where the user plays a significant role without necessarily being aware of it (Norman, 1993, p. 143), while activity theory focuses on intentional actions. In fact, activity theory “seeks to understand the unity of consciousness and activity” (Kaptelinin & Nardi, 2006, p. 8). In the design of extended television, it has been important to analyze situations where people’s intent has been unclear due to difficulties of interviewing people with dementia. Furthermore, it has been necessary to use a framework that highlights the role of artifacts. This is a frequently used and acknowledged property of distributed cognition (Nardi, 1996, p. 85). However, there is a risk that human intentionality and creativity can be neglected with a symmetric analysis like distributed cognition (Kaptelinin & Nardi, 2006, p. 241).

Suchman points out that situated action stresses the uniqueness of human action, compared to machine operation. Furthermore, she thinks that in order to make a situated analysis of human action, it is necessary to consider the “unique, unrepresented circumstances in which action in every instance and invariably occurs” (Suchman, 1987, p. 189). This means that human improvisation is emphasized (Nardi, 1996, p. 85).

The design in this project draws on the user’s earlier experience with television and on familiar patterns of actions concerning TV viewing. It seeks to limit the learning needed to use extended television, especially due to possible deficits in the explicit memory among older people. This means that the designed artifacts should well suit the implicit memory in which routine actions are stored. This makes distributed cognition a suitable framework of analysis, with consideration taken to human preferences.

3.2 Distributed Cognition and Aging

Distributed cognition, as mentioned in the introduction, can not be reduced to cognitive properties of individuals. This presents opportunities to support older people’s cognitive processes through adapting the environment to the mental processes of the individual. Such a procedure has been proposed by Palen and Aaløkke (2006). They made a study of how older people in a Danish home nursing setting remembered their intake of medications by how they were spatially arranged.

This arrangement was in accordance with the routines they normally carried out. For instance, one older person placed midnight medications on the nightstand. Furthermore, it was found that the spatial arrangement of medications was used in the communication between care workers and older people. Recommendations were made that a computer supported medication management system should take its starting point in the personalized, spatially distributed systems that the older person is already using.

In this way, the cognitive processes of the individual and the uniqueness of the human actors are taken into consideration by using a distributed cognition approach.

Older people's social networks and routines may have changed before the relocation due to fewer relatives or physical constraints (Östlund, 2005, p. 29). This means that the distributed cognition can have started to change before the relocation to the nursing home. As mentioned in the introduction, the relocation to a nursing home can cause a serious disruption in distributed memory. Furthermore, it is reported in medical literature that the relocation to a nursing home can cause much stress for the older person (as summarized in Manion and Rantz, 1995). Feelings of loneliness, sadness and crying can occur after relocation. However, this can be decreased if the older persons' decision making and contact with relatives is supported, and also if there are familiar artifacts in the apartment (Wilson, 1997). The need for contact with relatives and for familiar artifacts emphasizes the importance of designing in a way that retains the distributed memory that the older person had before the relocation. In other words, to change as little as possible.

3.3 Function, Attitude, Control, Enabling (FACE)

FACE is a conceptual design tool for analyzing individual functions and functional support, including support from both technology and other persons. According to Peter Anderberg, who developed the tool, functions are the activities we carry out in everyday living such as reading, telephoning, going to work or school, and meeting friends and co-workers. The chosen function is analyzed in the framework of the following three factors: attitude, control and enabling. "Attitude" means the social response to the function; how the user himself as well as others perceive it in the setting where it is used. "Control" involves how much the user, who is the owner of the function, has the power and right to define it and carry it out. "Enabling" deals with how the function support (the technological solution or artifact) is constructed and implemented. FACE can be used to compare the suitability of different function supports (Anderberg, 2005; Anderberg, 2006).

4. Setting and Participants

This section describes the setting, influx of older people, care workers, and the selection of participants to the project. In addition to this the older persons who will exemplify the use of extended television is described.

In a newly built nursing home for older people in the south of Sweden there are 53 apartments, common day rooms and TV rooms. The 53 apartments are distributed over three floors and five divisions, where the first floor is a short-term division that is excluded from the analysis and further description of the settings due to its late inauguration. In each of the apartments there was a 32" wall-mounted LCD TV screen. Both the number of older residents that were involved and the number of employed care workers increased during the period of observations, since the apartments were gradually becoming occupied. Floors two and three were filled to 40% capacity by May 2006, 80% at the end of August 2006 and 90% at the end of December 2006. The older persons that moved into the nursing home did so because they could not manage on their own, even if they had home nursing. A regional ethical vetting board scrutinized the project before the older people moved into the nursing home. The board found no obstacles to proceeding with the research project.

The unpredictable influx of older people and final adjustments of the nursing home made it impossible, at the outset, to find a static group of older persons who were using and evaluating the extended television. Instead, I included any of the older people who became acquainted with extended television, either because they themselves expressed a wish to use it, or because relatives or care workers expressed such a wish. Also the few older people who deliberately refrained from using extended television were included in the study. In total, 30 older persons were included. They were men and women of different backgrounds (fisherman, singer, engineer, athlete, care worker, etc.) and ages (between 60 and 100 years). They all had some kind of functional limitation due to, for instance, different kinds of dementia, physical disabilities caused by rheumatism or cancer, and stroke-related conditions such as aphasia or memory loss. Actually, the majority of the people living at the nursing home had some kind of memory loss.

All of the permanently employed care workers had traditional health care training, and several of them also had other training that was especially valuable at the nursing home, such as computer skills, painting, pottery, baking, hairdressing, etc. There are care workers present at the nursing home, ready to help the older persons, around the clock. Each older resident has a contact person among the care workers who has additional responsibility to assure proper care of the older person.

The apartment TVs were delivered with a remote control, denoted in this paper as "original remote control", which had approximately 40 grey buttons, several with multiple functions, see Figure 1a. In some apartments this remote control was exchanged for a universal remote control with ten buttons, see Figure 1b.

There is also an internal TV channel at the nursing home, showing movies and photos that the care workers have taken using digital cameras. The photos were taken of events at the nursing home such as the Midsummer Eve dinner, Christmas, bingo, visits to the nursing home, etc. A code of ethics was adopted. This meant that no photos were taken inside apartments or in situations that could be experienced as embarrassing by the older people. Two media centers were present in the nursing home in one day room and in one TV room. It was possible to view photo CDs, movies from the open archive at Swedish Television, DVDs, etc., using the media centers. They consisted of PCs connected to the Internet and the TV.

Of the older people who used the internal TV channel in their own apartment (at least 8 older people), Anna was chosen to exemplify the use. Anna was 90 years old and had memory deficits due to stroke and a brain tumor. However she did not have dementia. Her reaction to the photos in the internal TV channel is representative for most of the older people who watched the internal TV channel (including those who only watched it in the day rooms). Of the four older persons who had a universal remote control in their apartment, Marie was chosen to exemplify the use. However, her use can not be seen as representative since she was the only one with dementia. Marie was 92 years old and had a diagnosis of vascular dementia.

(a) (b)

Figure 1. The “original remote control” (to the left) and the universal remote control (to the right).

One older person, Berit, had a personal TV photo album installed in her TV. She had significant short-term memory loss due to mid-stage Alzheimer’s disease, and was almost 90 years old. The TV photo album was filled with scanned photos from her old photo albums and controlled by a special remote control in the shape of a photo frame, see Figure 2 in section 6.2.2. The decision to let Berit try a personal TV photo album was based on her interest in photo albums, and her relatives’ strong involvement in the design process.

In general the relatives included in this study were selected on the basis of two criteria. First, their older relative must participate in the usage of extended television, and secondly, they should have a great deal of contact with the nursing home.

5. Method

This section describes data gathering techniques, my presence at the nursing home and the treatment of data. The section also discusses measures to increase the validity of the results.

This project was action oriented, since the research persons' actions were analyzed and since I, as a researcher, also acted. Hence, and because the project sought to arrive at improvements, the project can be called action research (Robson, 2002, pp. 215-216). The action research approach was combined with participatory observation (Ely, Friedman, Gardner, & McCormack Steinmetz, 1993, p. 49), which is in line with Hutchins' suggestion of being situated through the use of cognitive ethnography to make a distributed cognition analysis (Hutchins, 1995b, p. 371). I was present at the nursing home about 15 hours a week between June 2006 and January 2007. However, during installation of technology and during observations of Berit's use of her photo album, my presence was considerably higher, up to 30 hours a week.

During my presence at the nursing home, I participated in the usage of all the functions of the extended television, routinely visited all the day rooms to listen to complaints or comments, conducted interviews with older people, relatives and care workers. I did not prepare questionnaires, but knew which areas of knowledge I was seeking to understand and concentrated of those during the interviews, as described in (Ely et al., 1993, p. 67). I wanted to know which functions a person used, how, why and when. In conversations with the older people, I had to take special care not to point out gaps in their memory, since this could risk upsetting them. I mostly scheduled the interviews with the care workers since they did not have time during working hours; however, interviews with older people and relatives were carried out on the spot as the opportunity came up. The care workers who were interviewed were either a contact person for an older person who used extended television, used it together with the older person, had ideas about usage or was a member of the TV team.

I also trained potential older users, relatives and four care workers in a TV team. The TV team used the functionality of extended television together with the older people, gave feedback to me and trained other care workers.

In addition to the design of extended television I helped out with everyday technical problems, which made me known and integrated in the research environment. In this role I visited more than 50 % of the older persons (mostly older people without dementia).

The design process was carried out in cycles containing data gathering, analysis, and manipulation of artifacts. I continuously wrote notes about observations and interviews. The notes were stored digitally on securely stored USB memory sticks, and never edited using an online computer, to avoid uncontrolled dissemination of the material. The notes were analyzed by posing questions like "Does she always react in this way to this photo?” or by utilizing analytical frameworks as in this article. The outcome of this analysis fuelled further changes in the design of extended television, with the aim of improving and adapting its functionality. I constantly sought after data that would be divergent from my interpretations of the data or would falsify conclusions, and I always used multiple sources of information. I also consulted research articles during the design process, in order to confirm or reject interpretations. This approach strengthens the validity of the method (Figueiredo & Cunha, 2007).

In accordance with the democratic motives of action research (Brydon-Miller, Greenwood, & Maguire, 2003), the different functions of extended television were evaluated by the older people, relatives and care workers, since they gave their continuous feedback in the project to me or to members of the TV team.

I interacted with all of the 30 older persons included in the study, but I spent significantly more time with Berit than with the other individuals. This is due to the detailed analysis needed to design the TV photo album. Hence, after Berit agreed to participate in the project, informed consents were signed by her trustee, involved relatives and care workers.

6. The Use of Extended TV

This section describes the usage and analysis of extended TV. The usage is divided into two categories: “Marie watching TV alone” and “with others in front of the TV”; the second category includes Anna’s reaction to the internal TV channel, Berit’s use of her TV photo album and the use of the media centers.

The analysis is divided into two parts, the FACE analysis and the distributed memory analysis. The FACE analysis is made separately for Marie, Berit and Anna, while the distributed memory analysis is divided into the categories of distributed memory in artifacts and together in front of the TV.

6.1 Marie Watching TV Alone

Marie is slightly over 90 years old and used to live in an apartment next to one of her daughters. She has been an avid viewer of national ice hockey competitions on TV for forty years. Before Marie moved into the nursing home, her daughter used to suggest programs for her to watch, and Marie was able to operate the TV herself. She knew that the TV was there and could come up with the idea of watching TV herself. She was told to always switch off the TV using the button on the TV set that made the red light go out. When Marie moved into the nursing home she had a diagnosis of dementia. The new TV was installed in her room with the “original remote control”, see Figure 1a. Marie could not handle this remote control, and most often forgot that there even was a TV in her apartment. Her daughter was unable to help her by telephone since the remote control was too difficult to use, and thus there was no point in suggesting a program to watch over the phone. The relatives who visited her on almost a daily basis suggested she have a universal remote control, and she was provided with one, see Figure 1b. Now it became possible for Marie’s daughter to call her to remind her of an ice hockey game on TV and guide her in using the remote control. Marie would then watch TV for hours, and she was able to change the channels herself, since she recognized the channel logotypes placed on the buttons on the universal remote control. Marie stated that she tried different buttons to find the right one. She could also turn off the TV; however, on several occasions the TV was unplugged. Marie probably did this to be sure that the red light went out. She did not remember this, however, when the TV was to be switched on again.

In one instance, a care worker came into her apartment as she was watching ice hockey. The care worker wondered if this was what she really wanted to see and the answer was positive. The care worker was surprised to find out that Marie liked to watch ice hockey. Later the care workers also understood that she liked to watch figure skating. After this, it seemed that the care workers remembered to turn on Marie’s TV if ice hockey was showing.

To sum up, after the relocation Marie could not operate the TV without assistance of care workers or relatives being physically present, and not all care workers knew about her TV watching habits. When Marie was handed the universal remote control, it became possible for her daughter to assist her in watching TV, an arrangement that had also existed before the relocation.

6.2 Together in Front of the TV

There are many different constellations of where, who and what people watch together. Three cases are described here: Anna’s reaction to the internal TV channel, Berit using her photo album, and viewing photos using a media center. The importance of the involved people having memories to share became evident in the following.

6.2.1 Anna's Reaction to the Internal TV Channel

One of the older resident's of the nursing home and the care workers told me that Anna, another resident who did not have a diagnosis of dementia, wanted to watch the internal TV channel in her apartment, and that her grandchildren where asking for this. Anna had earlier only been fond of watching SVT1 and SVT2 (the two oldest terrestrial broadcast channels in Sweden). I tuned Anna's TV to the internal TV channel and she was very happy to see the photos from her floor of the nursing home. She recognized many people, related things about them, and about what had happened as the photos where taken. She said, *"It is the Christmas celebrations,"* and, *"Look at that handsome man!"* She did not speak much if she did not recognize the event in the photo, and she thought she looked old in some photos. As I was about to leave, she raised her voice and pointed at the screen saying, *"Look!"* Shortly thereafter a care worker entered the apartment, and asked about the names of the people in the photos. The care worker corrected Anna when she said the wrong name of a neighbor shown on a photo. Anna mentioned that it is *"So good with these photos."* It was time for afternoon tea and the TV was turned off. A few days after this event, Anna did not remember that she had the internal TV channel, but she was happy to watch it when it was turned on.

To sum up, Anna's interest in viewing photos from the nursing home was not noted by the care workers until they watched the photos on the internal TV channel together with her. The photos made it possible for Anna to express her thoughts and for the care workers to communicate with her in a way not possible prior to the existence of the internal TV channel.

6.2.2 Berit's Photo Album

Berit had short-term memory loss and Alzheimer's disease, and was almost 90 years old. She did not take the initiative to look at her old photo albums, even if they were left visible. Berit was given and trained in using a universal remote control to watch TV programs, but she always had to be reminded to push the buttons on the remote control. The selection of photos for the TV photo album was made carefully, and required that Berit look through all her photos and that I learn the pattern of which photos she recognized. Berit could browse through the TV photo album if she were handed the photo frame, see Figure 2.

Figure 2. The photo frame remote control used with Berit's personal TV photo album.

However, there was no evidence that she did this alone or took the initiative to do so. The contact person did not have much time to sit down with Berit, but turned on the photo album, leaving a picture visible on the screen during the day. On the occasions when she sat down with Berit to discuss the contents of the photos, Berit mostly wanted to hand over the control to her. However, Berit became more of an individual to the contact person as she saw glimpses of her life. As Berit viewed the photos in my presence, she wanted to browse for herself, imitating how I did it. The relatives found that Berit became happy as she browsed among the photos. They noticed a considerable positive change in Berit's health and memory, a change that the contact person did not recognize. Berit reacted positively several times to a totally black image that was included in the TV photo album. She interpreted this as the end, which made it possible for her to interrupt the photo album activity. This could otherwise be difficult, as she sometimes forgot which pictures she had already seen.

Berit's understanding of which decade it is, or where she is, depends on where in the nursing home she is, or with whom she is interacting. As she sits in her room with me, she most often thinks that she still lives with her deceased husband, while with her relatives she is more aware of the present situation. Her understanding of where she is, is probably also affected by the photos she views.

To sum up, when Berit came to the nursing home the conversations were often guided by care workers or relatives, but Berit could guide the conversations with her relatives when the TV photo album was installed. Furthermore, Berit became more of an individual to the contact person, meaning that more of Berit's preferences and capabilities became visible to the contact person.

6.2.3 Viewing Photos Using the Media Centers

At first, the author was the only one who used the media centre to show photos taken inside the nursing home and from the surroundings. The care workers took on the task enthusiastically when they received written instructions on how to do so. They reported that the older people watching reacted positively.

Months later, one of the older persons with dementia still pointed at the media center in the day room when she saw me and said, *"When will you play?"*

On several occasions, relatives brought CDs with family photos, which they watched together with their older relative, either in the day room or in the TV room. This seemed to be a compensation for not having a personal media center in the older relative's room. The comments on photos by relatives and older people were interrupted by the automatic change of photos, and both groups proposed a manual browsing system.

6.3 FACE Analysis

The care workers support the older residents in many of their daily routines, resulting in their having less control than when they lived independently. The following sections analyze the usage of extended TV by means of FACE. The care workers and older residents both desire that the older residents have more control, and, to a certain extent, *extended TV* became a means of transferring some of this control. The media centers were not evaluated. It was clear, however, that the older person should have more control over the photo browsing function. First Marie's TV watching is analyzed, then the usage of Berit's photo album, and finally Anna's reaction to the internal TV channel.

6.3.1 Marie Watching TV Alone

The function to be evaluated is Marie's TV watching. She watched TV by herself before moving into the nursing home, and has used the universal remote control and watched TV for hours in her apartment at the nursing home. Hence, it is assumed that she is interested in watching TV.

Before Marie used the universal remote control, it was necessary for relatives or care workers to remind her to watch TV, operate the remote control, choose channels and switch the set off. The enabling aspect of this was low, and Marie's control was limited to reacting to others' initiative. The "original remote control" strengthened the attitude that older people need help and cannot be in control, and her relatives wanted Marie to have a universal remote control.

When Marie was given the universal remote control, she could change TV channels and turn off the TV. The enabling and Marie's control were higher than with the "original remote control"; however, she still relied on others to remind her to watch TV. Marie's attitude is difficult to assess, but her description of her trial and error handling of the universal remote control was made without seeing it as a problem. The care worker who understood Marie's interest in ice hockey was positively surprised, and the relatives thought the new remote control was a great improvement. The universal remote control strengthened the attitude of Marie as an individual with her own will.

To sum up, before Marie had a universal remote control the assistance given by care workers and relatives strengthened the conception that older people are helpless. However, the control Marie gained with the universal remote control proved this wrong.

6.3.2 Berit's Photo Album

The function to be evaluated is watching the TV photo album. Berit's interest in old photo albums and her comments while viewing her photos on the TV show that she is interested in having this function.

The technology did enable her to change photos, but she still needed to be reminded about the possibility of watching them. This was also the case with the old photo albums. She had control over what photo to watch, and it gave her influence over the discussions with her relatives. The black photo empowered Berit to end the viewing session if she was tired. Her relatives see the TV photo album as an improvement, and it mediated the message that Berit is an individual with a living history who can contribute constructively to a conversation.

On the other hand, Berit's use of the universal remote control to watch TV programs did not become enabling, since she always needed someone else to tell her to push the buttons. This made her control of the TV low, and she did not really accept the remote control as hers. This conveyed a message of a person being in need of assistance to make a decision.

To sum up, the universal remote control did not drastically alter the situation for Berit; her TV watching was still in need of human assistance in most aspects. This is in contrast to her use of the TV photo album. The difference between these two outcomes stresses the importance of individually designed artifacts.

6.3.3 Anna's Reaction to the Internal TV Channel

The function to be evaluated is talking about memories by using the internal TV channel. Anna's positive reactions to the photos and the reminders from her neighbors indicate that Anna desires this function.

The presence of photos with content to talk about enabled Anna to discuss her memories. Her control was low, since it was the older persons at the nursing home, care workers and relatives who reminded her to watch the internal TV channel. However, she could operate the remote control by herself. Anna was very positive about the internal TV channel, and a care worker used it to train her memory. Anna's need of reminders conveys a message of being in need of assistance; however Anna's reaction to the images shows that she has knowledge to share, and enjoys talking about her memories.

To sum up, it is clear that the presence of the internal TV channel enabled memory training that could not have taken place without it, and that care workers took the opportunity to use it.

6.4 Distributed Memory Analysis

The distributed memory analysis is divided into two parts. First the artifacts are analyzed and then the processes taking placing while watching together with Anna and Berit are analyzed.

6.4.1 Distributed Memory in Artifacts

Both Marie and Berit had memory loss, and actually both of them had problems coming up with the idea of watching TV programs. However, both of them could get the idea of turning off the TV. The activity of watching TV probably reminded them of this, but without this activity it is not certain that they would remember that the TV was there. If they were to turn the TV on or off, they occasionally looked for a button on the TV, as on their old TV. This button was missing, and the red light that came on when the TV was switched off was confusing, which resulted in their unplugging the TV set.

For Marie, the universal remote control worked well in several processes with distributed memory. Her relatives assisted her by telephone (thus enabling an arrangement of assistance that existed before the relocation); the care workers understood her interests; and she could operate the TV by understanding the logotypes on the remote control. For Berit, the universal remote control did not represent any memories that supported her TV watching. However, the black photo represented a memory since it gave her the information that she had watched all the photos and that the album had come to its end.

6.4.2 Together in Front of the TV

The photos in the internal TV channel, Berit's personal photo album, and the family photos shown using the media centers all made situations arise where the older person could contribute to the interaction. However, the interaction seemed to be most lively from all involved parties if everyone had shared memories concerning the photos. Otherwise, it was more of a situation where one party was telling and the other was asking questions. Older people and their relatives had shared memories or associations with the family photos shown at the media centers, in photos from the internal TV channel in which they had been present, and in Berit's personal photo album. The older person and the care workers had shared memories of the photos shown in the internal TV channel.

This means that the interaction was most lively if the memory was distributed between the older person, the photos and the relative or care worker. However, the relatives also expressed an interest in other photos in the internal TV channel than those they had experienced.

One of the care workers described the use of the internal TV channel as "an aid for gaining an understanding of what their [the older people] view is, and of what they apprehend." This care worker, who used the internal TV channel to ask the older people without dementia about its contents, also mentioned that it provided memory training for the older people. Her use of the photos in this way is possible because she knew what had happened as the photos were taken. She could easily compare the apprehension of the older person with her own mental image. It was observed that both people with dementia and people without learned the names of neighbors at the nursing home after they had repeatedly watched the internal TV channel.

Berit's use of her photo album depends on with whom she viewed it. With her contact person, she insisted on handing over the remote control, but with me she browsed more independently, and her relatives noticed a considerable change in Berit's memory, vitality and health. The relatives also noted that they always had enjoyable times when they met Berit to browse through the photos and that she participated constructively in the dialogues that took place. This can be because the selection of photos was made to reflect what Berit remembered, and furthermore she guided the conversation by changing photos. Probably the discussions came to focus on topics that Berit could contribute to, and that the relatives had associations with. This is different from the photos on the internal TV channel to which Berit did not have much to contribute, even if she was told about the contents.

7. Discussion

It is clear that some artifacts do not become involved in the intended use, for instance, poorly designed remote controls. However, the everyday problem solving attitudes among older people, relatives and care workers can break such barriers in innovative ways through distributed memory processes, although there are barriers that can be broken more easily if the “right” artifact is introduced or iteratively designed. It is also clear that TV watching at the nursing home exploits distributed memory, where the outcome cannot be solely attributed to the memory of individuals.

7.1 Designing for Distributed Memory

Marie’s use of the universal remote control, Berit’s use of the TV photo frame, the significance of Berit’s TV photo album for the interaction between her and her relatives and Anna’s interaction with care workers are examples of situations where the older person took an initiative. These examples show that it is important that the artifacts of *extended TV* are parts of the users’ distributed memory, meaning that they should represent memories to the individual that can be used in processes with distributed memory. This gives them greater opportunities to be in control and to choose what to do. In fact, this gives more freedom to the older person.

Furthermore, in cases where the older person enjoys a functionality, for example viewing photos, the designer should see the older person as the main user, asking questions such as, “How can the older person maintain control?” and “How can the older individual’s desired interaction with others be maximized?” A designer should not start with the thought, “Let’s make a simple interface; the functions that are too difficult for the older person to operate are for the care workers to handle.”

With a different approach, the care workers and relatives can better understand where their human support is needed. This means that their support becomes more individualized in nature. Furthermore, to design systems in this empowering way makes the older person's wishes visible, and thus the care workers and relatives can learn more about them. The preferences of the older person may change with time, and a deteriorating health status can indicate the need for more human assistance; hence the need for more memory functions distributed to care workers and other individuals.

One example of this design approach is an individualized TV channel, which shows all the programs the older person normally watches on the same channel. The selection of TV programs will be a direct translation from the older person, relatives and care workers to an artifact. Furthermore, browsing among channels should be possible through a trial and error method. This means that it should be impossible to end up in modes that the older person is unable to handle.

It can also be worth considering whether the older person's TV should accompany them into the nursing home to increase familiarity in the new setting. This could influence the distributed memory and draw on the older person's earlier experiences of operating the TV. However, if the TV is old, one can argue that it should be exchanged for a new one due to fire hazards and the space occupied.

Empirical findings suggest that assistive technology can reduce the need for the personal assistance given to older people (Hoenig, Taylor, & Sloan, 2003). As earlier mentioned, personal assistance can be of vital importance in lowering the feelings of loneliness among older people at a nursing home (Drageset, 2004). Hence, there is a danger that individually adapted technical solutions may leave older people lonelier than before, even if that was not the case in this study. In addition, the heightened level of social contact for the older people due to the interest shown by me as a researcher can have affected the results in a positive direction, since it undoubtedly also counteracted their loneliness. This can, of course, be very significant during periods of stress after relocations. However, Berit, Anna and Marie did not show clear symptoms of relocation stress.

7.2 Method

The design method for extended TV is described, analyzed and commented on in (Abdelmassih Waller et al., 2007). Especially, the considerations taken to include older people with dementia in the design process are commented. Hence the discussion here is limited to discussing the use of distributed memory and FACE as tools in the design process. The results indicate that the artifacts support the older people if the artifacts act as representations in distributed memory processes. This interpretation was used to suggest modifications during the design process, but has been made more explicit in this article. The design process would most probably benefit from using a distributed memory analysis as well as a FACE analysis during the ongoing collection of data.

The lack of analysis of creativity and intentionality in this article suggests that this analysis should be combined with another type of scrutiny, perhaps by the use of situated action or activity theory. Indications of this need are Berit's lack of spontaneous use of her TV photo album, Anna's need to be reminded about the internal TV channel and Marie's problems of remembering the existence of her TV. This all concerns design for situations were the older person does not have a plan to use extended television. However, my belief is that a distributed cognition analysis directs the designer towards consistent structures that act as cognitive contours (Svensk, 2001, p. 51). These contours increases the older person's understanding of what to do, when to do it and what to expect. This is vital for the easy of use of a device (Norman, 1999, p. 174) and can be expected to be important for older people with dementia, since they have difficulties in compensating for badly designed artifacts. Furthermore, such contours can also support creative activities (Jönsson et al., 2006, p. 175).

7.3 Validity, Reliability, and Generalization

Here I define validity as the degree to which I have identified the correct causal links and explanations. This can also be called internal validity (Yin, 2003, p. 34). Action research obtains much of its validity from testing improvements in action by most at-risk stakeholders. This can give action research, in some regards, strong validity as compared to conventional social science (Brydon-Miller et al., 2003). There are, however, several threats to the validity of the conclusions drawn. These threats can be reduced by adding rigor to the research process (Figueiredo & Cunha 2007, p. 89). However, I have not, as recommended, used a theoretical framework from the beginning. Furthermore, I have not challenged the analysis using competing theoretical frameworks as suggested by Robson (2002, p. 174).

The reliability of action research projects, in the sense that it should be possible to repeat them, is generally low due to the uniqueness of each intervention. However, by using the analytical frameworks in this article (FACE and distributed memory), my interventions have given rise to recommendations that are consistent to literature on older peoples' memory and literature on how to design for older people. This makes it possible to generalize from the research setting.

8. Conclusions

The shortcomings in the support that replaces the lost distributed memory (the memory aspects of distributed cognition) of older people moving to nursing homes were described, and it was shown, by using distributed memory analysis and FACE analysis, how artifacts can strengthen the older person's actions and control when they are a part of his or her distributed memory. It was found benificiary if the artifacts represented memories to the individual that could be used in processes with distributed memory. In addition, it was shown that used analysis, can be useful for the design processes for extended television in nursing homes, in order to provide cognitive contours. However, design processes would benefit from being combined with an analytical framework that takes human creativity and intentionality into account. Further research is recommended.

9. References

Abdelmassih Waller, P., Östlund, B., & Jönsson, B. (2007). *The Extended Television - using tangible computing to meet the needs of older persons at a nursing home*. Manuscript submitted for publication.

Allen-Burge, R., Burgio, L. D., Bourgeois, M. S., Sims, R., & Nunnikhoven, J. (2001). Increasing Communication Among Nursing Home Residents. *Journal of Clinical Geropsychology, 7*(3), 213-230.

Anderberg, P. (2005). Making both ends meet. *Disability Studies Quarterly, 25*(3).

Anderberg, P. (2006). *FACE Disabled People, Technology and Internet.* Doctoral Dissertation. Sweden: Division of Rehabilitation Engineering Research, Department of Design Sciences, Lund University.

Beck-Friis, B. (2000). *Det blir lättare när det blir svårare. Råd till personal och anhöriga som vårdar demenshandikappade [It becomes easier as it becomes more difficult. Advice to care workers and relatives that take care of dementia handicapped]* (3rd ed.). Stockholm, Sweden: Förlagshuset Gothia.

Bourgeois, M. S., Dijkstra, K., Burgio, L., & Allen-Burge, R. (2001). Memory aids as an augmentative and alternative communication strategy for nursing home residents with dementia. *Augmentative & Alternative Communication, 17*(3), 196-210.

Brydon-Miller, M., Greenwood, D., & Maguire, P. (2003). Why action research? *Action Research, 1*(1), 9-28.

Caprani, N., Greaney, J., & Porter, N. (2006). A Review of Memory Aid Devices for an Ageing Population. *PsychNology Journal, 4*(3), 205-243. Retrieved March 23, 2007, from http://www.psychnology.org

Cohene, T., Baecker, R., Marziali, E., & Mindy, S. (2007) Memories of a Life: A Design Case Study for Alzheimer's Disease. In J. Lazar (Ed.), *Universal Usability: Designing Computer Interfaces for Diverse User Populations.* Chichester, England: John Wiley & Sons Ltd.

Drageset, J. (2004). The importance of activities of daily living and social contact for loneliness: a survey among residents in nursing homes. *Scandinavian Journal of Caring Sciences, 18*, 65–71.

Ely, M., Friedman, T., Gardner, D., & McCormack Steinmetz, A. (1993). *Kvalitativ forskningsmetodik i praktiken - cirklar inom cirklar* [Original title: Doing Qualitative research: Circles in Circles]. Lund, Sweden: Studentlitteratur.

Fisher, J. C. (1998). Major Streams of Research Probing Older Adult Learning. *New Directions for Adult and Continuing Education, 1998*(77), 27-39.

Figueiredo, A. D., & Cunha, P. R. (2007). Action Research and Design in Information Systems. In N. Kock (Ed.), *Information Systems Action Research. An Applied View of Emerging Concepts and Methods*. Springer. Retrieved July 19, 2007, from http://www.springerlink.com/content/h01764t4343j2p07

Gamberini, L., Alcaniz, M., Barresi, G., Fabregat, M., Ibanez, F., & Prontu, L. (2006). Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project. *PsychNology Journal, 4*(3), 285-308. Retrieved March 23, 2007, from http://www.psychnology.org

Hoenig, H., Taylor, D. H., Jr., & Sloan, F. A. (2003). Does Assistive Technology Substitute for Personal Assistance Among the Disabled Elderly? *American Journal of Public Health, 93*(2), 330-337.

Hutchins, E. (1995a). How a Cockpit Remembers Its Speeds. *Cognitive Science, 19*(3), 265-288.

Hutchins, E. (1995b). *Cognition in the Wild*. Cambridge, Massachusetts: MIT Press.

Hutchins, E. (2001). Cognition, Distributed. In N. J. Smelser and P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 2068-2072). Retrieved July 13, 2007, from http://www.sciencedirect.com/science/referenceworks/9780080430768

Johansson, P., & Gärdenfors, P. (2005). Introduction to Cognition, Education, and Communication Technology. In P. Gärdenfors and P. Johansson (Eds.), *Cognition, Education, and Communication Technology* (pp. 1-20). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Jönsson, B. (2003). *Elderly and Design*. Dept. of Design Sciences, Lund University. Retrieved March 23, 2007, from http://www.design.lth.se/aldreochdesign/elderlypeopleanddesign_screen.pdf

Jönsson, B., Anderberg, P., Brattberg, G., Breidegard, B., Eftring, H., Enquist, H., et al. (2006). In B. Jönsson (Ed.), *Design Side by Side.* Lund, Sweden: Studentlitteratur.

Jönsson, B., Philipson, L., & Svensk, A. (1998). *What Isaac taught us.* Lund: Certec, LTH.

Lee, D. T. F., Woo, J., & Mackenzie, A. E. (2002). A review of older people's experiences with residential care placement. *Journal of Advanced Nursing, 37*(1), 19-27.

Kaptelinin, V., & Nardi, B. (2006). *Acting with Technology: Activity Theory and Interaction Design.* Cambridge: MIT Press.

Kaptelinin, V., Nardi, B., Bødker, S., Caroll, J., Hollan, J., Hutchins, E., et al. (2003). Post-cognitivist HCI: Second-Wave Theories. In *CHI '03 extended abstracts on Human factors in computing systems* (pp. 692-693). New York: ACM Press.

Manion, P. S., & Rantz, M. J. (1995). Relocation stress syndrome: A comprehensive plan for long-term care admissions. The relocation stress syndrome diagnosis helps nurses identify patients at risk. *Geriatric Nursing, 16*(3), 108-112.

Nardi, B., (1996). Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition. In B. Nardi (Ed.), *Context and Consciousness. Activity Theory and Human-Computer Interaction.* Cambridge: MIT Press.

Nordicom. (2006, March 3). *Television viewing: Daily channel reach by sex and age 2004 (per cent).* Retrieved January 24, 2006, from http://www.nordicom.gu.se/common/stat_xls/690_5520_daily_reach_sex_age_2004.xls

Norman, D. A. (1993). *Things that make us smart. Defending human attributes in the age of the machine.* Cambridge, Massachusetts: Perseus Books.

Norman, D. A. (1999). *The Invisible Computer, Why Good Products Can Fail. The Personal Computer Is So Complex and Information Appliances Are the Solution*. Cambridge: MIT Press.

Orpwood, R., Bjørneby, S., Hagen, I., Mäki, O., Faulkner, R., & Topo, P. (2004). User involvement in dementia product development. *Dementia, 3*(3), 263-279.

Orpwood, R., Gibbs, C., Adlam, T., Faulkner, R., & Meegahawatte, D. (2005). The design of smart homes for people with dementia − user-interface aspects. *Universal Access in the Information Society, 4*(2), 156-164.

Östlund, B. (1995). *Gammal är äldst, en studie av teknik i äldre människors liv [Old people are the most experienced. A Study of the meaning of technology in old people's every day life*. Doctoral dissertation. Sweden: Institute of Tema Research, Linköping University.

Östlund, B. (2005). Design paradigms and misunderstood technology: The case of older users. In B. Jæger (Ed.), *Young technologies in old hands. An international view on senor citizen's utilization of ICT* (pp. 25-39). Copenhagen: DJØF Publishing.

Palen, L., & Aaløkke, S. (2006). Of Pill Boxes and Piano Benches: "Home-made" Methods for Managing Medication. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 79-88). New York: ACM Press.

Plato, G. & Jönsson, B. (2001). *Art and Science - A different convergence* (Rep. No. 1:2001). Lund: Certec, LTH.

Rahm Hallberg, I. (2002). Att vara vårdare [To be a care person]. In A. K. Edberg (Ed.), *Att möta människor med demens [Meeting people with dementia]* (pp. 45-60). Lund, Sweden: Studentlitteratur.

Robinson, J. D. & Skill, T. (1995). Media usage patterns and portrayals of the elderly. In J. F. Nussbaum and J. Coupland (Eds.), *Handbook of Communication and Aging Research* (pp. 359-391). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

Robson, C. (2002). *Real World Research* (2nd ed.). Malden, Massachusetts: Blackwell Publishing.

Son, G. R., Therrien, B., & Whall, A. (2002). Implicit Memory and Familiarity Among Elders with Dementia. *Journal of Nursing Scholarship, 34*(3), 263-267.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human machine communication.* Cambridge: Cambridge University Press.

Svensk, A. (2001). *Design for Cognitive Assistance*. Licentiate thesis: Dept. of Design Sciences, Lund University. Sweden. Retrieved July 17, 2007, from http://www.english.certec.lth.se/doc/designforcognitive/designforcognitive.pdf

Svidén, G., Wikstrom, B. M., & Hjortsjö-Norberg, M. (2002). Elderly Persons' Reflections on Relocating to Living at Sheltered Housing. *Scandinavian Journal of Occupational Therapy, 9*(1), 10-16.

Topo, P., Mäki, O., Saarikalle, K., Clarke, N., Begley, E., Cahill, S., et al. (2004). Assessment of a Music-Based Multimedia Program for People with Dementia. *Dementia 3*(3), 331-350.

Wilson, S. A. (1997). The transition to nursing home life: a comparison of planned and unplanned admissions. *Journal of Advanced Nursing, 26*(5), 864-871.

Woodruff, D. S., & Lemieux, S. K. (2001). Memory and Aging, Neural Basis of. In N. J. Smelser and P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 9560-9563). Retrieved July 13, 2007, from http://www.sciencedirect.com/science/referenceworks/9780080430768.

Yin, R. K. (2003). *Case Study Research. Design and Methods* (3rd ed.). Thousand Oaks, California: Sage Publications.

Zimmer, Z., & Chappel, N. L. (1999). Receptivity to new technology among older adults. *Disability and Rehabilitation, 21*(5), 222-230.

PsychNology Journal, 2007
Volume 5, Number 2, pp. 133 -164

Development and Evaluation of a Method Employed to Identify Internal State Utilizing Eye Movement Data

Noriyuki Aoyama ♠♦ and Tadahiko Fukuda ♣

♦ Graduate School of Media and Governance, Keio University (JAPAN)

♣ Faculty of Environmental Information, Keio University (JAPAN)

ABSTRACT

In the attempt to recognize and estimate human internal states, such as varying emotions, psychological and conceptual conflicts pose interesting and challenging issues. In this paper, we explore a pattern recognition technique that can detect a state of confusion and can estimate human interest, each an internal state of mind. In order to automatically detect a state of confusion from the objective data made available to us, the technique we present relies upon eye movement data. We have conducted three experiments in which subjects are confronted with a task that includes a trap intentionally designed to confuse them. We have recorded their eye movement data. We demonstrate that approximately 89% of a state of confusion can be detected from eye movement data by using a backpropagation algorithm. Moreover, for estimating human interest, we present a technique that builds upon the foundation of our confusion detection technique. As a result, we can demonstrate that approximately 60% of human interest can also be estimated through eye movement data.

Keywords: *Confusion, Interest, Eye Movement, Human-Computer Interaction, Neural Networks.*
Paper Received 26/07/2006; received in revised form 22/05/2007; accepted 03/08/2007.

1. Introduction

As computer systems continue to improve upon their already high levels of performance and sophistication, the corresponding interactive systems that process information based upon user input are being utilized for numerous applications that improve functionality. This improvement process is predicated upon the fact that users tend to view sophisticated interactive systems as a servant who serves to satisfy various user demands (Reeves & Nass, 1996). Although originally intended to improve interactive system functions for the better, ironically, these improvements have also

♠ Corresponding Author:
Noriyuki Aoyama,
Graduate School of Media and Governnace, Keio University,
k205, 5322 Endoh, Fujisawa-shi, Kanagawa, Japan
E-mail: aonori@sfc.keio.ac.jp

increased the chances of creating user experiences in which complex problems are encountered, ones that cannot be solved by the users themselves. Most users expect the functionality of an interactive system to be flexible, yet scenario–specific, and responsive to any given situation. This very concept sets up the pitfall users can stumble into during the process of attempting to accomplish their program goals. When users are faced with interactive systems, they will attempt from the outset to work out quick strategies to accomplish their goals. The users then proceed in a methodical manner in pursuit of their goals. This is a perfectly acceptable strategy when the process progresses favourably and satisfactorily. However, if the interactive system generates an output that is different from a user’s expectation, the user becomes confused and can even suffer from psychological stress. As a consequence of this experience, some users may go so far as to try to avoid the use of interactive systems altogether. In a confirmed trend, some users have been found to feel the same type of responsibility that a manufacturing company may feel when a functionality or safety problem occurs, as well as feeling a similar responsibility when usability and convenience problems occur (Norman, 1988). Whether or not an individual will experience this phenomenon is dependent upon his/her unique sense or consciousness. That is, users can experience conflicts between their desires with respect to convenience improvements and their desire to use the right products and their overinflated confidence about their literacy in utilizing such convenient products. Indeed, these factors often stand in the way of completely successful user utilization of an interactive product. This dynamic is also made clear when examining the choices that industrial manufacturers make in adopting more human-centred product designs.

Firstly, the concept of human-centered design was originally proposed by Norman and Draper (Norman & Draper, 1986) and it became the front line of product development, preferencing utility over usability. This concept has since been propagating the model of giving more weight to actual usage scenes within the product design process. This user-oriented concept has now become a fundamental product development concept.

Secondly, the human-centred design concept has favourably enhanced user satisfaction in terms of the notion now referred to as “usability”. It was proposed by Nielsen (Nielsen, 1993), who had been inspired by several other human-centred design related concepts, such as universal design (Mace, 1985), cognitive engineering (Card, Moran, & Newell, 1983), and Kansei engineering (Nagamachi, Ito, & Tsuji, 1988). The usability concept, which incorporates three other fundamental concepts, is absolutely essential for product development.

Finally, manufacturers must currently attend to usability-satisfied users' increscent demand for interactive systems. That is, the market demands ever more personalized designs befitting individual users and their unique situations. The term "adaptive" best describes this new concept. Since the late 1990s, the term "adaptive user interface" has been used to describe this new interaction design within human-computer interactions (HCI) (Maglio, Barrett, Campbell, & Selker, 2000).

To better realize this adaptive user-interface, many researchers from a number of fields, such as artificial intelligence, ergonomics, and HCI, are working to recognize and estimate people's internal landscapes. These researchers are pursuing the possibility of developing diverse new interactive systems. For example, several research groups have attempted to build robots that possess human-like emotions (Hara & Kobayashi, 1996; Lim, Ishii, & Takanishi, 2000; Zecca et al., 2004). Conversely, another researcher is working to develop intelligent systems (e.g., an automatic breaking system for automobiles) that are designed to avoid dangerous or hazardous situations originally caused by human error. The critical factor is that these adaptive interactions need to operate automatically without the intervention of any user autonomous behaviours and intentions. For instance, when a user uses an interactive system (e.g., a Japanese bank ATM, which is highly complex in comparison to ATMs in other countries), that system should, first of all, provide the shortest possible path to the user's goal and preclude any information that is not indispensable to that user at that moment. If the user is unable to attain his/her preset goal as a result of problems, the system should automatically present new, detailed, and useful information that can reverse the situation and once again place the user on the right path to his/her goal. Consequently, the amount of time needed for a user to achieve the goal will overall decrease.

In this study, an attempt was made to develop a method that can identify user internal states by utilizing eye movement data during HCI. In order to conduct this research, the term *internal state* had to first be defined. Most relevant research studies in this field that had aimed to identify the "internal state" of humans had focused on identifying emotions that many researchers catalogued as basic emotions. However, basic emotions and research issues related to them had not been specifically defined, although researchers have been attempting to define emotions since the days of Plato and Aristotle dating back to circa 400 B.C. This research issue has been addressed contemporarily by Tomkins, Ekman, Izard, and others (Tomkins, 1962, 1963; Izard, 1991; Ekman, 1992). Although one might say that this type of work has been ongoing for more than two millennia, even the most basic emotions, serving as building blocks for the more complex emotions, remain undefined. The reason for this lies with the lack of clear and numerical criteria for using these terms that indicate an emotional state. In addition, emotional states differ in meaning according to who is experiencing them and in what context. Therefore, in this study, a specific definition of *internal state* was avoided. On that basis, *internal state*, in this study, will not be defined in detail but will instead be defined in the broadest of terms. The term will be defined in terms of a specific situation and context, such as the happiness of accidentally encountering best friends. *Internal state* eventually became defined in this work as those physiological and psychological state changes caused by the acquisition of external information and the internal processing of that information. In particular, the internal states of confusion and interest were focal points for us within the framework of potential internal states, because both confusion and interest are practical and highly versatile internal states when considering the real world of HCI.

Our hypothesis in this study was that the changes in a person's internal state are somehow reflected in his/her eye movement data, and this is the reason why we employed eye movement data to detect people's internal states in this study. Our hypothesis was proposed based upon our consideration of three factors: (1) previous studies conducted in related research topics; (2) our modification of the research concept for our study; and (3) the ease of measurement for our study.

Previous research has determined, as acknowledged in the first consideration point for our hypothesis, above, that although internal state information is not contained within eye movement per se, but that the eye itself, e.g. pupil size, indicate the changing human psychological makeup. Hess conducted experiments indicating that pupil size changed in accordance with the subject's interest in objects being viewed (Hess & Polt, 1960; Hess, 1965). An interpretation of this result can be that human interest affects their eye pupil sizes and this interpretation pointed to our own hypothesis. However, there was no previous study in which eye movement data was directly connected to changes within individual internal states.

Our second point in considering our hypothesis was the modified research concept. The traditional research concept utilizing eye movement data focuses more on how humans acquire visual information. The traditional concept focused on the activity of human information acquisition (e.g., Byrne, Anderson, Douglass, & Matessa, 1999;

Hayhoe & Ballard, 2005). In contrast, our new concept is based upon how humans communicate their internal information, bringing it into view through their eyes. In general, humans communicate more easily with other people in face-to-face situations (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Driver et al., 1999; Macrae, Hood, Miline, Rowe, & Mason 2002). This concept of face-to-face communication focuses on the activity of human information-providing. In addition, we found biological support for our hypothesis; for instance, human eyes have unique characteristics, such as the proportion between the exposed white sclera and the surrounding darker-coloured iris, the largest width-height ratio within the eye's outline, and the whitest of exposed sclera among all primates, which enables communicating with ease by using a gaze signal (e.g., Morris, 1985; Kobayashi & Kohshima, 1997; Kobayashi & Kohshima, 2001). Finally, the reason for using eye movement data is that such data can be easily collected from subjects without direct physical contact. The aforementioned factors guided us in the formation of our hypothesis.

In previous studies, three primary methods have been used to identify internal states by recording external data without any need for direct physical contact. The first method recorded eye movement related data; the second method recorded facial expression data, and the third recorded speech data. Umemuro and Yamashita (2003) attempted to detect states of confusion and states of surprise by assessing pupil diameter data within HCI. These authors defined users as being in a state of confusion whenever any unexpected disturbance obstructed the users from continuing their task and those users were unable to find an appropriate solution for handling that situation. Surprise, in contrast, was defined as the status of users upon recognizing an unexpected event, but such users were still able to continue their task. Researchers collected pupil diameter data from memory task trials, and then analyzed the data distribution. As a result, the subjects' confusion and surprise could be detected at rates of 25% and 65%, respectively. However, the pupil diameter data acquired with this method were inadequate because such data included the interference effects of the surrounding environment and other visual stimulation. For example, the amount of ambient light, as well as the colour and brightness of visual objects, all affect and impede pupil diameter data.

There have also been studies utilizing facial expression data (Yacoob & Davis, 1996; Lyons, Akamatsu, Kamachi, & Gyoba, 1998) and speech data (Dellaert, Polzin, & Waibel, 1996; De Silva & Ng, 2000; Nicolson, Takahashi, & Nakatsu, 2000). Yacoob and Davis explored facial expression recognition by using optical flow data. They focused on long-term image sequencing to recognize seven facial expressions—six different facial emotions and one "neutral" face—by integrating spatial and temporal information from video images. However, the recognition rate for the emotions of fear and sadness was low in comparison to the relatively high recognition rates for the emotions of happiness, anger, and surprise. Lyons et al. (1998), who conducted another study using facial expression data, tried to extract and code facial expression information from images by using a Gabor filter and calculating the geometric aspects of facial expressions. Lyons et al. analyzed the correlations between the Gabor-coded data and the semantic or geometrical ratings assigned to the respective stimulant images of the subjects participating in the experiment. They found the average correlation rate of a Gabor model to be approximately 0.57 and that of a geometrical model to be approximately 0.37.

These facial expression studies illustrate the same problem that affects the current recognition approaches being tested. These methods cannot recognize natural expressions. They are only capable of recognizing those artificially reproduced expressions that subjects have been asked to make with specific facial movements to display expressions for various emotions.

Using speech analysis, Dellaert et al. (1996) explored statistical patterns in an attempt to recognize four emotions based upon the pitch data of test subjects' speech. Dellaert et al. analyzed the data by comparing the performance of their proposed recognition technique in relation to human performance. The result was that the recognition technique indeed had some success with the carefully orchestrated human test performances. However, this technique had substantially high error rates that ranged from 20% to 44%. Nicolson et al. (2000) tried to recognize eight emotions by analyzing the speech data within HCI. They used neural networks to recognize emotions using both phoneme and prosodic features data. As a result of the analysis of the recognition rates of three networks namely, One-Class-in-One, All-Class-in-One, and Learning Vector Quantisation, the highest recognition rate recorded was approximately 57% for the All-Class-in-One neural network. This study served quite favourably as a reference but, nevertheless, the method was not sufficiently well adapted to use HCI without speech. Moreover, De Silva and Ng (2000) tried to recognize six basic emotions using both the data of audio waveforms to analyze emotional speech by Hidden Markov models and the data of video images to analyze facial expressions using statistical techniques. The recognition rate extracted from video-image data was approximately 65%, and the rate for audio data was approximately 35%, with a 72% rate for bimodal data. Although those recognition rates were relatively high, it is still difficult to say that this system is practical because it requires creating a user-by-user recognition model.

Studies that require physical contact for acquisition of the data necessary for estimating feelings and recognizing emotions have also been conducted (Musha, Terasaki, Haque, & Ivanitsky, 1997; Ishino & Hagiwara, 2003; Takahashi, 2004). Ishino and Hagiwara tried to estimate four different emotional states by reading electroencephalography (EEG) data and using a neural network. Their estimation rate was approximately 61%. Takahashi tried to recognize the states of pleasure and displeasure from EEG data by using a neural network and a support vector machine. The higher recognition rate was 62.3% for a neural network classifier. These results indicated that EEG could certainly provide a cluster of useful data for understanding internal states. However, the current devices used to acquire EEG data are impractical and unrealistic for use in HCI applications because such devices are prohibitively clumsy and disruptive. In addition, subject internal states could have been adversely affected by the extensive and complicated nature of the measurement system method itself.

Eye movement data can be provided without any physical contact using the above-described equipment and methods that were utilized in previous studies. However, because most commercially available eye tracking systems presently offered use infrared rays to detect eyes' positioning and orientation, it is necessary to ensure a lighting environment that is measured with absolute precision. A new technology that can adjust to a changing light environment has been recently developed using image processing technologies and algorithms (Yamanobe, Taira, Morizono, & Kamio, 1990; Zhu, Fujimura, & Ji, 2002). These features provide great advantage in working with internal states using eye movement data. Alternatively, the eye tracking method also poses a disadvantage. Namely, all existing equipment used to record eye movement requires calibration before commencing any eye movement recording. This calibration is necessary in order to measure eye movement with the potentially highest accuracy. This is a technical issue that many researchers have attempted to address (e.g., Ohno, 2006). However, in the case of a contact-free eye tracking system, the calibration procedure is easier for the researcher and less burdensome for the subjects when compared with other equipment and methods that could also be used to acquire objective data. In fact, eye movement data can be considered to be the most reliable and comprehensive method for assessing the internal state of a human being.

If another method could be developed that identifies human internal states using data obtained without any physical contact, then those systems that users interact with would be able to provide appropriate information to fit the use context and situation depending upon the user's internal state. This would enable not only the possibility of increasing users' chances of troubleshooting problems on their own without needing system help, but it could also decrease the chance of users experiencing psychological stress.

2. Confusion Detection Method

This term "confusion" had to be defined at the outset, before conducting any experiments. It is difficult to define the meaning and express it in a single word because of the high variability of human internal states. This is the very reason why researchers typically use more than one word to describe an internal state and the terms used to describe an internal state have varied by subject. For example, there are certain words that are related to confusion, such as puzzlement, embarrassment, and panic. However, these words have been used appropriately in accordance with each particular situation. Therefore, in this study, the meaning of the word "confusion" has been defined in relation to the actual context in which such confusion has occurred.

2.1. Experiment 1

Our first experiment was conducted to demonstrate the fact that eye movement data can be used successfully to detect a person's internal state. This experiment was designed to recognize a state of confusion from eye movement data in a context without any physical contact, while the interaction of multiple experiences among subjects still was taken into consideration.

2.1.1. Experiment System

The experiment system consisted of two personal computers (PCs). The first PC was used to output stimuli onto its monitor. The other PC controlled the eye tracking system. A simple system configuration is shown in Figure 1. Stimuli were presented on a SyncMaster 172MP-R (Samsung Japan Corporation, Tokyo, Japan) with a 17-inch liquid crystal display. Eye movement data was gathered using the image and the cursor, in which the subjects' sight points were superimposed. To record the superimposed images, two SSC120EX scan converters (Canopus Co., Ltd., Hyogo, Japan) were used. These images were then recorded by a Sony WV-DR7 (Sony Corp., Tokyo, Japan) DV/VHS combo cassette recorder. The eye movement data obtained from the dominant eyes were measured using the FreeView eye tracking system (Takei Scientific Instruments Co., Ltd., Niigata, Japan). The FreeView is a non-contact eye tracking system, which enabled easier operation of the computer mouse by seating each subject on a high-backed chair with a headrest, where each subject was instructed to place his/her head on the headrest to establish the correct distance between the eye camera and the subject's eye. This experimental setup is illustrated in Figure 2. The distance from the subject's eye was fixed at 70 cm. The FreeView system tracked each subject's eye movement automatically by employing a pupil-corneal reflection method using an infrared ray at a sampling rate of 30 Hz. This number was dependent upon the NTSC frame rate. In this paper, the word "frame" is always used in this context. The precision rate was far less than 0.1 degrees, even under the worst conditions, and the measurable range was ±20 degrees in both horizontal and vertical directions.

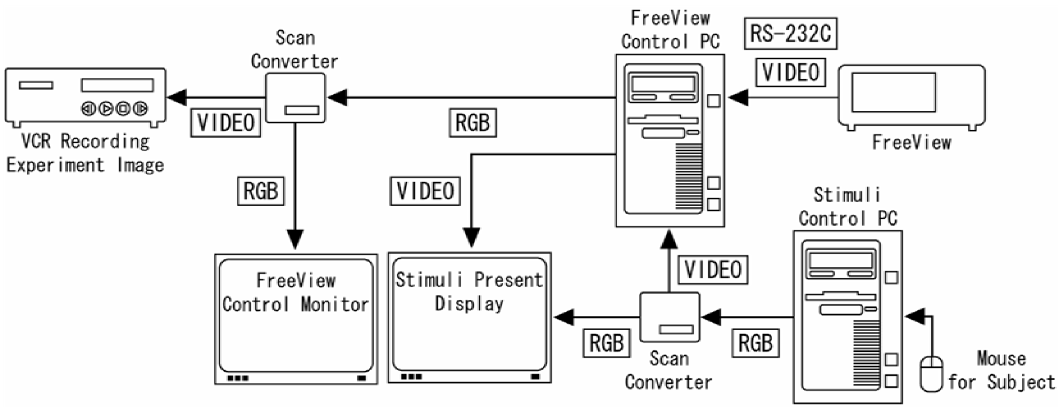


Figure 1. Configuration of the Experiment System. The experiment system consisted of a detection station for measuring eye movement, two PCs with displays for controlling the FreeView system and presenting stimuli, two scan converters for converting signal and branch signals, and a video cassette recorder.

Figure 2. Experimental Setup. The subjects sit in front of the display without any device making physical contact with them. Eye movements are measured while each subject performs tasks by manipulating a computer mouse.

2.1.2. Task and Procedure

In Experiment 1, the types of confusion caused by two different scenarios were identified as described below:

(1) The confusion caused by interrupting a plan that the subject has simulated; and

(2) The confusion caused by displaying a result differing from what each subject has anticipated.

Four types of stimuli, labelled as Task 1 through Task 4, were constructed for use with simple computational problems. Task 1 and Task 2 were designed to stimulate the first type of confusion described above, and Task 3 and Task 4 were designed to stimulate the second type of confusion. Each stimulus included a trap designed to confuse the subjects. The type and details of each stimulus are shown in Table 1. Sample screens of the first and last screens (the trap screen) in Task 2 are shown in Figure 3.

Task	Type	Detail
1	Adding and/or subtracting numbers to achieve a target value	The screen displays a present value and a target value, as well as two buttons, one assigned with a positive number and the other with a negative number. The subjects will try to reach the target value by selecting one button. They need to select the positive number because the present value is lower than the target value. However, in the process, subjects will notice that the positive number button option, when selected, creates a result exceeding the target value.
2	Answer selection of computational problems	The screen displays several computational problems and subjects must choose from among three answers. The next problem will be presented only when the subjects have correctly answered the current problem. During this process, however, the subjects will realize that no correct choice is being offered to solve the problem at hand.
3	Answer selection of computational problems	The screen displays several computational problems and subjects must select from among three answers for each problem. The next problem will be presented only when the subjects have correctly answered the current problem. During this process, however, the subjects will face a problem in which all the choices offered are the correct answer.
4	Adding and/or subtracting values to achieve a target value	The screen displays a present value and a target value, as well as two buttons, one assigned with a positive number and the other with a negative number. The subjects will try to reach the target value by selecting one button. They need to select a positive number button because the present value is lower than the target value. However, during this process, the subjects will discover that the selection buttons do not have any number at all assigned to them.

Table 1. Type and Details of Each Stimulus in Experiment 1

Figure 3. Sample Screen Shot of Task 2. The subjects were asked to choose the correct answer for the calculation problems (left). However, no correct answer is among the choices for the problem at hand (right).

The subjects of this experiment consisted of 65 males and 31 females. They were both undergraduate and graduate students between the ages of 18 and 24 years old. All reported that they had normal vision, including corrected eyesight.

Each subject carried out only one randomly selected task. And, in pilot test, several subjects were asked to comment few words about the experimental procedure and the tasks after the experiment for any revisions that may need to be made for the actual experiment. In doing so, the issues of interaction between tasks and the possible psychological impact of the task including a trap to stimulate the eye movement data were all considered. The breakdown of the final number of subjects assigned to each is shown in Table 2. In this Table, the subject number of Tasks 2 and 4 was larger than Tasks 1 and 3 because the large number of subjects was assigned to Task 2 and 4 based on the subjects' comment that among the four stimulant tasks, the operations within Tasks 2 and 4 were the two most comprehensible ones as a stimulus. The more comprehensible the task, the acquired data from the experiment was thought to be easier to recognize the pattern. Therefore, this experiment was conducted, based on the classification of two types of simple confusion state above, the variance of each type of simple confusion state and each Task were not considered to have major influence on the result of eye movement output by humans.

However, in the post-pilot test interviews, the subjects commented that among the four stimulant tasks, the operations within Tasks 2 and 4 were most easily comprehensible as stimuli. Therefore, in order to achieve our objectives and to accomplish our mission, Tasks 2 and 4 were given more emphasis within this experiment's structure. The breakdown of the final number of subjects assigned to each task is shown in Table 2.

Task	1	2	3	4
Number of Subjects	11	45	11	29

Table 2. Breakdown of the Number of Subjects Assigned Each Task

2.1.3. Detection Method

Based upon our hypothesis that internal state changes are somehow reflected in eye movement data, we included obtaining data about velocity changes for lines of sight within this study. This decision was based upon several sets of analyses acquired as a result of conducting the pilot test. A velocity change in a line of sight was calculated by using the difference from the previous frame. The data set of each frame was calculated by using the sight point location data shown on the display as acquired through the FreeView system. Degrees-per-second was adopted as the measurement unit because the sight point location data measured the degrees from each subject's eyes as a unit.

In order to create the confusion detection model, we attempted to use a learning neural network that utilized eye movement data as input data for extracting various features of a state of confusion when subjects encountered stimulating traps. Before compiling this report, we employed approaches that focused on the data of each frame as acquired through the FreeView system, but we were unable to grasp the specific difference between normal phase data and trap phase data that can detect a state of confusion. Therefore, the learning algorithm of a neural network was employed to treat eye movement data as a cluster of historical data extended to multiple frames to consolidate the features of a state of confusion during a relatively long duration. The model created could serve as a template and would include detecting the features of confusion as acquired from eye movement data that are considered as the time-series data.

2.1.4 Neural Network Learning Algorithm

The neural network used for learning in this study is a backpropagation network algorithm using the sigmoid function. This network had eight layers, including six hidden layers. The construction of units in each layer is shown in Table 3.

Layer	IL	HL1	HL2	HL3	HL4	HL5	HL6	OL
Number of Units	50	100	50	40	30	20	10	1

Table 3. Construction of Units in Each Layer

The input data mass number was set at 50 frames. This number was based upon the metal operator proposed by Card, Moran and Newell (1980). We set the time-duration for approximately 1.67 seconds. However, the setting of the metal operator was approximately 1.35 seconds. Because our setting scenario manipulated information through use of the Internet, it was a more complex one than the Card et al. setting, which simulated a simple HCI situation with a computer mouse.

The learning policy was such that the data set for the first 50 frames at the time of the subjects' initial encounter with the trap screen would be fixed to "1" as a target output. The other data sets for before the subjects encountered the trap screen would be fixed to "0" as a target output. However, all the data, excluding the first 50 frames of the trap screen phase, were not used in the learning because they would be considered to be noise in the learning of the neural network. The data input was carried out in such a manner as to input the neural networks by delaying one frame at a time. This learning policy is shown in Figure 4. The data was normalized while the input of the data of line of sight was carried out. The learning rate in the learning process was set at 0.001, and the neural network ceased learning when the value of the cost function reached 0.00001.

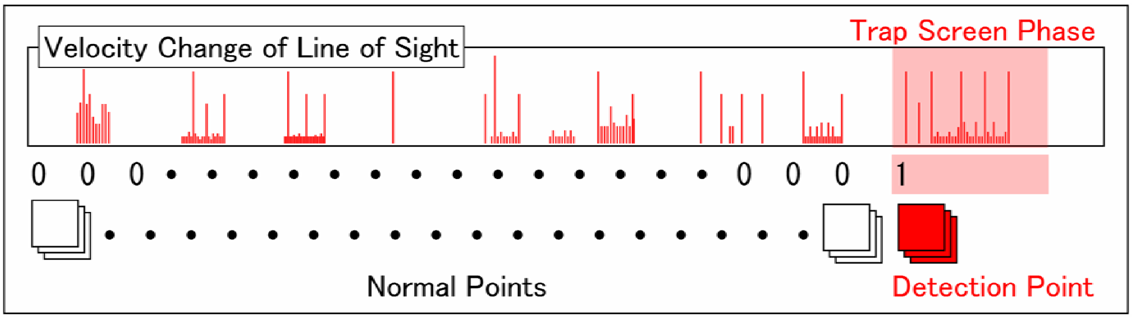


Figure 4. Learning Policy of the Experiment 1 Neural Networks. The unit of one data set was fixed at 50 frames. The data set other than the expected detection point was labelled as "0". For the anticipation detection data set, a label of "1" was given. The data set labelled "1" was counted from the point where the trap screen was presented and continued for 50 frames beyond that point.

Because the experimental objective was to develop a method for detecting the basic features of a state of confusion, the neural network learned by grouping together the four types of data that were acquired from this experiment.

2.1.5 Detection Results

As a result of the neural network learning using approximately 70% of the data acquired during the experiment, the learning coalesced and thus the confusion detection model was created. To confirm the detection accuracy of the model created, verification was conducted using approximately 30% of the data as testing data, data that had not been previously used for the neural network learning. A detection condition was judged to have occurred when the output unit of the confusion detection model provided a positive value at the points where it was expected that a subject's confusion would be detected. Therefore, we focused only on the results detected at those points where the detection of confusion was anticipated. We decided not to discuss false detection results, such as any output values of less than 0.5 at random points other than those points where it was expected that the output unit would make its detection. Moreover, in this report, the detection rate refers to the detection ratio of the anticipated points from the eye movement data as time-series data created by the model.

Verification showed that the detection rate for testing data was 100%. That is, all points that should have been detected as causing a state of confusion were indeed detected. The sample results are shown in Figure 5. The figure is comprised of three charts: the first chart (top) are the velocity changes within the lines of sight; the second chart (middle) shows the points where a state of confusion was expected; and the third chart (bottom) shows the output unit values. There are numerous peaks in Figure 5 because the verification was conducted in such a manner as to combine all the variables into one chart. However, only one point per subject is anticipated.

In this verification, the control situation with “no confusion” was not treated and discussed, although the reliability could improve if the same pattern with “confusion” were found from the data with “no confusion”. Since the neural network was learned by treating 50 frames as a data unit and delaying one frame at time, and then number of points given “1” as target output of a neural network on eye movement data was extremely smaller than the number of points given “0”, the reversing of our proposed learning algorithm in this study was not conducted and was not deal with in this paper.

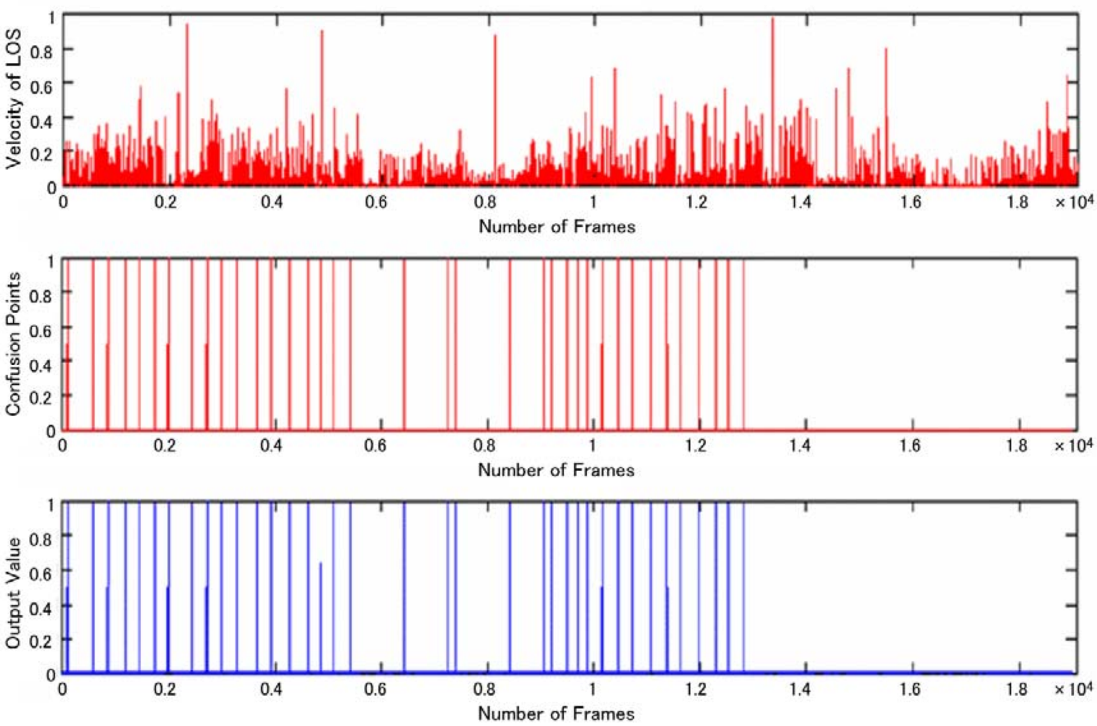


Figure 5. Verification Results for the Confusion Detection Model. The top chart shows the velocity changes with the lines of sight (LOS); the middle chart shows those confusion points where the model expected detection to occur; and the bottom chart shows the output values of the output units of the neural network.

Moreover, to verify the effectiveness of the model, we conducted another detection test using the noise-added data for further verification, including those data used for the neural network learning. The noise was calculated within a range of ±5% for each frame as a random value, and was then added to the original testing data. The detection rate of the confusion detection model with regard to such noise-added data was 100%.

2.2. Experiment 2

This experiment explored the possibility that a human state of confusion could be detected by a learning neural network using eye movement data limited to four kinds of stimuli in two key situations where subjects are contending with basic calculation problems, including a trap. The results of this experiment indicated that it was important to teach each neural network using the eye movement data recorded in each situation. Therefore, we tried to detect a state of confusion from eye movement data in a more realistic situation.

2.2.1. Experimental System

All factors in this experimental system were the same as in Experiment 1, except for the positioning of a video camera recorder with a microphone to film the display screen and simultaneously record the progress status of the task and the subject's voice. We used a Sony VX2000 (Sony Corp., Tokyo, Japan) video camera recorder.

2.2.2. Task and Procedure

The experiment focused upon using the Web and defined two situations to clarify states of confusion for Experiment 2, as described below:

(1) A state of confusion caused by not being able to find an item needed to accomplish one's goal; and

(2) A state of confusion caused by viewing a screen displaying an unexpected result and causing doubt as to whether or not one has committed an operational error.

As stimuli, two types of Web sites, each of which was constructed by drawing upon conventional Web services, were used. These stimuli were expected to provoke the two different states of confusion as described above. The type and details of each stimulus are shown in Table 4 and a sample screen shot for choosing a branch office in Task 5 is shown in Figure 6. The experimental procedure was modified for experiment 2 since it was more complex than the tasks in the experiment 1. In the experiment 1, the interviews were not conducted for all subjects. This decision was made from the point that no confirmation would be needed because the tasks were simple enough to understand. However, in the experiment 2, the interviews were considered to be a necessity for all subjects to confirm and to hold the reliability of the detected confusion because of the natural complexity of the task.

Task	Type	Detail
5	Online bank transfer	The subjects are asked to transfer a specified amount to a designated account displayed on the Web browser. However, when each subject selects the branch name, the system refuses to immediately display the target branch name on the screen (a type of time-lagged system). At this point, the subjects are confused.
6	Online hotel room reservation	The subjects are asked to make an online reservation on a Web browser in order to stay at a particular hotel with a friend. Each subject selects "two persons" on the third screen for the number of guests. However, the subjects will encounter incorrect output in the contents confirmation phase on the sixth screen, where the number of guests is displayed as seven persons. At this moment, each subject will start to harbour doubts about his/her own mistake, and will be confused.

Table 4. Type and Details of Each Experiment 2 Stimulus

Figure 6. Sample Screen of the Task 5 Shot. In this trap screen, the branch name choices are arranged in two columns and nine rows. However, the selected branch name cannot be found on the list.

The subjects of this experiment consisted of 44 males and 38 females. These subjects were both undergraduate and graduate students between the ages of 18 and 24 years old. All reported that they had normal vision, including corrected eyesight.

As was the case with Experiment 1, each subject carried out one randomly selected task. The number of subjects who carried out Task 5 was 31 and the number of subjects who carried out Task 6 was 51.

2.2.3. Detection Method

The detection algorithm was diverted from the algorithm that was used in Experiment 1. The neural network learning algorithm was the same used in Experiment 1. However, the neural network learning was conducted separately in the two tasks.

2.2.4. Detection Results

As a result of the neural network learning that used approximately 70% of all the data, the learning coalesced and a confusion detection model was created. To verify the detection accuracy of the model created, verification was conducted using the remaining approximate 30% of all the data as testing data. In addition, as a result of the recording during the experiment itself as well as the post-experiment interviews, the data revealed that several subjects had completed Task 6 without any awareness of the trap designed to cause confusion. These subjects' data were excluded from the neural network learning and from the detection rate verification. The post-verification detection rate was 100% in Task 5 and approximately 77.4% in Task 6.

2.3. Experiment 3

The results of Experiment 2 clearly indicate that a human state of confusion can be detected from eye movement data, even in realistic situations. Moreover, this shows that a human state of confusion can be detected by creating a model for each situation using eye movement data that were acquired from the respective situations. Consequently, an attempt was made to confirm whether age differences were a clearly indicated factor in the existence or lack thereof of a state of confusion. In order to verify the age difference variable, another test was conducted that contrasted the younger and older sets of test subjects from Experiment 3.

2.3.1. Experimental System

All components of this experimental system were the same as those used in Experiment 2, including use of a video camera recorder with a microphone.

2.3.2. Task and Procedure

In this experiment, Task 6, which was originally prepared as a stimulus for Experiment 2, was used because feedback and comments received in the post-Experiment 2 test subject interviews pointed out that Task 6 was easier to understand than Task 5.

The subjects of this experiment consisted of 62 older people, each with a minimum age of 65 years. However, among the data sets acquired, only 22 data sets were usable for objective analysis. This resulted from the eye movement data of older people being rather difficult to measure because of the diminished palpebra superior sizes caused by decreasing levator palpebrae superioris muscle tone. However, this particular subject group's computer and Web literacy skills did not pose any problem. The older test subjects belonged to a local computer club and possessed sufficient PC and Web navigation skills and knowledge.

2.3.3. Detection Method

The detection algorithm was diverted from the algorithm used in Experiment 1. The neural network learning algorithm was the same one used in Experiment 1.

2.3.4. Detection Results

As a result of the neural network learning that used approximately 70% of the total data, the learning coalesced and a confusion detection model was created. To confirm the detection accuracy of the model thus created, a verification of confusion was conducted using the remaining approximately 30% of all the data as testing data. In addition, it was confirmed that all subjects were confused at the trap point. The post-verification detection rate in Task 6 was 100%.

2.3.5. Confirming Age Differences

From the detection result of the model created using the data obtained from older people, it became clear that our proposed method had efficacy as a detection method for identifying the states of human confusion. Consequently, we also were able to confirm that the model built using the data obtained from young people was also reliable in detecting the state of confusion from the data of older people. After the confirmation, we checked as to whether the model created from the data of older people could also be applied to detect the state of confusion using the data obtained from young people. Both the Task 6 model and data in Experiments 2 and 3 were used for this confirmation work.

The results showed that the detection rate of the young people model using the data of older people was 80%. In contrast, the detection rate of the older people model using the young people's data was only 50%.

3. Interest Estimation Method

Utilizing both the results of the development and the evaluation of our proposed method for detecting states of confusion, we confirmed that the internal state of humans could be identified by a model created to use learned neural network by velocity changes within lines of sight as the input data in a similar manner for the measurement of interest. That is, user interest in HCI could be estimated by converting the technique employed in the confusion detection method.

Experiment 4, also designed on the basis of our hypothesis, was an experiment conducted to collect eye movement data from subjects that showed interest in visual objects on screen monitors.

3.1. Experimental System

The experimental system used was the same as that employed in Experiment 1, except the use of a computer mouse was omitted.

3.2. Task and Procedure

Three types of tasks were prepared for this experiment. The tasks comprised a slide show that incorporated eight images, each of which was created using Microsoft PowerPoint 2003. During the process of each task, an image was presented for five seconds. The subjects were asked to view these tasks. The type and details of each task are shown in Table 5. The female subjects carried out all the tasks. However, the male subjects carried out only two of the three tasks (Tasks 7 and 8) because the subject of Task 9 was women's fashions.

Task	Type	Details
7	Mixed topic	Eight images from varied topics, such as beautiful scenery, actors/actresses, young animals, etc.
8	Automotive topic	Eight images pertaining to automotives. These images ranged from a station wagon, SUV, sedan, and so forth.
9	Fashion topic	Eight images pertaining to women's fashions. These images included different types of clothing, such as knitwear, long sleeves, a shirt, a coat, and others. These stimuli were used only for female test subjects.

Table 5: Types and Details of Each Experiment 4 Stimulus

The subjects were asked to evaluate their interest in the eight images after each task. Five containers and printed out images of the stimuli presented in the tasks were prepared. These were aligned alongside the experimental system in order to measure the eye movement data. The five containers were numbered one through five. These numbers indicated the intensity of interest. The specific evaluation procedure was as follows:

(1) The subjects were asked to select the one image that piqued their keenest interest from among the eight images.
(2) The subjects were asked to evaluate on a scale of one to five their intensity of interest in that single image and to put the image in the container.
(3) The subjects were asked to select the image that they thought had most keenly interested them from among the remaining seven images.
(4) The same evaluation sequence was repeated with the remaining six images and so on, until the last image had been evaluated.

The goal of this procedure was to identify the exact image that piqued the subject's keenest interest. Moreover, within the evaluation phase, this procedure made it possible to avoid the grading fluctuations that typically occurred when the images were sorted into numbered containers.

The subjects of this experiment consisted of 53 males and 46 females. They were all undergraduate and graduate students between the ages of 18 and 24. All reported that they had normal vision, including corrected eyesight.

3.3. Estimation Method

The data quantifying the velocity changes within the lines of sight was utilized as a parameter because the method used for detecting a state of confusion had been successful. The velocity changes within lines of sight were calculated in the same manner as the confusion detection method velocity changes within lines of sight had been calculated.

Firstly, to create an interest estimation model, the data were screened in accordance with the evaluation conditions. That is, did the subjects evaluate the images immediately or not? The data were then organized in order to check whether the first image chosen had piqued the subjects' keenest interests and thus were evaluated immediately. Those data where the subjects were more careful in making an evaluative decision were screened to check the possibility of whether the data included noise that could be relevant to tricking the subjects' imaginations. In that case, some data were not adopted for both the neural network learning and verification phases. The ultimate number relationship between the learning and the verification phases of the neural network is shown in Table 6.

	Number of Data	
	Learning	Verification
Task 7	25	57
Task 8	24	56
Task 9	9	31

Table 6. Breakdown of the Number of Data Used for the Neural Network Learning and Verification Phases

Secondly, in creating the interest estimation model, we endeavoured to use a learning neural network that employed those data quantifying velocity changes within lines of sight as the input data for extracting the features of a state of confusion when subjects exhibited interest in visual objects. The point where the subjects showed interest in visual objects was specified from the evaluation results assessed after collecting each task's eye movement data. The model created was used as a template, and it includes features of interest for making estimations from eye movement data within time-series data.

3.4 Neural Network Learning Algorithm

The learning algorithm and the neural network construction were both the same as those used for the confusion detection method. The mass of input data was also the same but the learning policy differed. The labels used in the confusion detection method were modified. In this experiment, the input data mass was classified into two categories, the estimation point and a neutral point. The estimation point was set to start at 50 frames after the image was presented to the subject and it ended at 50 frames thereafter. The displayed image that was used to determine the estimation point's start was that image that the subject evaluated at the highest level. It was labelled as "1". This change was in keeping with the concept of time, which we measured in terms of the length of time that the subjects needed to take action after having acquired the visual information. In the confusion detection method, the preset time was based on the mental operation of Card et al. (1980). However, a person's interest was thought to require a longer process time than that of a user's confusion because the processing of the former type of information is more complex. This complexity is attributable to individual taste, a concept that itself formed over a period of many years in the field of psychological research. However, some uncertainty still remains. Details of the learning algorithm are shown in Figure 7. The input data were inputted one frame aside from the time-series data. During the learning process, the learning rate was set at 0.001 and the neural network ceased learning when the cost function value reached 0.00001. The neural network learned each piece of data that was acquired in each task.

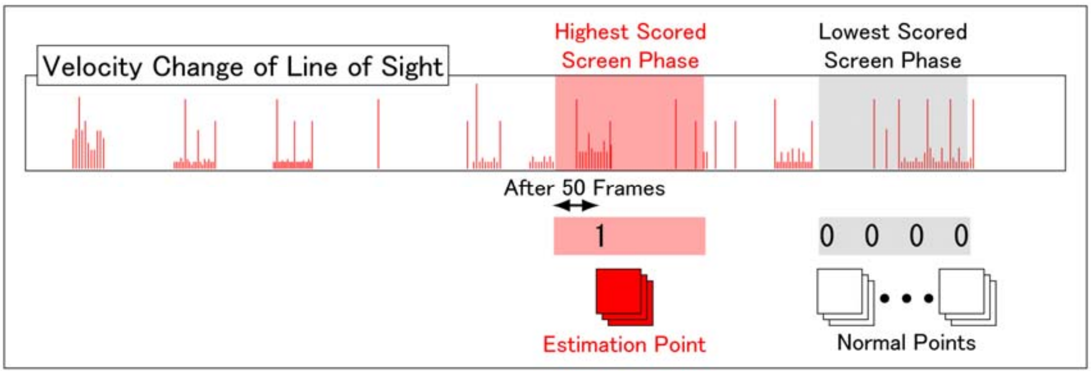


Figure 7. Learning Policy of the Experiment 4 Neural Networks

3.5 Estimation Results

As a result, the learning coalesced and a model to estimate a subject's interest was created. We confirmed the estimation ability by checking as to whether the model was able to estimate a subject's interest in a set of test data. In this study, we defined a condition of interest estimation as existing whenever the network output unit exceeded 0.5.

The estimation results of each model are shown in Table 7. In this verification, the data were not adopted as estimation objects, except for those data that were evaluated as the highest and lowest within a set of visual objects. For instance, when a subject accorded the highest valuation, 5 points, to visual object A, and accorded the lowest valuation, 1 point, to visual object B, and accorded 3 points to all the other visual objects, then, those visual objects accorded 3 points were not adopted as estimation objects. This chart reflects that, although the estimation point was only one per subject, all the data were merged into one.

	Estimation Rates [%]	
	Training Data	Testing Data
Task 7	72.0	64.9
Task 8	83.3	53.6
Task 9	100	83.9

Table 7. Estimation Results of the Neural Network

4. Discussion

This study showed that a method to detect a state of confusion, a human internal state, can be developed from eye movement data. This method demonstrated the importance of focusing upon a medium range history of eye movement data for approximately 50 frames.

In this paper, first we demonstrated the fact that a state of human confusion in simple HCI could be detected from eye movement data by focusing upon velocity changes within lines of sight as a parameter. This method used a learning neural network to employ eye movement data for pattern classification; those patterns became data labelled as confusion points and neutral points. The confusion points were those data acquired at the time that subjects encountered calculated traps embedded in the experiment's tasks and the neutral points were those data acquired before subjects encountered such traps. The learning of the neural network coalesced and a confusion detection model was created. The model created was successful in verifying the remaining test data that had not been used for the neural network learning. The verification results demonstrated that the detection rates were satisfactory.

Second, we demonstrated that the states of human confusion during realistic HCI could be detected in a similar manner. Thus, the learning policy was modified from the detection method used for simple HCI. The learning coalesced, and we were able to obtain a high detection rate. These detection rates were comparable to other previous studies that endeavoured to recognize basic emotions utilizing data obtained from contact-free devices measuring biological information (Dellaert et al., 1996; De Silva & Ng, 2000; Nicolson et al., 2000; Umemuro & Yamashita, 2003).

Finally, an additional experiment was conducted, one that employed older people as subjects to verify whether age group differences appeared within the features of the states of confusion. As a result of the neural network learning replicating the procedure used to create the first model, the learning coalesced. The second model created had a high detection rate for testing data along with the models that were made from the young subject data. Moreover, in verifications conducted by age group, the young subject data model successfully detected the older subject data, and vice versa. Consequentially, the detection rates obtained from verifying the young subject data model against the older subject data was on average comparatively high. In contrast, the detection rate obtained from verifying the older subject data model against the young subject data was not demonstrably high. We found it apparent that there were effective eye movement differences between these two generations. For instance, the saccadic eye movement is known to be affected by the aging process (Abel, Troost, & Dell'Osso, 1983; Warabi, Kase, & Kato, 1984). However, because the aging effect upon eye movement characteristics varied greatly from individual to individual, the effect could not simply be quantified as maintaining a strict barrier between generations (Warabi et al., 1984; Moschner & Baloh, 1994).

From these results, it became clear that our proposed method provided high detection rates with a learning neural network that used velocity changes in lines of sight as input data, alongside data acquired from various test situations. This illustrates the possibility that a single model, one that integrates a sub-network created for more than one situation, can detect various states of confusion. Although the issue of the age differences within this study still remains, we suspect that the problem lies not in the detection and algorithm but in the quality and quantity of our older test subject data. In addition, the knowledge gleaned in the verification by age group indicated the possibility that creating a single model could lead to data being detected without making any age or gender distinctions.

In order to establish a method for estimating human interest, we were able to create models displaying varying topical visual objects by employing almost the same algorithm as the one used in the confusion detection method. The human interest estimation models were also created with a learning neural network that used velocity changes in lines of sight as a parameter. However, the labelling process was modified based on the time-series differences pertinent to the internal states. In terms of a time-line, the state of confusion was thought to be the one next responsive in speed, following the state of situational reflection. In contrast, the time-line showing interest as an internal state was set at a longer time because humans require longer periods of time to make decisions. Therefore, we were able to demonstrate that human interest as an internal state can be estimated from eye movement data by focusing on the velocity changes in lines of sight. This method used a learning neural network to classify the labelled data of the estimation points and the neutral points according to patterns. The estimation points were those data measured at the time that the subject showed keen interest in the visual object and the neutral points were those data measured at the time that the subject exhibited absolutely no interest in the visual object. The neural network learning coalesced and the interest estimation models were thus created. As a result of those models' verification, these estimation models showed lower rates than the confusion detection rates were found to be in this study. The reason behind this is that there are differences in the fluctuating characteristics of each internal state, such as confusion and interest. For instance, in our research, confusion was a state that suddenly emerged from a normal situation. The difference between a normal and a confusing situation was clear and the difference likened to a digital waveform when the range of rise and fall of the state was delimited from "0" to "1". In contrast, a state of interest always emerged more or less in the form of an analog waveform. Although we were able to intuitively understand these features, no precise index capable of measuring the characteristic of suddenness in an internal state exists. This feature highlights the difference of ambiguity. Therefore, we took into consideration the fact that the interest estimation results were lower than the confusion detection results.

Using these results, we found that estimating human interest from eye movement data at a higher accuracy rate required changing the learning algorithm and neural network policy. However, the higher estimation accuracy rate might prove as difficult to reach as a higher accuracy rate in the confusion detection method. This is because the fluctuating characteristics vary greatly between human interest and confusion.

Finally, an assumption was made that it was important to pay attention to the fluctuating characteristics of each internal state in order for our proposed method to be applicable to other internal states. Moreover, if an aim was made to identify more stable characteristics of internal states other than confusion, e.g., interest or intention, the identification rate should be expected to be lower, without hoping for a near 100% rate. In the case of identifying other fluctuating characteristics, appropriate changes need to be made to the target identification rate.

5. Conclusion

The results of these experiments supported our hypothesis that the changes in a person's internal state are somehow reflected in his/her eye movement data. The investigation results demonstrated that human internal states could indeed be identified by using eye movement data. These results lead us to the conclusion; it is highly possible that the states of human confusion and human interest can be identified by employing a neural network learning algorithm. In such an algorithm, the velocity changes within lines of sight are adopted to handle time-series data more quantitatively as eye movement history. Taking this knowledge into consideration, in order to identify human internal states, a collection of eye movement data is needed in each particular situation in conjunction with the learning of a neural network based on such collected data. This work shows some indication of the possibilities that a model can identify an internal state without reference to age or gender in a particular situation.

6. Acknowledgements

This study was supported by the Softopia Japan Foundation in Japan. We wish to thank Yasunori Tanaka, Masashi Nakamura, and the Densan System Co., Ltd., as well as Dr. Minoru Sasaki and Dr. Yoshihiro Kaneko, both of Gifu University, for their many helpful suggestions during the course of this study.

7. References

Abel, L. A., Troost, B. T., & Dell'Osso, L. F. (1983). The effects of age on normal saccadic characteristics and their variability. *Vision Research, 23*, 33-37.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes?" Evidence from normal adults with autism or Asperger Syndrome. *Visual Cognition, 4,* 311-331.

Byrne, M. D., Anderson, J. R., Douglass, S., & Matessa, M. (1999). Eye Tracking the Visual Search of Click-down Menus. *Proceedings of the SIGCHI Conference of Human Factors in Computing Systems*, 402-409.

Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM, 23*, 396-410.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Dellaert, F., Polzin, T., & Waibel, A. (1996). Recognizing Emotion in Speech. *Proceedings of the 4th International Conference on Spoken Language Processing*, 1970-1973.

De Silva, L. C., & Ng, P. C. (2000). Bimodal Emotion Recognition. *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition 2000*, 332-335.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*, 509-540.

Ekman, P. F. (1992). Facial expression of emotion: New findings, new questions. *Psychological Science, 3*, 34-38.

Hara, F., & Kobayashi, H. (1996). A Face Robot Able to Recognize and Produce Facial Expression. *Proceedings of the International Conference on Intelligent Robots and Systems,* 1600-1607.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Science, 9*(4), 188-193.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science, 132,* 349-350.

Hess, E. H. (1965). Attitude and pupil size. *Scientific American, 212,* 46-54.

Ishino, K., & Hagiwara, M. (2003). A Feeling Estimation System Using a Simple Electroencephalograph, *Proceedings of the 2003 IEEE International Conference on Systems, Man and Cybernetics.* 4204-4209.

Izard, C. E. (1991). *The Psychology of Emotions.* N. Y.: Plenum Press.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*, 767-768.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*, 419-435.

Lim, H., Ishii, A., & Takanishi, A. (2000). Emotion Expression of a Biped Personal Robot. *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1*, 191-196.

Lyons, M., Akamatsu, S, Kamachi M., & Gyoba, J. (1998). Coding facial expressions with Gabor wavelets. *IEEE International Conference on Automatic Face and Gesture Recognition*, 200-205.

Mace, R. L. (1985). Universal design: Barrier free environments for everyone. *Designers West, 33* (1), 147-152.

Macrae, C. N., Hood, B. M., Miline, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me?: Eye gaze and person perception. *Psychological Science, 13,* 460-464.

Maglio, P. P., Barrett, R., Campbell, C. S., & Selker, C. (2000). SUITOR: An Attentive Information System. *Proceedings of the International Conference on Intelligent User Interfaces 2000*, 169-176.

Morris D. (1985). *Body Watching*. London: Equinox Ltd.

Moschner, C., & Baloh, R. W. (1994). Age-related changes in visual tracking. *Journal of Gerontology, 49*, M235-M238.

Musha, T., Terasaki, Y., Haque, H. A., & Ivanitsky, G. A. (1997). Feature extraction from EEGs associated with emotions. *Artificial Life Robotics, 1,* 15-19.

Nagamachi, M., Ito, K., & Tsuji, T. (1988). Image Technology Based on Knowledge Engineering and its Application to Design Consultation*. Proceedings of the 10th Ergonomics International, 88*, 72-74.

Nicolson, J., Takahashi, K., & Nakatsu, R. (2000). Emotion recognizing in speech using neural networks. *Neural Computing and Applications, 9* (4), 290-296.

Nielsen J. (1993). *Usability engineering*. San Diego, CA: Academic Press.

Norman, D. A., & Draper, S. W. (1986). *User centered system design: New perspective on human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.

Ohno, T. (2006). One-point calibration gaze tracking method, *Proceedings of the 2006 symposium on Eye Tracking Research & Applications*, 34.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge, U.K.: Cambridge University Press.

Rumelhart, D. E., McClellans, J. L., & PDP Research Group. (1988). *Parallel distributed processing*. Cambridge: The MIT Press.

Takahashi, K. (2004). Comparison of emotion recognition method from bio-potential signals. *The Japanese Journal of Ergonomics*, *40* (2), 90-98.

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1. The positive affect*, N.Y.: Springer.

Tomkins, S. S. (1963). *Affect, imagery, consciousness: Vol. 2. The negative affect*, N.Y.: Springer

Umemuro, H., & Yamashita, J. (2003). Detection of user's confusion and surprise based on pupil dilation. *The Japanese Journal of Ergonomics, 39* (4), 153-161.

Warabi, T., Kase, M., & Kato, T. (1984). Effect of aging on accuracy of visually guided saccadic eye movement. *Annals of Neurology*, 40, 462-469.

Yacoob, Y., & Davis, L. S. (1996). Recognizing human facial expressions from long image sequences using optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18* (6), 636-642.

Yamanobe, S., Taira, S., Morizono, T., & Kamio, T. (1990). Eye movement analysis system using computerized image recognition. *Archives of Otolaryngology and Head Neck Surgery, 116*, 338-341.

Zecca, M., Roccella, S., Carrozza, M. C., Miwa, H., Itoh, K., Cappiello, G., & Cabibihan, J., et al. (2004). On the Development of the Emotion Expression Humanoid Robot WE-4RII with RCHI-1. *Proceedings of the 4th IEEE/RAS International Conference on Humanoid Robots:, 1,* 235-252.

Zhu, Z., Fujimura, K., & Ji, Q. (2002). Real-time Eye Detection and Tracking Under Various Light Conditions. *Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, 139-144.

PsychNology Journal, 2007
Volume 5, Number 2, pp. 165 - 195

A First-time Wireless Internet Connection: More Than Just Clicking on a Link

Dimitri Voilmy ♠♣ and Karine Lan Hing Ting ♣

♣ GET / Ecole Nationale Supérieure des Télécommunications – Equipe Deixis – Sophia Antipolis (FRANCE)

ABSTRACT

In the context of understanding the particular use made of nomad Internet and mobile computing in its interactional dimension, this article examines the detail of a first-time connection to the university's Wi-Fi network. Through video ethnography, we analyse the collaborative *talk as work* between two participants in a public place and finely examine their use of artefacts and distributed information in the accomplishment of connection activities. Both their speech and their actions have been transcribed using the conventions of Conversation Analysis. We therefore follow the connection procedure step by step and demonstrate how handling computerized artefacts is not transparent and requires a certain degree of learning concerning this particular communicative and working tool.

Keywords: *Video ethnography, situated action, human computer interactions, cooperative activities, nomad Internet, ubiquitous computing.*

Paper Received 30/10/2006; received in revised form 09/06/2007; accepted 07/08/2007.

1. Introduction

From our ethnographic observations on the campuses, we have noticed that personal mobile equipment has been increasingly appearing in lecture theatres and teaching rooms in French public universities. Gradually, the campus has become a particularly good place for ubiquitous computing (Weiser, 1998), where we have been able to observe the taking place of heterogeneous nomadic and mobile activities using IT equipment. A campus environment with new practices is therefore being created and these new technological practices are closely observed *in situ*. It is therefore important to study members' practices as such, which at first may seem "essentially uninteresting" both for members and for the analysts. Video recording is a practical way of obtaining data, while preventing the problem of "the invisibility of common

♠ Corresponding Author:
Dimitri Voilmy
ENST, département SES, équipe Deixis, Institut Eurecom,
2229 route desCrêtes – BP 193, 06904 Sophia Antipolis,
Phone: +33493008416
E-mail: voilmy@enst.fr

sense procedures" (Have, 2002). The spatialization of a place emerges in and through the practices, where the spaces are made intelligible in terms of occupancy and patterns of hidden and potential familiarity (Williams & Dourish, 2006). Our understanding of this place is that rather than being an object or container, this place is involved in a process where it is continually being redefined by the usages (*ibidem*). Thus, only empirical detailed observation and precise analytic characterisation of the field can reveal the relevant phenomena.

Supported by video ethnography, our participant observation approach follows a five-month fieldwork observation of the campuses[1], where one of the present authors, *Bg*, was one of the actors in the accomplishment of a first-time connection. Being a member of the university, *Bg* has been "doing being" (Garfinkel, 1967; Heritage, 1984) both an ethnographer and a student-user of the Wi-Fi access, as he shares the same background knowledge about the day-to-day activities at the university. *Bg*'s participation is treated as an interactional event. We will question the relevance of the membership category (Sacks, 1972) of *Bg* related to the participants' orientation to one category or the other, and how this category is consequentially procedural to the development of the activity. We will also investigate the temporality and sequential organisation of the participants' speech and on-screen activities as well as their interactional and corporal orientation to the informative artefact. Both ethnographer (*Bg*) and first-time network user student (*Ji*) are seated side-by-side and facing the laptop screen. By examining the interaction's structural organisation, we focus on how the "collaboration-oriented" achievement of a first-time connection during the ethnographic observation, is rendered salient and relevant in the activity. How does the "collaborative" nature of this talk-in-interaction emerge? Is it as a matter of the contextual definition of the activity in which they are engaged: an observed first-time connection? Does it emerge by the participants' orientation to this specific talk and the special practices they employ to co-construct their conversation (Schegloff, 2002), for example an incipient state of talk during the activity of connection?

[1] We would like to thank Professor Chantal Charnet for having permitted our access to the research field of the campuses of the universities of Montpellier 2 and Montpellier 3 and for having obtained us the necessary authorisations for video recordings on the universities' premises for our five-month ethnographic observations. We also thank Marc Relieu for having raised relevant questions.

Figure 1. The entrance hall of the university's administrative building

The data presented in this article were video recorded[2] in an entrance hall. Although our corpus of data is composed of various types of activities (work, leisure or communicative) on the computer, we voluntarily limited our analysis to the uniqueness of this particular activity of a first-time connection. Through the deployment of situated practices articulated within ongoing processes of human interaction (Goodwin, 1994), we are able to analyse the "relationship between technology use and a local cultural practice" (Barkhuus & Dourish, 2004), that of getting connected and surfing the Internet. The unique nature of this activity focuses the analysis on the accomplishment of the specific steps of connection.

We therefore consider this situation of connection from a *flexible* procedurally consequential point of view (Schegloff, 1991) when, in the development of the activity and the talk-in-interaction, the participants orient to the distinctive features of the network-covered area. In this way, we avoid the "bucket theory" (Drew & Heritage, 1992) where the interactions are determined by their occurring in this space. Rather, the specific characteristics of the Internet space become relevant in the *situated* activity: the framework does not determine the type of interaction that will be produced; it only becomes relevant for our analysis, through the talk-in-interaction and in terms of the participants' orientation to it, and in relation to the accomplishment of the connection activities. We thus favour a situated approach where we centre the analysis on *how* the participants create, assemble, produce and reproduce the social structures to which they orient themselves, in their familiarity with, or adoption of, novel technologies, such as using the Wi-Fi network on the campus for the first time. Through the analysis of one single interpersonal collaborative interaction in the Wi-Fi covered entrance hall, we aim at examining how these two participants use physical artefacts (student card, computer screen) and affordances (Gibson, 1979; Conein & Jacopin, 1993; Conein, 2006) of the laptop's keyboard during *Ji*'s first connection to

---

[2] We would like to warmly thank Myriam for her kindness in collaborating.

the wireless network. The accomplishment of this first-time connection procedure allows us to grasp the difficulties and the specific usage made of mobile Internet.

Since using the Wi-Fi network depends on the student's identification and successful connection, our study focuses on the accomplishment (Garfinkel, 1967) of the procedures required to access the network, and more specifically on how the state of "knowing how to connect oneself" becomes relevant. Beyond the personal identification capabilities of personal mobile devices, we look into the collaborative knowledge work between actors, made accessible in their talk-in-interaction and their orientation to the activity. Above all, we question the difference between this connection and the processes of establishing a 'traditional' cabled connection, which is for *Ji* 'unremarkable', 'unobtrusive' and 'commonplace'. Do the differences in the connection procedures constitute an evolution, and is this evolution a characteristic of mobility and a specificity of the nomad internet? How should the difficulties in this first-time connection as regards to the characteristics associated to ubiquitous computing be considered in terms of a transparent interface and simplicity of utilisation? Weiser (1991) proposed the futuristic notion of a truly ubiquitous computing which would eclipse the "PC era" (Weiser 1998); the users would be less dependent upon the computer technology, and where it would be woven into the fabric of everyday life (Crabtree, Evans, Fraser, Tolmie, & McMullen, 2005). The first harbingers of ubiquitous computing are the small handheld computers – the laptops – which are starting to be widespread on the campus (Weiser 1998). Thus, nowadays, how far can we consider the use of mobile internet and laptops as being part of an augmented space? We also aim at examining in what manner and to what degree the new parameters brought along by mobile computing influence the way university members define and consider what is a 'working area' or a 'public space'. The term tends to become problematic for users because of the newly hybrid nature of this public space, this socially-organised setting (Williams, Kabish, & Dourish, 2005).

2. Local Circumstances of Practice

This practical action of establishing a first-time connection can be broken down for analysis into four distinct and successive steps, each of which has been transcribed in detail (extracts 1 to 8). These steps are: 1. the initial failure in the connection; 2. an evaluation of the situation with regard to previous practices; 3. recognition of the network by the laptop and 4. the student's identification on the university's intranet. We focus on the temporal development of the activity, that is, from the moment the participants orient to the potential use to be made of the Internet once they enter the network-covered area, until they are connected and actually surf the Internet. The analysis of the following extracts, made up of the connection steps and accessing the Internet, empirically questions the types of activities that take place in this Wi-Fi network space.

The activities linked to the use of the Wi-Fi Internet render this entrance hall an unusual place, both in terms of socialisation between the passers-by and the laptop users, or among the laptop users; and in terms of the specific, new use which is made of this space. This corridor, initially designed as a space for transient visitors to pass through, has been transformed and furnished, and can also, therefore, welcome students for didactic activities, since they have access to the university's pedagogical intranet. Nevertheless, by definition, this corridor and information space is noisy, and conversations among different groups of interactants who use the entrance hall constitute a constant but irregular background sound, even if at certain hours, the hall is much calmer, e.g early in the morning or during lunch time. We would therefore not expect working activities in such a place, and are more likely to observe them in the library.

2.1 Establishing an Ecological Framework

From the contextual information (Have, 2002) she gave us, we know that *Ji* is used to connecting to the Internet via the cabled network of the student residences where she lives. In her room, she has a broadband connection through a wired connection to her laptop.

In this sequence, the conversation between the two participants in action deals with the difference in the procedures involved in the connection between the cabled network at home and the wireless network at the university, where they are presently located. Most laptops, for instance, can connect through many different network set-ups, in the office, home, or on the road, and can move between different environments and network infrastructures (Dourish, 2004). The connection configuration is *a priori* not the same depending on whether one is using a desktop computer or a laptop, or, in the case of *Ji*, whether the connection is a cabled or wireless one depending on the locality. Therefore, the procedures and the practices linked to *Ji*'s Internet connection change. How can the characteristics of "ubiquitous computing" be applied to these two methods of connection?

In this extract, the two participants are sitting at a table in the entrance hall, and are preparing to establish a first-time connection. *Bg* is setting up the camera when he asks this question, while *Ji* is switching on her laptop. It is their first verbal exchange concerning the connection issue. *Bg* asks for an information concerning *Ji*'s usual use of her laptop. By her negative response "∇∇>no.<∇∇" (extract 1 – line 3) to *Bg*'s question "Δ>have you already? surfed on the internet↑<Δ (.) here?" (extract 1 – line 2), the 'first time' nature of this connection is made explicit as the participants orient to it.

1		(0.1)
2	Bg	Δ>tu es déjà? allée sur internet↑<Δ (.) ici?
		Δ>have you already? surfed on the internet↑<Δ (.) here?
3	Ji	∇∇>nan.<∇∇
		∇∇>no.<∇∇
4		(0.4)

Extract 1. Question-answer sequence concerning past experience

This question-answer sequence is not directly linked to the procedure needed for the connection, but deals with the very nature of the research activity itself. This pre-sequence constitutes the opening of the interaction once they are seated side by side at the table. It is anchored in a spatialized connection activity; the categories "ethnographer – laptop user" are accomplished by the reflexive articulation of a positioning on an observation activity which starts with a question (Relieu, 1999). This talk-in-interaction collaboratively sets up the ecology of the activity of connection to be observed, towards which the participants orient themselves. This conversation is organized as one in a series, or more specifically, as the first conversation in this observation activity, and is used by the participants to structure some parts of their conduct (Button, 1991). This talk-in-interaction - situated at the beginning of the activity of getting connected - reveals their orientation to the multi-level activity and the coordination of interaction both for the connection and the research observation, as well as their knowledge of standard connection procedures.

2.2 First Failure in Connection

Ji's first attempt here to obtain a connection is not successful. The result of this attempt is a message displayed on the screen: "`aucun réseau sans fil n'a été détecté à portée`", which means that no wireless network has been found. From our observations during our fieldwork, this type of error is common when students attempt to connect for the first time. This is probably due to the fact that a first connection implies several successive steps, which some students have difficulty in accomplishing. As it has often been described, the manipulation of technological tools (Suchman, 1987; Dourish, Grinter, Delgado de la Flor, & Joseph, 2004; Norman, 1998) or more specifically the interface of digital information (Ishii, 2004) requires a certain degree of learning and knowledge acquisition concerning this specific communicative and working tool.

In the extract transcribed below, the two participants try to obtain wireless access from the laptop, which is not the procedure to which *Ji* is used – a cabled connection from the student's residences. However, *Ji* knows that *Bg* has already observed and recorded a first-time connection with another student, and interestingly *Ji* exploits *Bg*'s experience as a resource.

1		Ji	∆ah mince::↑∆ >qu'est ce que **∆oh sugar::↑∆ >what's**	She reads the on-screen text (lines 1–3)
2			c'est?< ces mots? arrêté. **that?< these words? stopped.**	
3			((during 8.0sec Ji starts her software application))	
4		Ji	ah bé? y veut pas? **oh but? it doesn't want to?**	
5	→		il a fait que ça↑ l'autre **did he only do that↑ the other**	
6			jour. (0.1) >ou il a fait? aute **day. (0.1) >or did he do? another**	Ji turns her head towards Bg and looks at him (lines 6–11)
7			cho:se.< **thi:ng.<**	
8		Bg	pour s connecter↑ **to get connected↑**	
9		Ji	∆ouais↑∆ **∆yeah↑∆**	
10		Bg	non il faut chercher le réseau? **no you must search for the network?**	
11			d'abord. **first.**	

12		Ji	et tu cherches? euh:::	She reads the on-screen text
			and you search? uhm:::	
13	→		(0.3) euh::, (0.3) Δça↑Δ (.)	
			(0.3) uhm::, (0.3) Δthat↑Δ (.)	
14			>eh oui? c'est ça↑<	
			>oh yes? that's it↑<	
15		Bg	ouais?	
			yeah?	
16			(0.3)	
17	→	Ji	aucun réseau sans fil n'a été	She reads the on-screen text
			no wireless network has been	
18			détect- détecté↑ à portée↑	
			detect- detected↑ within reach↑	
19			(0.2)	
20		Bg	Δaucun: réseau↑Δ	
			Δno: network↑Δ	
21		Ji	Δnon,Δ	
			Δno,Δ	
22			(4.0)	
23	→	Ji	actualiser? ∇la liste? peut être,∇	
			refresh? ∇the list? maybe,∇	

Extract 2. No wireless network has been detected

By her interjectional phrase "**Δoh sugar::↑Δ**" (line 1), the actor *Ji* characterises the fact that a problem has occurred. This interjection shows a double orientation to the sequential environment of the activity. It is both 'context-shaped' (the interjection is produced in relation to the message that has just appeared on the screen) and 'context-renewing', being given that this same action is part of the current sequential environment in sight of, and for the next action (Wilson, 1991). Following this interjection, her question refers to elements which have appeared on the screen: the words "arrêté". After this question, there is *Ji*'s activity on the computer where she starts her software application, which lasts 8 seconds. She answers herself "**oh but? it doesn't want to**" and asks another question. Whereas in the first question, *Ji* was talking to herself, this second question (line 5-7) is addressed to *Bg*. *Bg* is not only the ethnographer, *Ji* orients to *Bg* as being a resource for the activity and therefore a legitimate participant to the active accomplishment of the task at hand. In her turn, "**did he only do that↑ the other day. (-) >or did he do? another thi:ng.<**" (extract 2 – lines 5 to 7), *Ji* is referring to her current activity on the computer and the results of it by using the deictic pronoun "**that↑**", while, by referring to "**the other day.**", makes *Bg* understand that the question is addressed to him, as it mobilises his previous experience*.* Previously, she had designated the absent student whose connection *Bg* had observed as "**he**" (lines 5-6) and the computer as "**it**" (line 4). In French, the same pronoun is used for animate and inanimate: the pronoun "il" ("y" when truncated). By using this pronoun, *Ji* personalises the artefact in the interaction (Suchman, 1987) and gives the laptop an orientation of its own. *Bg* asks for a confirmation of the meaning (line 8) by expanding *Ji*'s question and *Ji* answers positively. *Bg*'s category evolves from being a hearer of *Ji's* activity to an active speaker, as the activity becomes more and more conversational (Schegloff, 1996). After this inserted sequence, *Bg* answers the previous turn and verbally gives *Ji* the instructions to follow "**no you must search for the network? first.**".

After the conversational sequence (line 5 to 11) where *Ji* was looking at *Bg*, the talk-in-interaction which follows, far from being *topic talk* (Sacks, 1992) has a more practical goal to which both participants orient themselves, as they turn towards the screen: that of succeeding in connecting the laptop to the Wi-Fi network. By the demonstrative pronoun "**that↑**" that she uses twice (extract 3 – lines 13 and 14), *Ji* deictically refers to a window that has appeared on the screen.

12 Ji (0.3) euh::, (0.3) ∆ça↑∆ (.)
(0.3) uhm::, (0.3) ∆that↑∆ (.)

13 >eh oui? c'est ça↑<
>oh yes? that's it↑<

14 Bg ouais?
yeah?

Extract 3. A window appears on the screen

After her hesitation, she is still unable to name the object of her turn, and finally utters "**that↑**". Without seeing the video of the screen or as a simple overhearer of this conversation, one would be incapable of understanding the object of "**that↑**"; here *Bg* does (extract 3 – line 14). The sense of this deictic is clear for both participants, who are engaged in this situated activity and are looking at the same screen. The participants make use of all the resources at hand: the talk between the co-interactants, the on-screen text (error messages or instructions) that they read, the keyboard and trackpad that *Ji* manipulates all mutually inform each other within a single coherent activity (Goodwin, 1994). The interactants use an interactive construction of turns at talk to achieve courses of collaborative action, which is composed of both talk-in-interaction and gestures. The hesitation in *Ji*'s turn "**(0.3) uhm::, (0.3) ∆that↑∆**" is a form of delaying the utterance deictic as the object she is referring to has not become part of the scene yet (Licoppe & Relieu, 2005; Mondada, 2006). Another evidence of their common orientation is *Bg*'s confirmation "**yeah?**" (line 15) after the adjency pair of *Ji*'s question and answer: they are aligned on the same issue and the same task.

In the same way that she shows the different procedures and their mutual engagement in the attempt of connection, *Ji* shows their failure too. She reads the message aloud ("**no wireless network has been detect- detected↑ within reach↑**" - line 17 and 18) – which *Bg* could have read on the screen, by himself, without the need of the utterance. *Ji* confirms the failure through the second pair part, her answer "**∆no,∆**" (line 21). In this connection activity, *Bg*'s category in action evolves progressively from ethnographer to active participant in the connection activity. It is clear that *Bg*'s Membership Categorisation is not defined *a priori* and once and for all by the ethnographic observation context. Rather, his category becomes relevant in the course of the dynamic activity, as the participants themselves negotiate the relevance and orient to either category. The relevance of "being more than just the ethnographer" becomes essential for the development of the activity. *Bg* becomes an active member, the flexible cooperative modalities linked to the participants' co-presence give some kind of robustness to the collaborative team's work, thanks to different types of informal help, that, for example, experts can bring to lay persons (Grosjean, 2005). This is the nature of the help that *Bg* is giving *Ji* through their *talk as work* (Whalen & Zimmerman, 2005), and more particularly his explicit instructions (line 10) or questions (line 20). As he engages more and more in the activity as a participant, *Bg*'s contributions guide the activity and guide the course of the action. After a four-second pause in the talk-in-interaction, *Ji* produces, with a continuative tone, the last turn of the extract "**refresh? ∇the list? may be,∇**" (extract 2 – line 23). We notice a lower volume at the end of her utterance, as she suggests a possible action to remedy the failure.

2.3. A Gap between Online and Offline Experience

The interactants make a second attempt to connect the laptop to the wireless Internet network in the next extract transcribed below. Unlike the previous attempt, where the difficulty was centred around the state of being "online or offline", that is part of the "embodied" virtual space, it seems that, as the activity develops, the problem is extending to the real world. Wireless networking technologies are perhaps especially interesting due to their combination of tangible and intangible elements (Dourish and al., 2004). Caught up in the everyday usage of an Internet-connected computer, and using the possibilities of ubiquitous computing in a transparent way, the users sometimes forget this *sine qua non* state of connection. Since the wireless networks' interfaces offer the same properties and usage characteristics as wired ones, *Bg* tends to get confused in this extract.

But before coming to this issue, we will take a close look at the modality of the interaction between *Ji* and *Bg* in order to analyse the mutual and aligned sharing of information in the accomplishing of the activity. What contextual information does each participant effectively contribute to their accomplishment of the cooperative activities?

1		(0.4)
2	Ji	Δça? peut ête. que là↑ dfaçon?Δ
		Δit? can only be here↑ nyway?Δ
3	Bg	ou:[ais,
		ye: [ah,
4	Ji	[>afficher? [toutes les&
		[>display? [all the&
5	Bg	[Δouais, ouaisΔ on&
		[Δyeah, yeahΔ we're&
6	Ji	&connections que t'as fait.<
		&connections that you've done.<
7	Bg	&va faire ça? >afficher toutes les
		&going to do that? >display all the
8		connections::< (.) hhh euh:: hmmm::
		connections::< (.) hhh uhm:: hmmm::

Extract 4. Information sharing during the second attempt of connection

This is the first time overlapping occurs in the collaborative talk-in-interaction between *Ji* and *Bg*. These overlaps can be characterised as "supportive and cooperative interruptions" (Lerner, 2002), where the interactants collaboratively produce Turn Constructional Units (TCU) (Sacks, Schegloff, & Jefferson, 1974). The overlapping begins at *Bg*'s appreciative response (*ibidem*) at line 3, which displays *Bg*'s understanding of the deictic pronouns "**∆it?**" and "**here**" at line 2. *Ji*'s turn (extract 4 – line 4) overlaps *Bg*'s (line 3), which is in turn overlapped at line 5. These distinct forms of participation reveal their common orientation to the screen activity and the on-screen displays. *Bg*'s overlap of *Ji*'s turn, which has just begun and where the TCU has not yet been uttered, displays his understanding of the situation. Both participants have access to the screen resources and understand, at the same time, which procedures need to be accomplished, as they share the same "collaborative floor" (Lerner, 2002). Their overlaps show that they both acknowledge what's going on in terms of action and sequences of actions. *Bg* does not wait for the end of *Ji*'s utterance before starting to treat it: he follows the emergence progressively (Goodwin, M.H., 1980; Goodwin, C. & Goodwin, M.H., 1987). *Bg* pays attention not only to the talk itself but also to the surrounding environment (Goodwin, C. & Goodwin, M.H., 1996). The environment contains the relevant elements for the task in hand, and more particularly the relevant informational elements necessary for decision making.

9	Bg	•hhh (0.3) hhh ∆tu connais∆ le:
		•hhh (0.3) hhh ∆do you know∆ the:
10		euh le site de la fac↑, >eh beh non
		uhm the university's website↑, >eh no
11		on peut pas y aller< tant qu'on est
		we can't go there< as long as we're not
12		connecté,=
		connected,=
		#((dismay hand gesture))
13	Ji	=eh #ou:ais,
		=oh #ye:ah,
14	Bg	(0.2) euh::, (0.5)
		(0.2) uhm::, (0.5)
15		∇conne::xion?∇ quand tu >quand
		∇conne::ction?∇ when you >when
16		tu es?< quand tu es? à ta cité euh
		you're?< when you're? at your student's uh
17		universitaire↑ tu fais comment.=

residences↑ how do you do it.=

18 Ji =moi c'est là↑ >mais c'est parce

=for me it's here↑ >but it's be-

19 que< c'est branché↑

-cause< it's plugged in↑

20 (.)

21 Ji [>parce que là il est pas branché<]

[>cause here it is not plugged in<]

22 Bg [∇ah oui::∇]

[∇oh yes::∇]

23 (.)

24 Bg là c'est branché↑ directement

there it's plugged in↑ directly

25 ∆avec un [câble↑∆=

∆with a [cable↑∆=

((Ji's head nod))

26 Ji [hum:

[hum:

27 Bg =tu fais pas wifi. tda↑ (--) euh::

=you don't use wifi. tda↑ (--) uh::

28 hum::et? si tu fais? rechercher::?,

uhm::and? if you do? search::?,

29 créer-? ∇non pas créer une

create-? ∇no not create a

30 nouvelle connexion:?, ∇ (0.2) euh::

new connection:?, ∇ (0.2) uhm::

31 (0.2)∇tda euh∇

(0.2)∇tda uhm∇

32 Ji ∆et? C'est quoi?∆ qui nous

∆and? what is it?∆ that

33 bloque::↑ euh

blocks:: us↑ uhm

34 Bg ∇∇je me souviens plus comment y

∇∇i don't remember how we

faut faire∇∇

must do∇∇

Extract 5. the multimodal interaction between *Ji* and *Bg*

In lines 9 and 10, *Bg* asks her interlocutress if she knows the university's website, "**∆do you know∆ the: uhm the university's website↑,**". Considering this question in the context of the activity, this pre-sequence allows an 'opening' on the projected activity. *Ji* as well as the analysts are able to understand that *Bg* is referring to the instructions concerning the connection instruction that are available on the website. The speaker *Bg* seems to orient to the fact that it is difficult to reconcile a manifestation of being "offline" (the laptop not being connected yet) with an "online" experience (visiting the university's website). Without the instructions to help them in their attempt, they cannot connect the laptop and since the laptop is not connected, they are unable to obtain these instructions. *Bg* himself realises the paradox: after having uttered the question, he produces an explicit explanation (extract 5 – lines 10 to 12). By her verbal confirmation, accompanied by the left hand gesture at line 13, *Ji* makes visible that she has seen the absurdity too. This visible and recognised contradiction, to which they orient themselves, points to the role of the Internet as a medium for a communicative and informative artefact. But one of the main characteristics of this artefact, as we have seen in this extract, remains embodiedly virtual and cannot be accessed until a connection is established and gives the Internet its full capabilities.

By elaborating on this issue, which is the difference between a "traditional", or at least usual, cabled connection and a wireless one, *Ji* and *Bg* makes the characteristics of the wireless network relevant for the current activity. This specificity of the Wi-Fi connection, as opposed to a cabled one, is made relevant in the practical development of the activity: it is more difficult to connect to the wireless network. These connection procedures constitute, in the initial steps of the connection activity, a problem for the participants themselves, because there are more actions to accomplish and these procedures are still unknown to *Ji*, who is in the process of learning them. At line 18, *Ji* verbally refers to her usual manipulation and orients herself to the difference between her usual routine procedures and the present procedures needed to obtain a wireless connection that she is experiencing for the first time in this open space. At lines 17 and 18, *Ji* and *Bg*'s respective turns follow each other rapidly, without any interruption. In coordination with her utterance, *Ji* produces a pointing gesture with her right hand (line

18), before the lexical component to which the gesture is affiliated (Schegloff, 1984). This gesture is used to show to *Bg* the window that has just popped up on the screen. The cabled connection parameters, which have previously been saved in the operating system, are written in this window. *Ji* refers to the distributed resource (Hutchins, 1995; Conein & Jacopin, 1993) concerning the computerized artefact in order to answer the question. In general, the computerized artefact which helps the users plays the role of a cognitive tool, that is usually non intrusive and which remains under the control of the human operator (Salembier & Pavard, 2003).

Ji relies on the knowledge shared by *Bg* and herself in the course of the interactional activity: she makes the situation explicit when answering *Bg*'s question about her usual procedures of connection (extract 5 – line 17). As the activity unfolds, *Ji* mobilises her knowledge of "what she usually does" as a resource for the current action, as the on-screen text becomes relevant. After uttering her answer, she continues to manipulate her laptop: she keeps on looking at the screen while she activates her mouse cursor. At line 26, her answer **"[hum:"**, which overlaps *Bg*'s turn, is accompanied by a head nod (Goodwin, M.H., 1980; Goodwin, C., 1987). The head nod becomes a pertinent resource, which is available and made visible within the activity (Goodwin, 1981), placing, for the analysis, the indexical proprieties of the practical action on the foreground. Some studies (Heath, 2002; Goodwin, 1981) have legitimated the gaze and vocal elements as essential vehicles for the production of actions and social activities in the socialisation process. In this particular collaborative interaction, which occurs in a public place, and where the co-present participants are seated side by side and orienting themselves to the same on-screen activity, the production of a gestural manifestation, coupled to the verbal **"[hum:"**, appears to be the appropriate action for the current activity. Usually the speakers use the visual channel to encourage the partner in producing a gaze readjustment or a series of head nods showing their engagement in the current activity (Heath, 1989).

We have just seen that the technical object "a cable" served as a medium for *Ji*'s connection to the Internet: without its functionality and the appropriate parameters, the usual home connection is not made possible. In the next step of connection, we will examine further the characteristics of Wi-Fi network, which is essentially virtual. Therefore, within the network-covered space, the interactants are engaged in the accomplishment of a new and different procedure, in order to have access, first to the Internet, and finally to the shared workspace.

2.4. Connection Button's Affordance

In the following extract, the users familiarize themselves with the virtual workspace after accomplishing the prior actions needed for connection, which have to be learned, through successive compulsory steps. In this extract, the co-participants finally succeed: the laptop accepts the wireless connection. It is possible to question here the particular hybrid nature of the open space, where the university's Wi-Fi space merges with a public place. One must remember that *Ji* and *Bg* find themselves in an open and public corridor, where different categories of university members constantly pass along. The notice board which they are facing is consulted by two students, *Pm* and *Rl,* who want to know about the list of the appointees for the exam. The two girls are discussing and are engaged in their own activities. They are standing next to *Ji* and *Bg* at the very moment they get connected to the Internet. Thus, students surfing on the Internet from their laptop and students consulting the notice board, passing by or anyone having a drink meet in this same space while they are engaged in quite different forms of activities. For example, *Ji* and *Bg* (try to) use the Wi-Fi network available within the public space, while the two students reading information on the notice board do not orient themselves to the possibilities of the Wi-Fi network at all. Yet, the four of them find themselves in the same restricted area, engaged in activities that are expected to happen there, and rub shoulders without treating the other persons' activities as disturbing. As their activities and topic talk skirt without merging, their separate conversations have been transcribed in parallel columns while preserving the temporality: *Ji* and *Bg*'s conversation on the left and *Pm* and *Rl*'s talk on the right. We can therefore observe that the organisation of the wireless network functions *within* a public space, as well as it is made available *for* the public space. This feature of the open space is directly linked to the mobility associated to it, where Internet users are nomads and can appropriate all kinds of places (including public places) if they are Wi-Fi zones.

1 (0.4)
2 Ji >et dans l'aide< de ce truc? y::=
>and in the help menu< of this thing? there's::=
3 Bg =<et pourquoi? y:, y met activé,>
=<and why? he:, does he put activated,>
4 (0.5)

5 Pm ça↑ c'est tous les noms? de tous
this↑ that's all the names? of all
6 les délégués. (0.3) t'as vu: Δon
the delegates. (0.3) did you see: Δwe
7 commence par le a↑Δ hihi hhh
start by the a↑Δ hihi hhh
8 £y doit ête là bas↑£
£he must be there↑£
9 Rl moi? ∇j'en ai rien à fou:tre du a.∇
me? ∇i don't give a damn about the a.∇
10 (6.0)

11 Bg les boutons sans fil, situés
the wireless button, situated
12 sur l'ordinateur? (.) °qu'est? ce
on the computer? (.) °what?
13 que c'est que ça,°
is that,°
14 (8.0)

15 Pm ∇regarde.∇ (0.3) £y a tout ça:?
∇look.∇ (0.3) £there's all that:?
16 comme djouba£ [(inaudible)&
like djouba£ [(inaudible)&
17 Rl [ΔΔhi hmm↑↑ hmmΔΔ
18 Pm &(inaudible)

Physical
affordance
of the button

19 Ji &Δah bé↑ c'est c truc↑↑ no:n↑Δ=

&∆oh but↑ it's this thing↑↑ no:↑∆=

20 Bg =ah bé voilà↑ [hrrr

=oh that's it↑ [hrrr

21 Ji [hrrr↑ hhh

22 ∆∆heuheuheuheu [heuheuheu↑↑∆∆&

23 Bg hhh [∆merci? ∆ (-)&

hhh [∆thanks? ∆ (-)&

24 Ji &hhh hrrr ah? putain::↑ heuheuheu

&hhh hrrr oh? shit::↑ heuheuheu

25 Bg & merci? jiji?

& thanks? jiji?

26 Ji [heu hhh £j'm'étais toujours

[uhm hhh £i had always

27 Bg [et voilà

[and here you go

28 Ji demandé ce que c'était ce bouton£

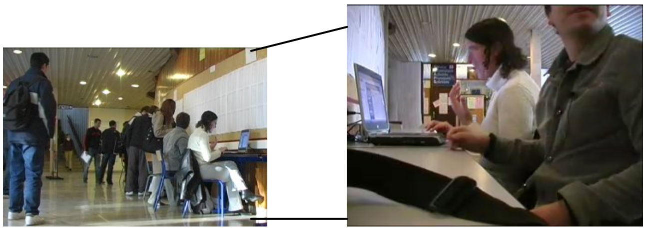
wondered what this switch was for

29 heuheu

30 (.)((physical contact between Rl and Bg. Bg looks at her and moves his chair))

31 Bg par?don

so?rry



Extract 6: the connection to the university's Wi-Fi network mixes up with the public space

Since the public areas, which are Wi-Fi zones, have evolved and have been adjusted and furnished in order to meet the new needs and use of space, it is also interesting to question the way in which different types of activities are accomplished. Behind *Ji* and *Bg* (who are sitting side-by-side facing the laptop), there is a corridor, which is an open public space that students and administrative staff constantly pass through. The wall in front of them serves as a notice board, where students can obtain information, like the two girls we will see later who consult the list of the appointees for an exam. A coffee machine and a drinks machine are also available in the entrance hall. The two campuses that we have observed are public places. We observed that they are shared spaces as well as workplaces of a particular nature: the students are engaged in personal work tasks, sometimes in collaboration with other students, or in leisure activities. This connection activity therefore occurs in a place that can be categorised as “public”, where particular types of interaction are also produced. But beyond what “appears to be” at the same time a work and leisure area, how, precisely, do these places progressively emerge as an exclusive context of interaction, giving rise to verbal exchanges and practices that may be characteristic of a Wi-Fi zone on the campus?

At line 5, *Pm* is addressing Rl. The two speaker pairs *Ji-Bg* and *Pm-Rl* are in such close proximity that the conversation between *Pm* and *Rl* is audible on the video recording. The conversation between *Pm* and *Rl* is available for analysis because of the *continuing state of incipient talk* (Schegloff and Sacks, 1973) between *Ji* and *Bg.* These gaps of silence (as from line 4) are not occasions to close the conversation, but a silence set by the contingencies and constraints of the screen activity in which they are primarily engaged. As a result, the collaborative talk-in-interaction is suspended when they are accomplishing some individual tasks, like reading. Moreover, we can even interpret these silences as uninterrupted action. In these different extracts, the talk-in-interaction has a collaborative goal. Most of the time, *Ji* and *Bg* share information and try to find a solution when there is difficulty in accomplishing the activity.[3]

In the same way, the *body torque* allows us to structure and make sense of the two parallel conversations in context. The bodily orientations of the respective groups (see images at line 4) inform us about their relationships and the nature of their activity. Therefore, an utterance or a bodily movement has to be analysed in relation to the local context of its production in order to grasp its full meaning; all the more in such environments where different types of activities occur in the same place. While the first group sitting at the table is engaged in their main activity of obtaining a wireless laptop connection to the Internet, the second group beside them, who are standing up and moving slightly, are looking for names of students on the posted lists. *Ji* does not orient to the potential “disturbance” (she continues to face her laptop) and despite the corporal proximity and the parallel conversations, she continues to explore the configuration of her laptop, with *Bg*’s help. *Ji* does not seem to consider the girls’ conversation as a nuisance because there are no contextual features of privacy relative to a workplace as such. It is worth noting that *Bg* and *Ji* are simultaneously in

[3] Therefore, this empirical analysis of action is not only confined to the analysis of interpersonal communication (Conein, 1990: 103).

two places (Schegloff, 2002): the real physical environment, the entrance hall, and also the more "embodied virtual space" (*ibidem)* where they are trying to connect the laptop to the Wi-Fi network. The main characteristic of this place is that activities which traditionally take place on desktop computers via wired Internet connections, which are placed in computer labs or libraries at the university, are transferred to open places where the Wi-Fi network can be picked up. They are hybrid places where heterogeneous activities can, and are carried out as seen here. As a result, *Ji* and *Bg*'s activities on the laptop can, of course, be analysed from a generic Human-Computer Interaction, where, once the users obtain an Internet connection, the procedures do not really differ whether the users are mobile or not. However, in sequences like here, where students using their laptop are close to others who are engaged in completely different activities, the relevance of the specificity of using a Wi-Fi network from a laptop on the university premises emerges in and through the activity, and is analysable through the talk-in-interaction.

The long pauses (six seconds at line 6, and eight seconds at line 14) in between their different utterances show *Ji* and *Bg*'s orientation to their engagement in the current activity. *Bg* finds information in the help menu provided by the computer's operating system, "**the wireless button, situated on the computer?**" (extract 6 – lines 11 and 12). This distributed information concerning the computerized artefact allows them to find the solution to the network connection problem. But before it becomes a solution, the button must first be located on the laptop. Until a *double coordination* (Norman, 1991) has been accomplished by *Ji* and *Bg* between the written instruction and the task environment, the spatial information is not relevant (Conein, 2006). *Ji* accomplishes a pointing gesture (Goodwin, 1981; Schegloff, 1984) with her right hand towards an integrated button on her computer, simultaneously to her turn "**&∆oh but↑ it's this thing↑↑ no:↑∆=**" and presses the button (extract 6 – line 19). *Bg*'s turn "**=oh that's it↑ [hrrr**" quickly follows *Ji*'s without any break and confirms the success of *Ji*'s gestural action. Therefore, it seems that the action models the object "button" that is manipulated and is at the same time modelled by it. By coordinating their action with the relevant information and artefacts, the two co-participants have successfully co-operatively completed this new step in the connection procedure. From the point of view of the interaction, the extract cannot be reduced to simple conversational interaction, since it presents itself as a scene of three-party joint attention: *Ji* and *Bg* orient themselves to the button which becomes the common object of focus (Conein, 2006). The laptop is now finally connected to the university's Wi-Fi network. It is only when they have reached their goal of connecting the laptop that *Bg* orients himself to the presence of the two girls and realises that his chair could possibly be preventing them from accessing the notice board. Until then, *Ji* and *Bg* were engaged in a single activity, reading the instructions on the screen and looking for solutions. They were concentrating on the task at hand and probably had their awareness (Schmidt, 2002) about the surrounding environment reduced. The fact is that it is only when they are connected that *Bg* moves his chair, looks at the girls and says "**so?rry**".

The computerized artefact – informational and communicative – needed to be configured first before they could access the Internet: the activation of the Wi-Fi Card was crucial in accessing wireless Internet. We have observed the first steps of this procedure. Now that they have Internet access, the two interactants intend to accomplish different pedagogical activities using the Wi-Fi network.

2.5. Restricted Intranet and Internet Access

The fourth and last step in the proceedings of the first connection is the one that we have identified as the authorisation to access the university's server. Access to the faculty's pedagogical intranet requires a personalized identification number. *Ji* needs to enter the information relative to her user account which appears on her student card. The identification stage is, therefore, a determining one in the intranet access procedure, since the university's server only allows access to the students who are registered at the university. The identification number only allows access to the Wi-Fi network if the student's card is valid for the current academic year.

The student's card is therefore an official document, which allows the student to be identified, for all possible activities related to the university, including access to its restricted network. The card is a paper support where the identification number is written down, and kept by the student. In this respect, it is a cognitive artefact (Norman, 1991), where information is saved. We will, therefore, examine the role of this paper document in the digital collaborative activity between *Ji* and *Bg*, which becomes especially relevant because at first *Ji* does not use her student card. She types her university mail address and the connection fails.

1	Ji	>et c'est normal qu'il mette↑<
		>and is it normal that it takes↑<
2		trois plombes? ∇comme ça? ∇
		three hours? ∇like this? ∇
3		(0.2)

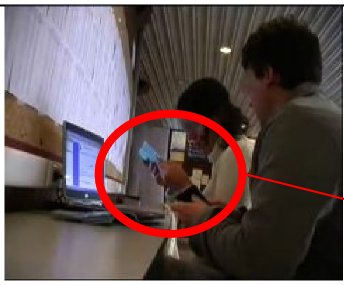
4	Bg	qu’est ce qui y a marqué sur
		what is there written on
5		l’autre fenê:tre?
		the other wi:ndow?
6		(2.0)
7	Bg	euh::: ∇ah nan∇ (.) là↑ faut que
		uhm::: ∇oh no∇ (.) there↑ you need
8		tu? recommences.
		to? start again.
9	Ji	ah bé min::ce↑ (.)∆ c’était pas
		oh da::mn↑ (.)∆ it wasn’t
10		ça↑ que vous aviez tapé↑∆
		that↑ that you had typed↑∆
11		l’aute jour::↑ parce que mon
		the other day::↑ because my
12		courri[e::l euh c’est pas
		mai[::l uhm that’s not
13	Bg	[∆mais je j’ai pas rgardé↑∆]
		[∆but i i didn’t have a look↑∆]
14	Ji	marqué là d’]dessu:s
		on] i:t written
15		(.)

Extract 7. ‘what the interactants see is what the interactants get’

At lines 1 and 2, *Ji* examines the situation and verbally remarks that there is another problem by her question. She compares the situation with the normal (fast) connection time that she expects. Collaboratively, *Bg* accesses the laptop’s interface and looks for an explanation (extract 7 – lines 4 and 5), and expects to find a message in the form of a window.

After this temporary failure, the action accomplished by the user *Ji* combines an understanding of the information she reads on the screen and an identification of the relevant data available on her student card. At line 7, *Bg* tells *Ji* that she needs to start again. By asking *Bg* whether the situation is normal or not, she attributes a form of expertise to *Bg,* who gathers information for the diagnostic before producing a prescription “**you need to? start again.**” (extract 7 – lines 7 and 8). The task of obtaining a connection is co-constructed by *Bg* and *Ji* in the progress of the interaction. The connection activity is accomplished by using *Bg*’s knowledge, which is called upon during the accomplishment of the activity. Both interactants are able to evaluate the situation in real time by reading the computer's response, that is, accessing to the intranet has failed because *Ji* has not entered the appropriate login details. What the interactants see on the screen is what the interactants get: the effect of the rapid incremental reversible operations on the object of interest is immediately visible (Shneiderman, 1997). Sequentially, the successive turns of *Bg* (extract 7 – line 7-8, 13) and *Ji* (extract 7 – line 9-14) are the consequences of the error message that appears on the screen. In order to correct the data, this time, instead of entering her university's email address as login, *Ji* takes her student card, on which her student's identification number is written.

1 Bg `il a: utilisé:: les- il a`
he ha:s used:: the- he has
2 `utilisé::↑`
used::↑



Student's card

3 Ji `ah si. ∆ah:? c'est∆ avec que::`
oh yes. ∆oh:? it's∆ with only::
4 `euh maju::scule↑`
uhm capital:: letter↑
5 Bg `il a utilisé? ça:: là.`
he used? that:: here.
6 `(0.2)`
7 Ji `>non mais en fait?,< y faut une`
>no but in fact?,< you need a
8 `majuscule:: ∇pour l'écrire∇`
capital letter:: ∇to write it∇
9 `(29.0)`
`((Bg points the on-screen text with his left hand))`
10 Bg `non: c'est mar:?qué (.) euh nnn`
no: it's wri:?tten (.) uhm nnn
11 Ji `[en minuscule↑ et sans la clef`
[in small letter↑ and without the key
12 Bg `[en minuscule (.) et sans la clef,`
[in small letter (.) and without the key,

13	Ji	ah↑ min:ce euh::::↑↑
		oh↑ su:gar uhm::::↑↑
14		(27.0)((typing her identification number))
15	Ji	ça fait dix? ça↑
		does it make ten? this↑
16		(6.0)
17	Bg	voilà.
		here you go.
18	Ji	ouais:::↑↑
		yeah:::↑↑

Extract 8. access authorization to university's server

Using the document containing the data is locally relevant for the two interactants. The inscriptions do not identify the student by name, but provide a different form of information, which is not readily accessible. Instead of a name or an email address, the student card gives a code that has to be deciphered according to the instructions given on the graphic user interface in an adequate and acceptable fashion. Then, the student has to enter the number according to the the instructions on the screen.

In this extract, the card serves as a tool to create the connection, and becomes a cognitive artefact to complement the computer screen's interface (Hollan, 2000). Both the student card and the information appearing on the screen are used to correct the login, allowing successful access to the university's intranet. During subsequent connection attempts, the card will certainly continue to be used as a cognitive artefact, as a memory aid (Norman, 1991). This use of the artefact reflects its situated and embodied character in the context of action and use, as the actions are distributed in time, and between persons who accomplish the activity. Thus, the cognition is "distributed" (Hutchins, 1995) between the student *Ji* and the object "the student card" as a physical element of her environment which helps her in successfully obtaining a connection.

> *"The term 'distributed cognition', for instance, is increasingly used to demarcate a concern with (socially) shared representations and the co-ordination of action by individuals in organizational environments"* (Heath et al., 2000: 306).

When accessing the server, some information which is available on the university's home page appears on *Ji*'s laptop screen. *Ji* follows the instructions exactly. Therefore, the computer, in turn, becomes an informational artefact, which displays information that becomes an available resource for the participants in action. At lines 11 and 12, *Ji* and *Bg* simultaneously utter what they synchronously read: "`[en minuscule↑ et sans la clef`", which is the message displayed on the screen. In the setting of an observation of how usages are constructed, the interface creates a distribution of the shared resources between the user and the information displayed on the screen. The interface as a cognitive artefact presents informational proprieties. It is, thus, possible to observe a shared contextual representation between the interactants *Ji* and *Bg* (extract 8 – lines 11 and 12). The overlapping of the two utterances demonstrates the interactants' common and simultaneous orientation to the displayed information relevant for the completion of this situated activity.

Moreover, the handling of, and the interaction with, the interface allow a look into the participants' specific orientation to the activity. *Ji* at first tries to enter her password using capital letters, before starting again using lower-case letters. The identification, accomplished by the users themselves, of what is the correct login and password, is a situated activity (Suchman, 1987), closely linked to a specific context and usage framework. The interactants' orientation in the learning process is dynamic and develops progressively as the interaction unfolds.

We have closely observed that once the connection steps are accomplished and the restriction overcome, it is possible for the students to obtain an Internet connection from almost anywhere on the campus. In order to connect, they rely on certain pieces of information spatially oriented towards the accomplishment of the task: the information concerning the connection procedures are directly available on the university's website and the student's card contains the necessary identification number. Therefore, the space and spatial arrangements in the participants' immediate environment (technological space of the laptop and the entrance hall as a public space) influences the action and at the same time is influenced by it. This contextual configuration of the space and distributed cognition is closely linked to the technology usage.

3. Conclusion

With the arrival of Wi-Fi network on the premises of the campuses of Montpellier, new digital environments have been created through users' appropriation of the space, their attempts in connecting to and using the wireless network, like *Ji* whose connection steps we have described. These places were traditionally practical spaces where social interaction took place or where people passed by. With the coming of the wireless network and the furnishing of this crossing and notice area, heterogeneous activities started to take place within the same space, rendering the term "place of study" difficult to define: pedagogical and digital activities move from the library or computer lab to this public open space. Therefore, like in any public space, the activities of different groups of people co-occur without merging, and mobile wireless Internet activities occur alongside others. In the light of the temporally unfolding activities we have described, and the characteristics which have emerged empirically through the accomplishment of the specific steps of connection, we propose the term of *Ambient Internet Space*.

Thus, the *Ambient Internet Space* is more than simply a wireless network covered area or a public and open space. The two students *Pm* and *Rl* find themselves in the same place as *Ji* and *Bg,* but as opposed to them, *Pm* and *Rl* are neither engaged in computer activities nor even orienting to the possibilities of using the Wi-Fi network. The *Ambient Internet Space* emerges from the moment laptop users orient to the possibility of accessing the internet and accomplish connection activities until the identification requirements are met and the Wi-Fi network is accessed and used. However, from the description of *Ji*'s successive steps in connecting and the issues that were raised during the activity, we notice that even though *Ji* is accustomed to using her laptop and a cabled connexion, she does not orient to the wireless connection procedures as being straightforward.

The personal computer is widely and commonly considered as a valuable piece of technology for people and purposes. As technological and communicative infrastructures, computers and wireless networks should ideally be "invisible" (Norman, 1998) or "transparent" (Ishii, 2004). Empirically, through video ethnography, we have seen that the use of the Wi-Fi network, which is new to her, is far from being transparent for *Ji*. At home, her routine accomplishment of connecting her laptop is naturally integrated in her daily activities, and is as transparent to her as "plugging in and using my hairdryer". Unlike a simple and non-problematic home connection, her failures and successive attempts to obtain a connection are central to the activity, and constitute the core of our analysis. Far from being "invisible" for the user, the activity of connection necessitates full attention and involvement in this specific task, and precise interpersonal coordination concerning the use of artefacts. It involves the users' reading of the computer responses in the form of error messages, the interpretation of information contained in artefacts, like the student card, and situating this information in the spatial environment of action (finding and pressing the button) in order to accomplish the required connection steps.

Therefore, studying the role and the place of the artefacts in the interaction as they fit into the structure of the connection activity within this newly-created hybrid space, the *Ambient Internet Space,* necessarily goes along with the study of the situated talk-in-interaction as they are engaged in the activity. It constitutes the main data for analysing the participants' orientation to the artefacts or to the coordination of action, even in a multimodal analysis of interaction. That is why we chose an *in situ* approach, where activities are examined *in situ*, as they unfold in the *hic* and *nunc* of the interaction, through the close analysis of the participants' conversations, actions and the distribution of their activity with the environment.

4. Transcript Symbols

Data were transcribed according to conventions developed mainly by Gail Jefferson and commonly used in conversation analysis.

[	overlapping talk
=	Latching
(.)	micro pause
(8.0)	Pause
:	extension of the sound or the syllable it follows
.	stopping fall in tone
,	continuing intonation
?	rising inflection
Mine	Emphasis
°uh°	quieter fragment than its surrounding talk
.hh	Aspiration
Hh	out breath
&	a continuation of a same turn after overlapping talk
><	quicker than surrounding talk
↑	mark an overall rise in pitch across a phrase
∆	rise in volume

∇	lower volume
£	laughing tone in the voice
hrr heu	laugh transcript

Figure 2. The table of transcript symbols

5. References

Barkhuus, L., & Dourish, P. (2004). Everyday encounters with context-aware computing in a campus environment. *Proceedings of the sixth international conference on Ubiquitous Computing Ubicomp 2004,* Nottingham.

Button, G. (1991). Conversation-in-a-series. In D. Boden & D. H. Zimmerman (Eds.), *Talk and social structure: studies in ethnomethodology and conversation analysis* (pp. 251-77). Cambridge: Polity Press.

Conein, B. (2006). Agir dans et sur l'espace de travail avec des objets ordinaires. In Mondada L. – Espace, inter/action & cognition – *Intellectica*, n41-42, pp. 163-179.

Conein, B. et Jacopin, E. (1993). Les objets dans l'espace. La planification dans l'action. In Les objets dans l'action, *Raisons Pratiques*, 4, pp. 59-85.

Conein, B. (1990). Cognition située et coordination de l'action. La cuisine dans tous ses états. In Opinion, savoir, communication, *Réseaux* n43, pp. 99-110.

Crabtree, A., Evans, M.A., Fraser, M., Tolmie, P., & McMullen, D.F. (2005). Situating Ubiquitous computing in everyday life: Bridging the social and technical divide. *Workshop proposal submitted to the UbiComp 2005 International Conference*, Tokyo, Japan, September 11-15.

Dourish, P. (2004). What we talk about when we talk about context. *Proceedings Personal and Ubiquitous Computing*. Springer, 19-30.

Dourish, P., Grinter, R., Delgado de la Flor, J., & Joseph, M. (2004). Security in the Wild: User strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing*, *8*(6), 391-401.

Drew, P., & Heritage, J. (1992). *Talk at work: interaction in institutional settings*. P. Drew, & J. Heritage (Eds), Cambridge: Cambridge University Press.

Garfinkel, H. (1967). *Studies in ethnomethodology.* Englewood Cliffs, N.J.: Prentice-Hall

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Houghton Mifflin, Boston.

Goodwin, C. (1996). Transparent vision. In: Ochs, E., Schegloff, Thompson, eds., *Interaction and Grammar*. Cambridge: Cambridge University Press: 370-404.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606-633.

Goodwin, C. (1987). Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments (with Marjorie Harness Goodwin). *IPrA Papers in Pragmatics, 1*(1), 1-55.

Goodwin, C. (1981). *Conversational organisation: interaction between a speaker and hearer.* London: Academic Press.

Goodwin, M. H. (1980). He-said-she-said: Formal cultural proceedings for the construction of a gossip dispute activity. *American Ethnologist, 7*, 674-695.

Grosjean, M. (2005). *L'awareness à l'épreuve des activités dans les centres de coordination.* Retrieved April 2006, from http://www.activites.org/v2n1/html.grosjean.htlm.

Have, P. ten, (2002). Reflections on transcription. *Cahiers de praxématique, 39*, 21-43.

Heath, C. (2002). Demonstrative suffering: the gestural (re)embodiment of symptoms. *Journal of Communication, 52*(3), 597-617.

Heath, C. (1989). Institutional form and interactional detail: the delivery of diagnosis in a medical consultation. *Langage et travail*, 280-289.

Heath, C., Knoblauch, H., & Luff, P. (2000). Technology and social interaction: the emergence of 'workplace studies'. *The British Journal of Sociology*, *51*(2), 299-320.

Heritage, J. (1984). *Garfinkel and Ethnomethodologie*. Cambrige: Cambridge Polity Press.

Hollan, J. (2000). Human Computer Interaction. In R. Wilson & F. Keil (Eds), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 379-380). MA: MIT Press.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.

Ishii, H. (2004). Bottles: A transparent interface as a tribute to Mark Weiser. *IEICE Transactions on information and systems*, *E87-D*(6), 1299-1311, June.

Lerner, G. H. (2002). Turn-sharing: The Choral co-production of talk-in-interaction. In C. Ford, B. Fox & S. Thompson (Eds.), *The language of turn and sequence*. New York: Oxford University Press: 225-256.

Licoppe, C., & Relieu, M. (2005). Entre système et communication. Une approche située de la compétence des téléopérateurs dans les services d'assistance technique, dans *Le travail avec les technologies de l'information.* Hermes Science Publications, pp. 177-199.

Mondada, L. (2006). La construction de l'origo déictique comme travail interactionnel des participants : une approche praxéologique de la spatialité, *in* Mondada L. – Espace, inter/action & cognition – *Intellectica*, n41-42, pp. 75-100.

Norman, D. (1998). *The invisible Computer*. Cambridge, MA: MIT Press.

Norman, D. (1991). Cognitive artefact. In J. Caroll (Ed.), *Designing interaction psychology at the human-computer interface*. Cambridge: Cambridge University Press.

Relieu, M. (1999). Parler en marchant: Pour une écologie dynamique des échanges de paroles. *Langage & société*, *89*, 37-67.

Sacks, H. (1972). An Initial Investigation of the Usability of Conversational Data for Doing Sociology. In D. Sudnow (Ed.), *Studies in Social Interaction.* New York: The Free Press.

Sacks, H. (1992). *Lectures on conversation*. Oxford: Basil Blackwell.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696-735.

Salambier, P., & Pavard, B. (2003). Analyse et modélisation des activités coopératives situées. Evolutions d'un questionnement et apports à la conception. In *Activites.org*. Retrieved April 2006, from http://www.activites.org/v1n1/salembier.pdf

Schegloff, E.A., & Sacks, H. (1973). Opening up closings, *Semiotica, 8*, 289-327.

Schegloff, E. A. (2002). Beginnings in the Telephone. In J.E. Katz & M. Aakhus (Eds.), *Perpetual Contact: Mobile communication, private talk, public performance* (pp. 284-300). Cambridge: Cambridge University Press,.

Schegloff, E. A. (1996). Confirming allusions: toward an empirical account of action. *American Journal of Sociology, 104*, 161-216.

Schegloff, E. A. (1991). Conversation Analysis and Socially Shared Cognition. In L. Resnick, J. Levine, & S. Teasley (Eds.), *Perspectives on Socially Shared Cognition.* Washington, D.C.: American Psychological Association, 150-171.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action (*pp. 266-298). Cambridge: Cambridge University Press.

Shmidt, K. (2002). The problem with "Awareness" – Introductory remarks on 'Awareness in CSCW'. *Computer Supported Cooperative Work, 11*, 285-298.

Shneiderman, B. (1997*). Designing the User Interface: Strategies for Effective Human-Computer Interaction (3rd Edition)*. MA: Addison Wesley, Reading.

Suchman, L. (1987). *Plans and situated actions – The problem of human-machine* *communication*. Cambridge: Cambridge University Press.

Weiser, M. (1998). The future of Ubiquitous Computing on Campus. *Communications of the ACM* 41(1): 41-42.

Weiser, M. (1991). The computer for the 21st century. *Scientific American, 265*(3), 94-104.

Whalen, J. , & Zimmerman, D.H. (2005). Working a call: Multiparty management and interactional infrastrucure in calls for help. In *Calling for Help: Language and Social Interaction in Telephone Helplines*, Baker, Carolyn, Michael Emmison and Alan Firth (eds.). Amsterdam: John Benjamins, pp. 309–345.

Williams, A. and Dourish, P. (2006). Reimagining the city: the cultural dimensions of urban computing. *IEEE Computer*, *39*(9), 38-43.

Williams, A., Kabisch, & Dourish, P. (2005). From interaction to participation: Configuring space through embodied interaction. *Proceedings of the International Conference on Ubiquitous Computing (Ubicomp 2005)*, Tokyo, Japan, September 11-14, pp. 287-304.

Wilson, T.-P. (1991). Social structure and the sequential organization of interaction. In: Boden, Zimmerman, (Eds.), *Talk and Social Structure*, Polity Press and U.C. Press pp. 22-43.

PsychNology Journal, 2007
Volume 5, Number 2, pp. 197 – 206

Identifying the (Tele)Presence Literature

Matthew Lombard ♣ ♠ and Matthew T. Jones ♠

♠ Temple University, Philadelphia

ABSTRACT

This paper discusses the value of identifying the expanding interdisciplinary scholarly literature on the topic of (tele)presence, proposes a detailed procedure for doing so, and presents a list of 1,831 journal articles, books and other publications that constitute the (tele)presence literature as of May 2007.

Keywords: *Telepresence, Telepresence Literature, Telepresence Bibliography, Spatial Presence, Social Presence*
Paper Received 19/06/2006; received in revised form 10/06/2007; accepted 09/08/2007.

1. Introduction

Since Short, Williams and Christie (1976) wrote about social presence in organizational communication and Minsky (1980) wrote about telepresence and the dynamics of teleoperation, a growing number of scholars in communication, computer science, psychology and many other disciplines have studied (tele)presence and (tele)presence phenomena in a variety of contexts ranging from art to engineering. Some (e.g., Barfield & Weghorst, 1993; Biocca, Harms, & Burgoon, 2003; Draper, Kaber, & Usher, 1998; Freeman, 2004; Heeter, 1992; Lee, 2004; Lombard & Ditton, 1997; Mantovani & Riva, 1999; Sheridan, 1992; Waterworth & Waterworth, 2001; Zahorik & Jenison, 1998) have contributed detailed, often quite diverse, concept explications and many others have reported studies of presence phenomena and used conceptualizations of presence without specifically using the term. Because the research and theory related to (tele)presence is highly interdisciplinary, diverse and scattered, there is no well defined and generally agreed upon definition of what

---

♣ Corresponding Author:
Matthew Lombard,
BTMM Department, Annenberg Hall (011-00),
Temple University, Philadelphia, Pennsylvania 19122
phone: (215) 204-7182;
fax: (215) 204-5402
e-mail: lombard@temple.edu

constitutes the literature on this topic. We argue here that identifying such a literature will have important benefits for the advancement of (tele)presence scholarship, allowing us to build on work not previously identified as relevant to presence. We then describe a procedure we've developed to identify this literature before presenting and briefly discussing the results of its application: a list of 1,834 publications that constitute the current (as of May 2007) (tele)presence literature.

2. Benefits of Defining the (Tele)presence Literature

A comprehensive list of the scholarly works on (tele)presence would have three important benefits.

First, it would provide an overview of the development and current status of the work in this topic area. It would reveal the volume of work in the area produced to date, the origin and growth trajectory of presence scholarship, and the disciplinary, geographic and institutional diversity of the work (and changes in these over time), all in objective and quantitative terms. Regular updates to the list would allow us to chart future changes in the field's maturity. Evidence of this maturity would have important advantages in justifying (tele)presence scholarship to various stakeholders including individuals and institutions charged with making funding decisions regarding telepresence scholarship and technology.

Second, despite many discussions and explications, there is still substantial disagreement about the meanings and scope of the telepresence and presence concepts. Telepresence is generally used to refer to the sense of 'being there' experienced by users of advanced media such as virtual reality, and more generally, the psychological state or subjective perception in which a person fails to accurately and completely acknowledge the role of technology in an experience. Presence is used both as a shortened version of the term telepresence and to refer to the actual or perceived physical presence of objects and entities, in technology-mediated and/or nonmediated contexts. Further, scholars have used several variants of these terms to distinguish different dimensions or types of the concepts, and there is little agreement about how to appropriately define and measure them. A list, and then a detailed review, of all relevant publications and the conceptual and operational definitions within them would provide a needed perspective on these debates and help us identify and set aside redundancies so we could focus on the important outstanding issues and increase the cohesiveness, and value, of future scholarship in the field.

Third, especially because it's a particularly interdisciplinary and rapidly changing technology-based field, a comprehensive list of (tele)presence publications is likely to contain theories, findings, and descriptions of phenomena, technologies and applications of technologies related to (tele)presence that many presence scholars wouldn't ordinarily encounter. This exposure would likely create new synergies and suggest new research questions that would help us advance our collective work. Conversely, such a list would create a resource researchers and theorists outside of the domain of (tele)presence might conveniently use to consider their work from the perspectives of (tele)presence scholarship.

3. Difficulties Involved and Previous Attempts

Constructing a valid and comprehensive list of (tele)presence literature poses several challenges. Because of the interdisciplinary nature of the topic, relevant works exist in many different, established literatures each with their own indices, databases and search protocols. Even when scholarly databases are selected there is the problem of what search terms to use. The word presence is generic and often appears in scholarly and other publications in contexts completely unrelated to the concept of interest, producing hundreds of thousands of irrelevant 'hits.'

The debates about what constitutes presence (or telepresence) have led to the use of several different variants of the term for the same and similar concepts – telepresence, virtual presence, social presence, sense of presence, subjective presence, etc., as well as terms that don't include the word at all – parasocial, perceived reality, computers are social actors, and others. While the more specific compound terms yield more useful search results, one has to compile a long list of search terms to capture the complete literature, and many of them, such as virtual presence, produce irrelevant publications that must be "manually" sifted.

A few previous attempts to compile formal and informal collections of presence literature are useful and instructive. IJsselsteijn, Lombard and Freeman (2001) have proposed a core bibliography of presence. While this list, which contains 95 publications, is valuable for determining some of the most influential work in the field, the focus is narrow and the authors therefore didn't confront the same problems that an exhaustive survey of literature presents.

Emerson, Steed, and Billinghurst (1999) compiled a "Presence Bibliography" as a technical report for the University of Washington Human Interface Technology Laboratory (HitLab). While the list of 255 works is more inclusive than the core literature, it lacks explicit criteria for inclusion, contains items not generally available, and is 6 years old as of this writing.

The presence-research.org web site, the official site of the OmniPres project funded by the European Union's Presence Research Proactive Initiative, contains a bibliographic list of 92 online papers related to presence, but the inclusion criteria are not provided. Many of the works are early versions presented at conferences, and while some of the works are on the web site, others are only linked and may not remain accessible.

Another listing of presence publications was on the web site of the International Society for Presence Research (ISPR), http://ispr.info, until 2007. While it was a large (n=463) list and contained abstracts for many entries, it too was compiled without explicit or systematic criteria for inclusion.

What is needed is a replicable, systematic and updateable document that encompasses all (or as much as is possible) of the theory and research that is directly relevant to the study of (tele)presence. The above examples, although inadequate in various ways, are important resources that are capable of providing a base to begin strategizing as to how such a comprehensive list can be created and maintained.

4. Method: A Strategy for Defining the Literature

We first developed a set of parameters and criteria regarding works that should and should not be included in a (tele)presence literature and procedures for identifying them.

The works should contain academic/scholarly writing that is relevant to and preferably focused on the concepts and/or phenomena of presence or telepresence, as defined broadly by presence scholars. Recognizing the interdisciplinary nature of the topic, works from any field or discipline – not just psychology, communication and/or computer science - should be eligible for inclusion. Articles in the popular press represent a different, though interesting, venue for information related to (tele)presence and should be excluded. To be useful, works in the literature should be generally available, that is, catalogued and published or archived in some way that is accessible to the academic community. Journal articles, conference proceedings (but not other conference papers), books and doctoral dissertations normally meet this criterion. Technical reports and masters' theses normally do not qualify (the procedures and standards for these also vary particularly widely and in the case of reports may be biased by commercial concerns).

To insure that works across as many fields as possible were systematically considered, we evaluated several database alternatives, conducting a series of searches and evaluating the results (for example, the database Google Scholar (n/d) is very extensive, but test searches contained many duplicate and irrelevant entries and Google, at this writing at least, has not divulged which resources are searched).

We eventually selected four scholarly citation databases: ComAbstracts, Computer Abstracts, PsycINFO, and ISI Web of Science. ComAbstracts (n/d), maintained by the Communication Institute for Online Scholarship (CIOS), is a database of "sources of relevance to researchers, scholars, and students interested in fields related to human communication studies (mass communication, human interaction, rhetoric, health communication, communication and new media, journalism, communication history, etc.)." Computer Abstracts International Database (CAID) (n/d), published by Emerald, "covers the major computer science subjects including: artificial intelligence; communications and networks; computer theory; database and information systems applications; hardware; human-computer interaction; mathematics of computing; programming; and systems organization." PsycINFO (n/d), published by the American Psychological Association (APA), "provides abstracts and citations to the scholarly literature in the behavioral sciences and mental health. The database includes material of relevance to psychologists and professionals in related fields such as psychiatry, management, business, education, social science, neuroscience, law, medicine, and social work." ISI Web of Science (n/d), from Thomson, is a database index of "current and retrospective multidisciplinary information from approximately 8,700 of the most prestigious, high impact research journals in the world." It includes works in the hard sciences (via Science Citation Index and two chemistry databases), social sciences (via Social Sciences Citation Index), and the arts and humanities (via Arts & Humanities Citation Index).

To insure adequate coverage of the scholarly work on (tele)presence, all of the references identified by Ijsselsteijn et al. (2001) as part of the "core bibliography of presence" were included. Because they are logical and common venues for the publication of work on (tele)presence, all references published in the MIT Press journal *Presence: Teleoperators & Virtual Environments* and its online, peer-reviewed supplement, presence-connect.com, and the Mary Ann Liebert, Inc. journal *Cyberpsychology & Behavior* were included if their title and/or abstract included the use of the term presence or a variant of that term (other than in an incidental or irrelevant context).

The next step was to establish a list of appropriate search terms. A diverse set of words and phrases have been associated with the concept of presence; only those that directly refer to the concept itself rather than related concepts and technologies (e.g., flow, suspension of disbelief, immersion, virtual reality) were included.

The selected search terms fall into two categories: (1) terms that include the word presence and (2) specialized terms that do not contain the word presence but refer to a concept that represents (tele)presence or a key dimension of (tele)presence (e.g. parasocial, CASA). The 16 selected search terms are: telepresence, tele-presence, (tele)presence, spatial presence, social presence, parasocial, computers are social actors, copresence, co-presence, subjective presence, virtual presence, sense of presence, perceived realism, perceived reality, perceptual realism, and social realism.

It's important to note that references identified through database searches, especially for certain search terms (e.g., virtual presence), require a check of the context in which the terms are used. Reading the title of the work in question often clarifies the relevance of the work, but it is not uncommon for further contextualization, usually found within the abstract, to be necessary. Search terms that yield a high proportion of irrelevant references, e.g., the single word presence, were deemed impractical and not used (for this reason it may be useful for scholars to avoid using only the term presence as a shortened term for telepresence). Many terms, such as virtual reality, simulation, illusion, engagement, flow, etc., are clearly related to (tele)presence but yield too many references that don't directly discuss or concern presence or telepresence per se, and are therefore not useful in identifying this literature.

5. Results

Over 1,800 (n=1,831) published works were identified as constituting the (tele)presence literature as of May 2007. The complete list of references in APA format is available here and on the "Bibliography" page of the web site of the International Society for Presence Research (ISPR) at http://ispr.info.

The works date back to the 1930s and the majority of them are journal articles (n=1,614). They feature an impressive diversity of topics and come from a wide array of fields. While *Presence: Teleoperators & Virtual Environments* (n=206) and *Cyberpsychology & Behavior* (n=318) are the most represented, works come from journals in art (*Journal of Art & Design Education*), business (*Production Planning & Control*), communication (*Human Communication Research*), computer science (*Journal of Network and Computer Applications*), education (*Journal of Educational Television*), engineering (*IEEE Transactions on Robotics and Automation*), linguistics (*Cognitive Linguistics*), medicine (*International Journal of Adolescent Medicine & Health, Journal of General Internal Medicine*), music (*Journal of New Music Research*), philosophy (*Philosophy & Phenomenological Research*), physics (*International Journal of Modern Physics C*), psychology (*Journal of Psychology*), religion (Studies of Formative Spirituality), social work (*Journal of Social Issues*), sociology (*American Journal of Sociology*), and more. Many nations are also represented, including Germany (*Zeitschrift fur Psychotogie*), Italy (*Rassegna Italiana di Sociologia*), Japan (*Japanese Journal of Experimental Social Psychology*), The United Kingdom (*Elsevier Science*), The United States (*Communication Theory*), and Scandinavia (*Acta Psychiatrica Scandinavica*).

Many interesting phenomena are examined in the works which can serve to expand our conceptualizations and understanding of presence and telepresence. Just a few examples are the sense of the presence of a recently deceased spouse (Conant, 1996) and the illusion of the presence of others as a result of sleep paralysis (Cheyne, 2001), brain injury (Persinger, 1994), psychiatric condition (Koehler & Sauer, 1984) and experimental induction using specific (frequency-modulated) complex magnetic fields (Cook & Persinger, 1997; Tiller & Persinger 2002).

Although there is substantial overlap, many works come from only one of the databases or other sources, suggesting the value of scholars using a broad initial search strategy to obtain material when conducting literature reviews.

6. Conclusion

Building on the work of others, we've developed and described here a systematic procedure for identifying the scholarly literature on (tele)presence. The resulting set of works demonstrates the impressive volume and diversity of work in this area and suggests new avenues for conceptualization and understanding. Detailed analyses of the works identified, regular updates of the literature, and perhaps even the use of the identified works as a basis for a centralized database for (tele)presence researchers and theorists would further help demonstrate the health of the field, resolve disagreements, focus efforts, and increase understanding of presence and telepresence phenomena.

7. References

Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy & M. Smith (Eds.), *Human computer interaction: Software and hardware interfaces* (pp. 699-704). Amsterdam: Elsevier.

Biocca F., Harms C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoperators and Virtual Environments, 12*(5), 456-480.

Cheyne, J. A. (2001). The ominous numinous: Sensed presence and 'other' hallucinations. *Journal of Consciousness Studies, 8*(5-7), 133-150.

CommAbstracts. (n.d.). [Database]. Communication Institute for Online Scholarship (CIOS). Retrieved July 17, 2007, from http://www.cios.org/www/abstract.htm

Computer Abstracts International Database (CAID). (n.d.). [Database]. Emerald. Retrieved July 17, 2007, from http://lysander.emeraldinsight.com/info/products_services/abstracts/CAID/index.jsp

Conant, R. D. (1996). Memories of the death and life of a spouse: The role of images and sense of presence in grief. In D. Klass & P. R. Silverman (Eds.). *Continuing bonds: New understandings of grief. Series in death education, aging, and health care* (pp. 179-196). Philadelphia: Taylor & Francis.

Cook, C. M., & Persinger, M. A. (1997). Experimental induction of the "sensed presence" in normal subjects and an exceptional subject. *Perceptual & Motor Skills, 85*(2), 683-693.

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40*(3), 354-375.

Emerson, T., Steed, A., & Billinghurst, M. (1999). *Presence Bibliography.* University of Washington Human Interface Technology Laboratory (HitLab) Technical Publication: R-99-9. Retrieved July 17, 2007, from http://www.hitl.washington.edu/publications/r-99-9/

Freeman, J. (2004). *Implications for the measurement of presence from convergent evidence on the structure of presence*. Paper presented at the may 2004 conference of the International Communication Association, New Orleans, LA.

Google Scholar. (n/d). [Database]. Google. Retrieved July 17, 2007 from http://scholar.google.com/

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1*(2), 262-271.

IJsselsteijn, W. A., Lombard, M., & Freeman, J. (2001). Toward a core bibliography of presence. *Cyberpsychology & Behavior, 4*(2), 317-321.

ISI Web of Science. (n.d.). [Database]. Thomson. Retrieved July 17, 2007 from http://scientific.thomson.com/products/wos/

Koehler, K., & Sauer, H. (1984). Jaspers' sense of presence in the light of Huber's basic symptoms and DSM-III. *Comprehensive Psychiatry, 25*(2), 183-191.

Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27-50.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). Retrieved February 25, 2005 from http://jcmc.indiana.edu/vol3/issue2/lombard.html

Mantovani, G., & Riva, G. (1999). 'Real' presence: How different ontologies generate different criteria for presence, telepresence, and virtual presence. *Presence: Teleoperators and Virtual Environments, 8*(5), 540-550.

Minsky, M. (1980). Telepresence. *Omni,* June, 45–51.

Persinger, M. A. (1994). Sense of a presence and suicidal ideation following traumatic brain injury: Indications of right-hemispheric intrusions from neuropsychological profiles. *Psychological Reports, 75*(3, Pt 1), 1059-1070.

PsycINFO. (n.d.). [Database]. American Psychological Association. Retrieved July 17, 2007 from http://www.apa.org/psycinfo/products/electronic.html

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments, 1*, 120-126.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: Wiley.

Tiller, S. G., & Persinger, M. A. (2002). Geophysical variables and behavior: XCVII. Increased proportions of the left-sided sense of presence induced experimentally by right hemispheric application of specific (frequency-modulated) complex magnetic fields. *Perceptual & Motor Skills, 94*(1), 26-28.

Waterworth, E. L., & Waterworth, J. A. (2001). Focus, locus, and sensus: The three dimensions of virtual experience. *Cyberpsychology & Behavior, 4*(2), 203-213.

Zahorik, P., & Jenison, R. L. (1998). Presence as being-in-the-world. *Presence: Teleoperators and Virtual Environments, 7*(1), 78-89.

PsychNology Journal, 2007
Volume 5, Number 2, pp. 207 -222

Searching for Information on PDA in a Naturalistic Environment with or without Music

Andrea Zucchi[♠♦] and Luciano Gamberini[♦]

[♦]HTLab, Dept. of General Psychology
University of Padua, Italy

ABSTRACT

The present paper investigates the effects of different kinds of music on information searching in a crowded cafeteria. Our experiment employs four texts that differ in length only, not in content or syntactic complexity. Each text is loaded on a web page and displayed on a PDA. We ask to participants 20 questions about each text and then compare their accuracy and performance time. Participants carry out their task in three different background conditions: normal environmental noise, earphones with classical symphonic music or earphones with modern Italian songs. We assume that classical symphonic music improves information searching by isolating the participant from the noise and background chitchatting of the cafeteria, while modern Italian music compromises performance, because processing the lyrics interferes with the task. In line with our expectations, classical music significantly improves information searching, but contrary to our hypotheses, Italian music improves performance, although not in a significant way. We conclude that in a situation with background noise, listening to classical music increases the speed of information searching with respect to a condition without music.

Keywords: *PDA, Searching for information, Background music, Noise, Irrelevant speech, Information processing, Arousal.*

Paper Received 15/05/2007; received in revised form 08/08/2007; accepted 10/08/2007.

1. Introduction

Mobile technologies play a central role in today's world (Schaffers, Brodt, Pallot, & Prinz, 2006). The advantage of mobile technology is that you can always carry it with you. A palmtop, more than a laptop, can be carried anywhere. Such mobile devices are often used by those who travel because of their profession. Therefore, PDAs are generally employed in dynamic and noisy environments in which users are moving (MacKay, Dearman, Inkpen, & Watters, 2005).

[♠] Corresponding Author:
Andrea Zucchi
HTLab, Dept. of General Psychology, University of Padua
Via Venezia 8, 35131 Padua, Italy
E-mail: andrea.zucchi@gmail.com

Mobile technology has recently added web browsing that provides PDA access to web information. The Internet offers all kinds of information. When most people surf the web with their PDA, they are not looking for complex information that requires longer time to be found and a full-size desktop screen to be displayed. Instead, “the searching behavior supported by PDAs tends to be one in which the user seeks specific information” (Kim & Albers, 2001, p. 193) in noisy and crowded environments. For instance, while having breakfast or lunch in a cafeteria, people may like to read the daily news. They need to find the right website among many online newspapers. In such a situation, the PDA user is typically more exposed to background noise (in particular, background conversation) than the PC user in an office. Recently, Kallinen (2004) conducted a study in which participants in a noisy environment read texts on PDA when listening to music with earphones. The data obtained demonstrate that music not only improves reading abilities by reducing distraction and enhancing attention but also increases positive emotional responses and the level of general satisfaction. According to Kallinen, music enhances the social richness of the medium “because it increases one meaningful information channel (i.e., music instead of background noise)” (p. 1227). But when searching for information on a small screen display, does listening to music facilitate the performance, or cause more errors? Moreover, are the accuracy and the performance time influenced by the kind of music?

This paper attempts to answer these questions. There is a lack of studies on searching for information on PDA in ecologically valid realistic settings[4]. Our study investigates information retrieval in a web page displayed on PDA in an ordinary environment with different background conditions.

2. Experiment

2.1 Study Design

There are three variables in our experiment:

- background condition (without music, with classical symphonic music, and with Italian music);
- length of the body of the text (450, 550, 650, and 800 words);

[4] Kim and Albers (2001) focus on the differences between handheld and desktop interfaces and explore ways in which design of handheld may affect information retrieval; however, the experimental setting is not naturalistic.

- location of key information within the body of the text (beginning, middle, and end).

The location of the key information is a within-subjects factor, whereas the other two variables are handled between-subjects. Since the aim of our research is to investigate the effects of different kinds of music on information searching, the background condition is the main independent variable. We have employed texts of different lengths with different positions of the key information, because there could exist some interaction between these variables and the background condition.

2.2 Background Condition

When searching for textual information, one must focus on the task of searching, without being distracted by the external environment. After the desired information is found, one must decide if it is correct and exhaustive enough. The information should be repeated, so that it can be maintained in the conscious workspace, until it is stored in the long-term memory, or written somewhere (on a spreadsheet, for instance). Furthermore, most information cannot be found directly "but by narrowing the search down based on implicit relationships between what the person is currently viewing and what is desired" (Kim & Albers, 2001, p. 195). These cognitive capacities could be affected by external stimuli such as noise or music.

Noise

There are different kinds of noise. White noise[5], road traffic noise, meaningless irrelevant speech, and meaningful irrelevant speech are the most common and studied. The former two are nonverbal, whereas the latter two are verbal.

Immediate serial recall declines in the presence of meaningless and meaningful irrelevant speech but is unaffected by white noise (Rouleau & Belleville, 1996).

Studies of irrelevant speech and short-term memory underline marked impairment from both meaningful and meaningless speech, particularly on memory tasks with a serial component (e.g., Martin, Wogalter, & Forlano, 1988; Jones, Miles, & Page, 1990; Oswald, Tremblay, & Jones, 2000; Tremblay, Nicholls, Alford, & Jones, 2000). In reading comprehension tasks, meaningful irrelevant speech is more disruptive than irrelevant speech without meaning (see Martin et al., 1988; Smith, 1989; Oswald et al., 2000).

---

[5] White noise is defined as 'the effect of the complete range of audible sound-wave frequencies heard simultaneously, analogous to white light, which contains all the frequencies of the light spectrum' (*Encyclopedia Britannica,* 2007).

A recent study (Boman, Enmarker, & Hygge, 2005) summarizes the data of three separate previous experiments (Hygge, Boman, & Enmarker, 2003; Boman, 2004; Enmarker, 2004) performed with the same experimental design. The results show that, in the reading comprehension task, cued recall and recognition are more impaired by meaningful irrelevant speech than by road traffic noise. Moreover, the obtained noise effects are not related to the participants' capacity to perform the task; indeed, there is no interaction between the noise and age groups.

There are at least three different possible explanations to these results:

1) The semantic character in speech interferes with the semantic content in the material to be processed, especially for cognitive complex tasks. Indeed, irrelevant speech could be regarded as an additional task, making the situation comparable with divided attention situations where multiple stimuli must be processed simultaneously (Smith, 1985; Oswald et al., 2000).

2) The critical aspect of the noise causing irrelevant speech effects is neither its meaningfulness nor its linguistic character but the auditory signal that exhibits distinct acoustic variation or changes between its components (Jones, Madden, & Miles, 1992).

3) An auditory verbal background interferes with the retention of visually presented text, because under normal conditions both types of material have access to the *phonological store*[6] (Baddeley, 1986, 1990); therefore, one may impede the other. Thus, the irrelevant speech effect occurs at the level of the phonological store and is unrelated to the acoustic quality of the stimuli (Colle, 1980) and the semantic content of the verbal noise (Salamé & Baddeley, 1982). However, "recall is affected by auditory stimuli that include a verbal component, as is the case with speech or songs, but it remains intact when the auditory stimuli are nonverbal, as is the case with white noise" (Rouleau & Belleville, 1996, p. 357).

To summarize, irrelevant speech is more disruptive in a cognitive task than white noise and road traffic noise.

Music

Some studies suggest a disturbance of background music to message processing (e.g., Furnham & Bradley, 1997). According to our hypotheses, this effect is related to the kind of music. With classical symphonic music background, there should not be

[6] In working memory, the *phonological loop* includes a *phonological store* where verbal information is held in a phonological code. Since the phonological store is time-limited, a rehearsal procedure is used to prevent material from decaying. This procedure is also used to transfer the written digit into the phonological store (see Baddeley, 1986, 1990).

interference with the information search, given that there is no linguistic message to be processed. Italian music leads one to process the lyrics; such processing happens in a superficial way, according to the paradigm of *attenuated processing* (Eysenck & Keane, 1990). This processing is also an elaboration of lyric meaning and could interfere with the information search, subtracting attentional resources or complicating information maintenance in working memory.

Music itself should not interfere with the task, but the meaning present in the lyrics should. For this reason, in the present study we compare classical symphonic music by different composers (Haydn, Mozart, and Beethoven) with songs by modern Italian songwriters (De André, Battisti, Guccini, and De Gregori). Italian is the language of the participants.

Predictions

It is reasonable to anticipate that classical symphonic music improves the performance of information searching, because it suppresses meaningful irrelevant speech. Indeed, the condition without music is characterized by the presence of meaningful irrelevant speech, because the experimental setting is a crowded cafeteria and, therefore, many linguistic messages are present. On the contrary, we expect that Italian music impairs the performance of information searching, because there is only one linguistic message that is very clear. Therefore, the condition with Italian music is expected to produce more impairment than the condition without music, which is expected to produce more impairment than the condition with classical symphonic music.

2.3 Hypotheses

There are five experimental hypotheses in this study:

H_1: Participants' performance (accuracy and speed) declines when the text length increases. Moreover, participants should have more difficulty finding information in the middle and in the end of web pages than at the beginning. These findings should be more extreme in the Italian music condition.

H_2: Participants should exhibit fewer errors when they search for information while listening to classical symphonic music than while searching without music.

H_3: Participants should take less time to perform the same tasks when they search for information while listening to classical symphonic music than when searching without music.

H_4: Participants should exhibit more errors when they search for information while listening to Italian music than in other conditions.

H_5: Participants should take more time to perform the same tasks when they search for information while listening to Italian music than the other conditions.

Hypotheses 2 and 3 suppose that classical symphonic music improves information searching, while hypotheses 4 and 5 suppose that Italian music decreases performance.

2.4 Participants and Location

The experimental subjects are 24 Italian native speakers (10 males and 14 females), with varying educational backgrounds, ranging from 18 to 38 years of age. All participants have normal or corrected-to-normal vision. They were unaware of the purpose of the experiment. After the experiment, some participants received food incentives for their participation, since the environment was a cafeteria.

The participants were randomly assigned to one of three independent groups: without music, with classical symphonic music, or with Italian music. Each group consists of 8 participants.

The experiment was conducted in a comfortable, well-lit cafeteria in Padua centre during lunch time to ensure that participants would be exposed to similar levels of noise. Participants used a Qteck 9000 Pocket PC. Each participant had time to be comfortable with the mobile device. The Pocket PC was held in their hands or put on the coffee table and used with a stylus pen. Answers were given using the PDA keyboard. Participants listened to the music with earphones.

2.5 Materials

Texts

The 800-word article “Fornarina, Velata o Galatea” by Corriere della Sera (Bonazzoli, 2006) was gradually shorted by eliminating words that did not interfere with the overall meaning of the article itself to obtain the other three texts of different lengths (650, 550, and 450 words). According to the Flesch-Vacca scale[7] (Vacca & Franchina,

[7] To evaluate the *readability* from the syntactic point of view, you consider the length of some linguistic units (sentences and words), that make up the text, as a factor that indicates how difficult a reading passage is to understand. The sentences and words length is used as a linguistic variable in many statistic *readability index*. One readability index is a mathematical formula that, thanks to a statistic calculus, predicts the text complexity, using a specific scale of values.
In our study we use the *Flesch index* (Flesch, 1948). To apply this index you only need the average sentence length (total words/total sentences) and the average number of syllables per word (total

1986), all four texts can be considered as *easy*. In other words, they do not differ in content or readability (Flesch, 1948).

Each text was loaded on a web page. The background of the web page was a texture, whose principal color was light grey (#EAEBE6). A 14-pt Sans-serif Arial font , was used for all words except the title, which was 14-pt Arial in bold.

Questions

Since “the searching behavior supported by PDAs tends to be one in which the user seeks specific information” (Kim & Albers, 2001, p. 193), 20 clearly articulated questions requiring text answers were developed from the article. We ask the same questions for each text and compare the accuracy and performance time.

Key information

The key information is the right answer to the question; this information can be located in the beginning, middle, or end of the body of the text. Key information located at the beginning is the only condition in which the participants do not have to scroll multiple screens in order to view it. Therefore, information located at the beginning of the text might be easier to find than information in the middle or the end (see Kim & Albers, 2001).

2.6 Procedure

Participants met the experimenter individually in the cafeteria. At the beginning, they were given the PDA, were informed of the equipment and the procedure, and were asked to sign a consent form. The task consisted of answering 20 questions about a text appearing on their PDA by typing them in a specific web page. Participants were also informed that the PDA would record their performance time, starting from the appearance of each question and stopping when the answer was entered.

The web site appearing on the PDA with Internet *Explorer* started with a short questionnaire collecting demographic and other information (i.e., age, gender, education, computer and PDA user experience, habit of listening to music while working or studying). Afterwards, a second xhtml page appeared, composed of two frames: the article appeared in the superior frame, and a start link appeared in the

---

syllables/total words) in a sample of 100 words. The Flesch index has been adapted to the Italian language in 1972 by Roberto Vacca and after that readapted by Vacca and Franchina (1986).
The following data show the scores for each text employed in our experiment:

Flesch-Vacca index (450 words) = 68	*Flesch-Vacca index* (550 words) = 69.2
Flesch-Vacca index (650 words) = 61.6	*Flesch-Vacca index* (800 words) = 67.6

According to the *Flesch-Vacca scale*, from 60 to 70 the text is defined as *easy*.

inferior frame. When the participants clicked on this link, the first question appeared in the inferior frame. For each remaining question, two new links appeared: answer and next.

Participants read the question and began searching through the article for the answer. When the participants found the answer, they clicked on the answer link. This opened a web page where they could enter their answer and their related confidence level on a Likert scale from 1 ("not at all") to 5 ("absolutely self confident"). Once they closed this web page, the PDA stored their answers automatically. This process continued for all twenty questions; the experiment took about 35 to 45 minutes per participant, and the researcher was present for the entire experiment.

The questions were presented in random order. Participants answered the questions while listening to classical or Italian music or without listening to any music. For the first condition, music started when the participants opened the second xhtml page. At the end, participants filled in a final questionnaire to evaluate the accuracy of their answers and the principal disadvantages of the medium, layout, and text. Participants in the condition with music also indicated if the music helped or hindered their information search.

2.7 Measures

The results were measured in a number of ways: correct/incorrect answers; participants' confidence levels for each question, measured by a five-point Likert scale; and performance time in seconds.

The participants' answers were recorded in an online database, and the accuracy of the answers was controlled by the researcher, who calculated the number of errors for each participant. The self-confidence level was automatically stored in the online database as a numeric value.

Concerning the first question, timing began once the participants click on the start link and ended when they clicked on the answer link. For all other questions, timing began once the participants clicked on the next link and ended when they clicked on the answer link. In this way, the performance time does not include the time employed by the participants to write the answer, which depends on the typing rate and does not have any relation with the searching task. Instead, the performance time is strongly influenced by the users' scrolling behavior, reading rate, capacity of working memory, and information search strategy.

3. Results and Discussion

	Text length (in words): 450			550			650			800		
	wm[a]	cm[b]	Im[c]	wm	cm	Im	wm	cm	Im	wm	cm	Im
Average error	1	0.5	0.5	2	1.5	1.5	2	0.5	2	1.5	1.5	1.5
Average time	63	47	53	68	56	60	77	61	69	89	76	81
Standard deviation	24	18	21	25	22	23	24	22	24	27	23	24

Note. [a] without music. [b] with classical symphonic music. [c] with Italian music.

Table 1. Average error, Average time and Standard deviation across background conditions, without music (wm), with classical music (cm) and with Italian music (Im), for the different lengths of the text.

Location of the key information:	Text length (in words): 450			550			650			800		
	wm[a]	cm[b]	Im[c]	wm	cm	Im	wm	cm	Im	wm	cm	Im
Beginning	62 (19)	36 (12)	47 (19)	54 (16)	51 (20)	63 (21)	59 (17)	54 (14)	58 (17)	74 (6)	59 (15)	78 (27)
Middle	60 (22)	50 (17)	53 (19)	71 (27)	52 (17)	58 (22)	81 (24)	56 (21)	66 (15)	85 (22)	84 (24)	76 (16)
End	77 (27)	62 (21)	65 (22)	83 (31)	81 (18)	66 (27)	91 (15)	77 (22)	90 (29)	117 (31)	85 (17)	95 (25)

Note. [a] without music. [b] with classical symphonic music. [c] with Italian music.

Table 2. Average time and (Standard deviation) across background conditions, without music (wm), with classical music (cm) and with Italian music (Im), for the different positions of the key information.

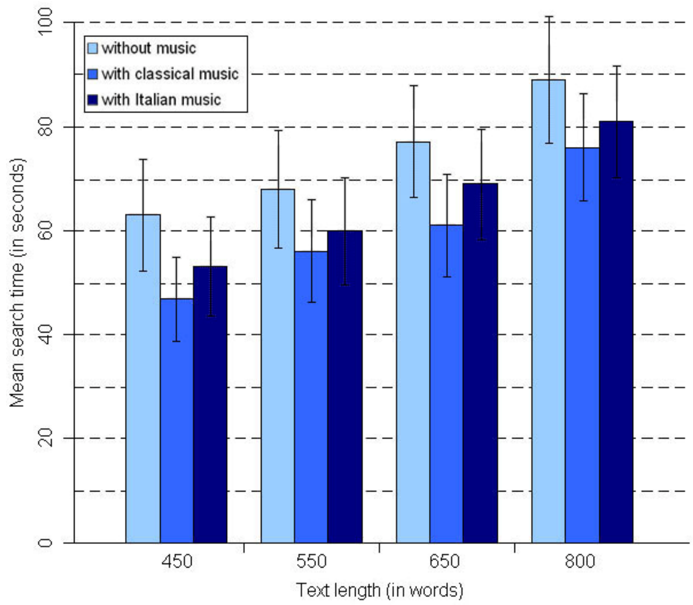


Figure 1. Mean search times across background conditions for the different lengths of the text.

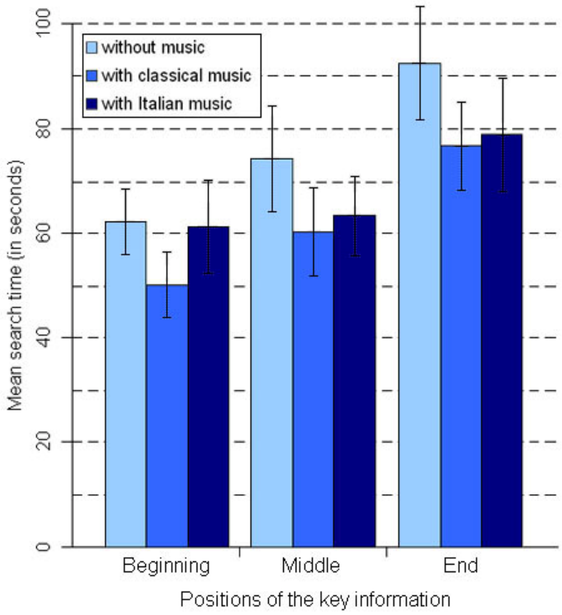


Figure 2. Mean search times across background conditions for the different positions of the key information.

3.1 Hypothesis 1

A multivariate analysis of variance (MANOVA) showed that the text length significantly influenced information searching ($F_{6,24} = 5.69$, $p < .05$). At a univariate level, accuracy was not influenced by the text length, whereas speed was significantly influenced; for all background conditions, when the text was rather long, the searching time increased (Figures 1 and 3). This suggests that the amount of text presented on the PDA affected the users' ability to retrieve information but did not prevent users from finding the right answers.

The background condition also significantly influenced performance ($F_{4,24} = 5.44$, $p < .05$). There were no significant interactions by MANOVA between the length of the text and the background condition.

We expected that participants should have more difficulty finding information in the middle and in the end of web pages than at the beginning. Time performance (Table 2) supported this assumption: a significant main effect of the location of the key information was found by repeated ANOVA measures across the three experimental conditions ($F_{2,42} = 89.72$, $p < .05$). Moreover, pair-wise comparisons reveal that searching time for information significantly increased when the key information was placed at the end of the web page rather than at the beginning or the middle ($p < .001$ for both).

Contrary to our expectations, the worst condition was searching for information without music and not with Italian music, in particular when the key information was placed at the middle and at the end of the page (Figure 2). No significant interaction effect was found between background condition and location of the key information.

3.2 Hypothesis 2

The level of accuracy was overall quite high (see Table 1), and the differences between the two conditions, with classical symphonic music and without music, were not significant by ANOVA. Thus, the second hypothesis is not supported.

3.3 Hypothesis 3

Results confirm our third hypothesis: classical symphonic music significantly improved the performance of the participants, reducing mean times to perform the tasks. Indeed, the differences in mean times between the two conditions, with classical symphonic music and without music, were significant by ANOVA ($F_{1,14} = 6.42$, $p < .05$). By suppressing meaningful irrelevant speech and isolating the subject from the surrounding world, classical symphonic music decreased distraction and enhanced attention, concentration, and arousal. According to the Yerkes-Dodson Law (Yerkes & Dodson, 1908), an optimal level of arousal causes an optimal level of performance.

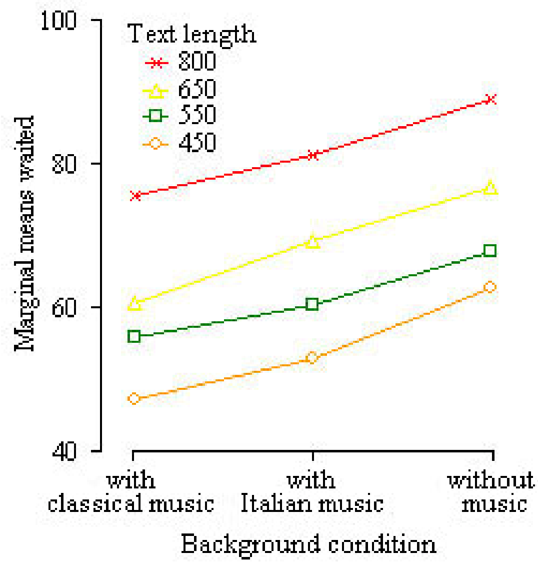


Figure 3. Marginal means of time.

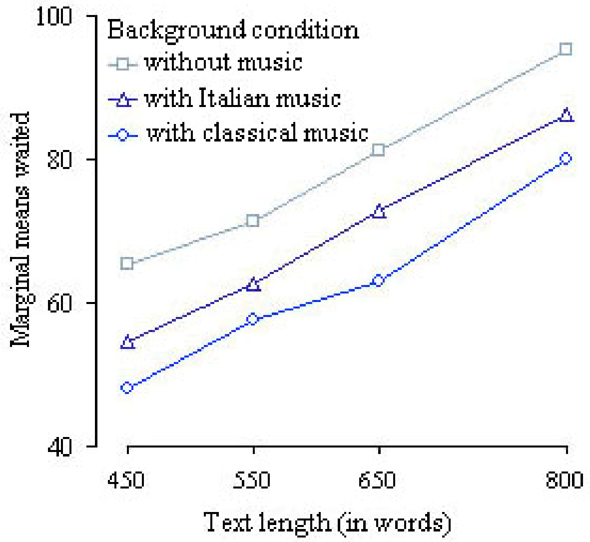


Figure 4. Marginal means of time.

3.4 Hypothesis 4

The fourth hypothesis was not supported by the results of the experiment. No significant differences by ANOVA were found for accuracy scores between the two conditions, with Italian music and without music; likewise, differences in error rate between the two conditions, with Italian music and with classical symphonic music, were not significant.

3.5 Hypothesis 5

Differences in mean times between the two conditions, with Italian music and without music, or with Italian music and with classical symphonic music, were not significant by ANOVA. Participants did not spend more time answering questions while listening to Italian music; while listening to either Italian or classical music, performance was better than without any music (Figures 1 and 4). Italian music reduced mean search times and improved performance, although not in a significant way.

This outcome contrasts with our initial hypothesis that the linguistic message present in the lyric would impair performance. A possible explanation is linked to the Yerkes-Dodson Law. Music increases the perceived social richness of the PDA and therefore the level of pleasure and arousal (Kallinen, 2004). Increased arousal corresponds with increased performance efficiency, which compensates for the impairment caused by the attenuated processing of the lyric.

Figure 4 shows that Italian music, which contains linguistic messages, is placed between classical symphonic music, which suppresses background conversation, and without music, where background conversation is present. Perhaps, participants were distracted by the lyrics; however, this needs to be investigated in future studies, where also the effect of their level of familiarity with the songs is analyzed.

3.6 Likert Scale Point

Our study finds that the participants had good self evaluation ability: a high Likert score corresponds with a correct answer, while a low Likert score corresponds with a wrong answer. In the final questionnaire, almost all participants (about 92%) said that they were globally confident of their answers. Moreover, a high Likert score corresponds with key information placed at the beginning of the text, while a low Likert score corresponds with key information placed at the middle or end of the text. Indeed, a low Likert score often corresponds with longer performance time. This means that the subjects were less confident of the answer when they took more time to find it.

3.7 Final Questionnaire

In the final questionnaire, the participants considered the sensitivity of the stylus pen and the small dimension of PDA screen and keyboard as the principal disadvantages of the medium employed. The small character and page homogeneity were considered the principal disadvantages of the layout. The participants suggested that the main disadvantages in the text were the complexity of syntax and the lack of hierarchical layers, such as subtitles for each paragraph. A possible solution would be short and concise blocks of information, called *chunks*. The principal features of chunks are topic sentences placed in the opening position and the use of the upside-down pyramid method to give information (conclusion followed by supporting material). A web page organized in chunks is not syntactically complex and helps the reader access information quickly (Morkes & Nielsen, 1997); however, an excess of chunking could fragment the ideas expressed in the text and impair the user’s ability to understand the information (Chu, 2001).

4. Conclusion

Some people in a crowded place, where the noise level is high, might think that listening to music improves concentration and helps us perform tasks more easily than intermittent background noises. Instead, in the present study, 45% of subjects believed that music had compromised their performance; 35% said that music had no influence; and only 20% felt that music improved performance. The results of the experiment show that both classical symphonic music and Italian music improve searching ability. In particular, this study shows that participants who listened to classical symphonic music while searching for information on PDA in a noisy environment had a better performance.

5. References

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Baddeley, A. D. (1990). *Human memory: Theory and practice.* London: Erlbaum.

Boman, E. (2004). “The effects of noise and gender on children's episodic and semantic memory. *Scandinavian Journal of Psychology*, *45*, 407-416.

Boman, E., Enmarker, I., & Hygge, S. (2005). Strength of noise effects on memory as a function of noise source and age. *Noise & Health*, *7*, 11-26.

Bonazzoli, F. (2006, May 9). Fornarina, Velata o Galatea Appare sempre l’amore segreto. *Corriere Eventi*, p. 5. Retrieved July 6, 2006 from http://www.corriere.it. Chu, S. (2001). The Possibilities Are Wireless: Designing and Delivering Information in the Wireless Space. *Technical Communication*, *48*, 49-58.

Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial locations. *Journal of Verbal Learning and Verbal Behavior*, *19*, 722-735.

Enmarker, I. (2004). The effects of meaningful irrelevant speech and road traffic noise on teachers' attention, episodic and semantic memory. *Scandinavian Journal of Psychology*, *45*, 393-405.

Eysenck, M. W., & Keane, M. T. (1990). *Cognitive Psychology. A student handbook.* Hove: Erlbaum.

Flesch, R. F. (1948). A new readability yardstick. *Journal of applied Psychology*, *32*, 221-233.

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extroverts. *Applied Cognitive Psychology*, *11*, 445-455.

Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scandinavian Journal of Psychology*, *44*, 13-21.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech: The role of changing state. *Quarterly Journal of Experimental Psychology*, *44A*, 645-669.

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory?. *Applied Cognitive Psychology*, *4*, 89-108.

Kallinen, K. (2004). The Effects of Background Music on Using a Pocket Computer in a Cafeteria: Immersion, Emotional Responses, and Social Richness of Medium. *CHI'04 extended abstracts on Human Factors in Computing Systems*, 1227-1230.

Kim, L., & Albers, M. J. (2001). Web Design Issue when Searching for Information in a Small Screen Display. *Proceedings of the 19th annual international conference on Computer documentation*, 193-200.

MacKay, B., Dearman, D., Inkpen, K., & Watters, C. (2005). Walk 'n Scroll: A Comparison of Software-base Navigation Techniques for Different Levels of Mobility. *Proceedings of the 7th international conference on Human computer interaction with mobile device & services*, 183-190.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language*, *27*, 382-398.

Morkes, J., & Nielsen, J. (1997). *Concise, scannable, and objective: How to write for the Web*. Retrieved April 23, 2007, from http://www.useit.com/papers/webwriting/writing.html.

Oswald, C. J. P., Tremblay, S., & Jones, D. M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory*, *8*, 345-350.

Rouleau, N., & Belleville, S. (1996). Irrelevant speech effect in aging: an assessment of inhibitory processes in working memory. *Journal of Gerontology: Psychological Sciences*, *51B*, 356-363.

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implication for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*, 150-164.

Schaffers, H., Brodt, T., Pallot, M., & Prinz, W. (2006). *The Future Workspace. Prospective on Mobile and Collaborative Working.* Telematica Institut.

Smith, A. (1985). The effects of different types of noise on semantic processing and syntactic reasoning. *Acta Psychologica*, *58*, 263-273.

Smith, A. (1989). A review of the effects of noise on human performance. *Scandinavian Journal of Psychology*, *30*, 185-206.

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. (2000). The irrelevant sound effect: Does speech play a special role?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *6*, 1750-1754.

Vacca, R., & Franchina, V. (1986). Taratura dell’indice di Flesch su testo bilingue italiano-inglese di unico autore. *Linguaggi*, *3*, 47-49.

Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, *18*, 459-482.