The University of Sheffield
Department of Information Studies

Re-thinking Lifelogging: Designing Human-Centric Prosthetic Memory Devices

by
Vaiva Kalnikaitė

A thesis presented in the fulfilment of the requirements for the degree of

Doctor of Philosophy at the University of Sheffield

May 2009

Dedicated to the memory of my grandparents

Adolfina Papievienė & Antanas Papievis

Abstract

Building Prosthetic Memory (PM) technology has been an active research area for the past few decades, with the primary aim in supporting Organic Memory (OM) in remembering everyday events and experiences. Through building and evaluating new PM tools, this thesis attempts to explore *how* and *when* PM tools are used to help OM in everyday memory tasks.

The focus of this thesis is to investigate PM tools as an extension of, or a supplement to, OM and to understand *why* people choose to use PM as opposed to their OM to help them retrieve information. Further aims of this thesis are to investigate the role of Metamemory and social processes. Finally, the work aims to support Autobiographical memory through building new PM tools.

The studies apply mixed experimental and naturalistic methods, and include 3 controlled lab studies and 3 field trials involving a total of 217 participants. Overall, there were 5 new PM devices built and evaluated in long-term and controlled contexts.

Results obtained through lab studies suggest that PM and OM function in a synergetic relationship. In particular, use of PM increases when OM is particularly weak and this interaction is mediated by organic Metamemory processes. PM properties also have an influence – people prefer efficient over accurate PM devices. Furthermore, PM cues help in *two* ways: 1) at encoding to help focus OM; and 2) at retrieval to cue partially remembered information.

Longer term studies also reveal that PM is not used to substitute for OM. Instead users prefer to use recordings to access specific parts of a lecture rather than listen to the whole thing. Such tools are extensively used by non-native speakers, although only native speakers' coursework benefits from usage. PM tools that support social summarisation demonstrate that people exploit social feedback and cues provided by other users and that these improve recall.

Finally, evaluations of new autobiographical memory tools show that people upload mementos based on their importance. There is evidence for preference for mementos that are associated with other people and home.

I conclude with a discussion of the design and theory implications of this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Organic Memory (OM) is crucial in everyday remembering, and the fallibility of OM has been much studied and well understood (Schacter 1987; Schacter 1997; Baddeley 1998; Anderson 2004). With the proliferation of mobile and ubiquitous technologies, it is now possible to imagine capturing vast amounts of personal information about our past on an external device in analogue or digital form. This concept we call Prosthetic Memory (PM). Our daily lives are full of to-dos and shopping lists, diaries and post-it notes as well as mobile phones, blackberries, PDA and wearable sensing devices. All of this technology has a common purpose: to aid Organic Memory in everyday remembering tasks.

Everyday memory experiences can be varied, for instance, going to meetings and remembering conversations, attending university, and learning new materials or collecting memorabilia and remembering personal events from the past. PM tools to support any of these domains have to be designed not only from the Computer Science field of view, but also consider Psychology research in OM and Human Computer Interactions (HCI) – dedicated to the natural facilitation of information transfer between humans and machines, synthesising these diverse research fields.

There is little evidence, however, of how these research areas might interact in order to design new usable PM tools. Previous PM research has tended to focus on technical innovation rather than understanding when and why these tools are used. The research presented in this thesis attempts to bridge these research areas in order to address *four* related research questions:

1) What factors determine the use of PM tools and how does this relate to OM?

2) What different strategies do people use with PM tools and what is the effect of these strategies on OM?

3) What are the Long-term benefits of using PM tools, who uses them, why do they do so and what benefits do they derive from them?

4) How can we design better PM tools to support Autobiographical memory?

This chapter presents both an overview and challenges common to these areas, introduces the motivation for this work, and establishes the contribution of the thesis.

1.1 Overview

Prosthetic Memory tool building has been popular in the past few decades. However, as yet there have been few attempts to situate these tools in the context of psychological research on memory or to explore empirically how these tools are used. The psychology of OM has a lot to offer in understanding how these tools should function in order to help OM.

There has been a recent proliferation of mobile and ubiquitous PM technologies such as mobile phones, PDAs and wearable devices intended to support OM in retrieving experiences, conversations and personal information. Bush's original *Memex* vision (Bush 1945) for Prosthetic Memory has been well explored with the growth of new OM support tools (Whittaker, Hyland et al. 1994; Karger and Quan 2004; Hodges, Williams et al. 2006; Vemuri, Schmandt et al. 2006). However the goals of these tools often seems to be to replace OM with PM rather than complement it (Gemmell, Bell et al. 2006). There is still little understanding as to what leads users to choose PM tools over their OM for capturing and retrieving specific information and events.

Further, despite the availability of all this technology, there has been relatively little effort in investigating how these tools are used in long-term, and in real settings, in terms of the benefits they might bring to real users.

In terms of autobiographical memory, there has been substantial research in material cultures and the meaning of memorabilia (Czerwinski, Horvitz et al. 2004; Petrelli, Whittaker et al. 2008). However there are few tools that investigate how these autobiographical mementos could be captured and organised digitally.

1.2 Motivation

Everyday people are surrounded by more technology, there is greater potential to capture everyday events and experiences more easily and move frequently. Furthermore, storage search and capture technologies are improving rapidly. However, what determines the usage of these PM devices and how does their usage relate to people's unaided OM?

These PM tools can provide prolonged capture, over long periods of time. Thus there is a need to investigate how people use these tools and how the information generated can help OM over long-term in contexts where it is important for them to remember. To examine this I looked at student's use of multimedia recording tools in a real-life learning situation.

Finally, autobiographical memories have a profound influence on individual identity, but there are virtually no tools that support digital capture and organisation of this type of data. I therefore designed and critically assessed a new tool intended to capture Autobiographical memories digitally.

1.3 Statement of Thesis

In this thesis I show that over the long-term people can retrieve more information with digital PM rather than their OM or analogue PM. I also show that if there is low user confidence in OM there is a tendency to use PM more. Furthermore PM use depends on the properties of the PM tool. Users prefer fast and inaccurate PM searches over accurate but less efficient retrieval. I also investigated strategic use of these tools, finding that exhaustive note-taking on a PM device where notes can be used to retrieve speech recordings distracts rather than focuses. Thus users who strategically decide what to capture perform better in memory tasks overall.

Over the long-term PM tools offer benefits for remembering: users tend to do better in their coursework when using new digital multimedia augmented PM tools, performing better with PM tools compared with using analogue traditional tools such as pen and paper.

In personal autobiographical remembering, I have shown that important mementos tend to focus on people rather than things or places. I have also shown that while previous work has argued for the value of narration in social settings, users preferred more lightweight textual methods for annotating mementos.

1.4 Contributions

As part of the larger research effort in the HCI community to build effective PM technologies, my research focuses on both building new PM tools as well as providing useful further understanding into *how* and *when* people exploit these PM tools. My strategy in pursuit of this goal has been to build and evaluate these novel PM tools with real users over long-term Testing Sections in the context of everyday memory for conversations, learning and autobiographical memorabilia.

The contributions of this thesis are the following:

Conversation-based PM – I first conducted a short-term study evaluating *when* and *why* people use PM tools. As well as building and evaluating a new human-centric PM tool for this study, I have determined critical PM and OM attributes that influence PM choices. I have also established the benefits of using PM. Furthermore, non-strategic PM usage techniques distract the user rather than providing a focus.

Learning-based PM – There have been few long-term studies evaluating PM tools that support OM over the long-term. I have implemented four versions of a novel PM tool that have been deployed in real classrooms to evaluate their effectiveness. Use of PM improves student performance and course results.

Context-based Autobiographical PM – Finally I carried out a field trial evaluating the capture and organisation of personal memorabilia. This study looks at how PM tools can be applied to the domain of personal autobiographical memories. PM tools that support this domain require an understanding of what *types* of personal mementos people would want to capture and *why*. There are factors such as emotion, importance and context to be considered before understanding participant memento choices. In order to answer these questions, I built and evaluated a novel PM tool for capturing and organising personal memorabilia.

1.5 Structure of the Thesis

Each research contribution is introduced in a separate chapter. Chapter 2 introduces and outlines the topics addressed in the thesis, by surveying related work. In addition, existing research that specifically relates to each contribution is presented in the first section of the corresponding chapter. Chapter 3 outlines the methodology followed in the design and various user evaluations of the new PM tools.

1.5.1 Research Design

The first step toward addressing the issues stated in Section 1.3 was to identify the drawbacks of current PM tools and how current findings in OM from Psychology might help improve future PM design. Following this, I designed, implemented and evaluated a series of novel PM tools to help OM in specific domains of remembering conversations, lecture materials and personal autobiographical mementos. Based on these evaluation results, I report further knowledge on *how* and *when* people rely on PM tool use in these domains and also propose future PM design principles.

1.5.2 Outline of the Proposed Research

The thesis is structured as follows:

Chapter 2: Overview of Research in Organic and Prosthetic Memory. This introductory chapter presents an overview of psychology research into OM and provides a review of current PM tools that are aimed at supporting OM. It identifies the drawbacks in current PM tool design.

Chapter 3: Methodology. This chapter gives an overview of the methodology used in requirements gathering, design and an evaluation of systems. All the PM tools presented in this thesis have been built according to iterative design principles, evaluated and analysed using a combination of methods presented in this chapter.

Chapter 4: ChittyChatty: When Do People Use Prosthetic Memory to Augment Organic Memory? The research presented in this chapter aims to understand and evaluate the benefits of using PM tools for retrieving everyday conversations over different time intervals. As part of the research into improving PM tools, several new design principles are proposed here that address the problem of aiming to replace OM with PM.

Chapter 5: Digital Notes: How and Why Do They Cue Memory? This chapter explores the effects of different note taking strategies and how these relate to memory. It compares digital and traditional analogue note taking strategies as well as evaluating what note taking techniques cue OM better.

Chapter 6: Augmented Digital Records to Support Organic Memory for Learning. This chapter investigates how PM tools are used in a large scale long term real learning environment, where demand for memory is very high. An in-situ study revealed an added benefit of PM in the form of augmented digital records for lectures and suggested new refinements to learning based PM tools.

Chapter 7: Social Summaries to Augment Organic Memory for Learning. Further to Chapter 6, I devised and deployed new types of PM in the classroom for user-centric evaluations. I explored web 2.0 type social summarisation tools and their effects on memory for learning new materials. The new PM tools summarised lectures based on the popularity of user access, revealing that students who access these summaries do better in their course.

Chapter 8: MemoryLane: Contextualising Autobiographical Mementos. In this chapter, I explored a new direction by looking at what types of mementos people digitise. I designed and implemented a novel PM tool for managing personal digital mementos. The new PM tool was evaluated with a group of users and was preferred over traditional tools currently available for memento management. There was also a bias towards people – centric mementos and a dislike for speech narratives associated with mementos.

Chapter 9: Contributions, Conclusions and Future Work. The final chapter contains further discussion of contributions and final conclusions as well as presenting future directions stemming from this work.

1.5.3 Main Questions Evaluated

1) *When* and *why* do people choose to use PM tools instead of relying on their OM?

- Since PM tools capture exhaustive and accurate records of everyday conversations, why do people sometimes still choose to retrieve this information from their OM?
- What influences this decision?
- How do different PM strategies affect recall?

2) Do PM tools help OM over long-term?

- How do people use PM tools over the long-term and how do they help recall?
- Who uses these tools and what are the measurable benefits of using PM over long-term?

3) PM tools for Capturing and Contextualising Autobiographical Mementos

- What types of physical mementos do people digitise and how frequently?
- How do people choose what to digitise?
- How do people verbally annotate their memorabilia and associated emotions?

Chapter 2

Overview of Psychology Research in Organic Memory and Review of Prosthetic Memory Tools

2.1 Introduction

'The serious difference between computer and human memory is that we don't pop out a pristine copy of the original event... We guess." (Loftus and Calvin 2001).

Organic memory is in use daily to help with simple tasks from grocery shopping to remembering the location of a misplaced house key. In the long-term it contributes to our sense of self and our social relationships. Often Organic Memory (OM) uses the strategy of pursuing pathways to individual internal reminders and building extensive personal collections of memories, which in turn shape individual identities. However our awareness of the fallibility of OM motivates that memory storage be outsourced to a Prosthetic Memory (PM) tool. PM tools often utilise external prompts to aid OM retrieval (e.g. a yellow sticky note left by our computer, or a note in our diary). The goal of these external reminders is to help retrieve everyday prospective and retrospective events and people make frequent use of shopping lists, to-do lists; diaries or journals.

These PM tools focus future recall of OM through cataloguing the past as well as reminding about future actions.

2.2 Prosthetic Memory

The aim of this chapter is to give an overview of existing research in OM. Research into OM has been extensive and diverse, spanning decades, most of which is outside the scope of this thesis. However, I will provide an overview of OM functions and OM studies that can directly inform the design and evaluation of PM tools, that effectively support OM.

This chapter also provides a brief overview of current PM capture tools. I focuse on understanding how these tools fit within the general PM framework on the relationship between OM and PM.

OM is fallible, and this leads people to actively prepare for these failures by using external PM tools, e.g. pen and paper, PDAs, cameras and other specialised devices. I define PM as being the constellation of tools that extend the capacity of OM that have the potential to capture important mnemonic information. There are two main forms: *active* Prosthetic Memory and *passive* Lifelogging.

Active PM tools can use either: analogue, digit or hybrid interactions. For example, pen and paper uses analogue representations, but newer digital tools e.g. Whittaker's Filochat (Whittaker, Hyland et al. 1994) augment digital recordings with digital annotations. In contrast, Stifelman's Audio Notebook (Stifelman, Arons et al. 2001) uses hybrid augmentation methods: digital voice recording annotated with analogue paper based handwriting. *Active* PM tools require *Intentional* capture where people make *a strategic decision* to capture the information being presented to them.

In contrast Lifelogging tools employ *passive* capture – without active intervention from the user, often engaging sensor triggers e.g. SenseCam (Hodges, Williams et al. 2006) or EyeTap (Mann, Fung et al. 2005). These tools perform *non-delierate* capture of everyday events. People do not have to make a decision as to what is being captured, e.g. they put on a wearable piece of technology that captures relevant memory data when automatic triggers become actively engaged.

However, the predominant questions for either of these forms of capture are: how well do *intentional* or *unintentional* capture tools support OM overall?

With a few exceptions (Whittaker, Hyland et al. 1994; Sellen, Fogg et al. 2007; Harper, Randall et al. 2008) there has been little work on systematically evaluating these tools. In particular:

1) When and why would people choose to capture everyday events in this manner? and

2) How would people exploit these tools to help them re-construct their memory of the past?

2.3 Organic Memory Processes

Organic Memory (Baddeley 1998) is at the intersection of a number of diverse disciplines such as neuroscience, psychology, and human computer interaction (HCI). One key focus of this thesis is to apply the knowledge from psychology research of human memory functions to better understand the interactions between OM and PM in the area of HCI. The first part of this chapter reviews Organic Memory as it is presented in psychology research and what processes drive organic long-term remembering.

According to Schacter - *"Memory is the retention of, and ability to recall, information, personal experiences, and procedures such as skills and habits."* (Schacter 1997).

There is no universally agreed model of the mind according to Schacter (Schacter 1997) and there is no agreement on how the memory works overall. Other memory models are based on cognitive psychology and assume that the brain operates like a computer (Johnson-Laird 1988), but these are rejected by Schacter because they cannot account for the subjective and present-need just-in-time basis of memory. Subjectivity in remembering, Schacter says, involves at least three important factors:

1. Memories are constructions made in accordance with present needs, desires, influences, and other factors;

2. Memories are often accompanied by feelings and emotions; and

3. Memory usually involves the rememberer's awareness of the memory - Metamemory.

According to Schacter, a good model of how memory works must not only fit with specific knowledge but also fit with the subjective nature of memory itself.

Memory debates amongst psychologists produced different theories and explanations of how memories might be produced, stored and reconstructed. However, most of those explorations have been conducted in a laboratory setting or based on the study of small numbers of brain damaged patients. Rather fewer naturalistic studies involve people being asked to remember more everyday materials such as conversational stories, films, personal events or maps (Linton 1982; Cohen, Kiss et al. 1993).

Next this chapter reviews the hierarchy of OM areas and discusses the individual functions of each of these areas.

2.4 Organic Memory Hierarchy

This section aims to introduce and discuss the hierarchical structures of organic memory. A new concept of Metamemory is also introduced as part of the hierarchy. This represents the decision process at conscious encoding and reconstruction of memories. Metamemory acts as an interactive process or a gateway guiding information search and evaluation. When the right information is found, Metamemory reassesses the relevance and suitability of this information in light of the original query. Metamemory is important in the context of this thesis, because it is the process that can motivate people to rely on PM as opposed to OM and that also controls the PM strategies that people use. Further detailed description of Metamemory is presented in the subsection below.

Figure 2.1 illustrates the location of Metamemory within the overall hierarchy of OM, affecting all areas of Organic Memory.

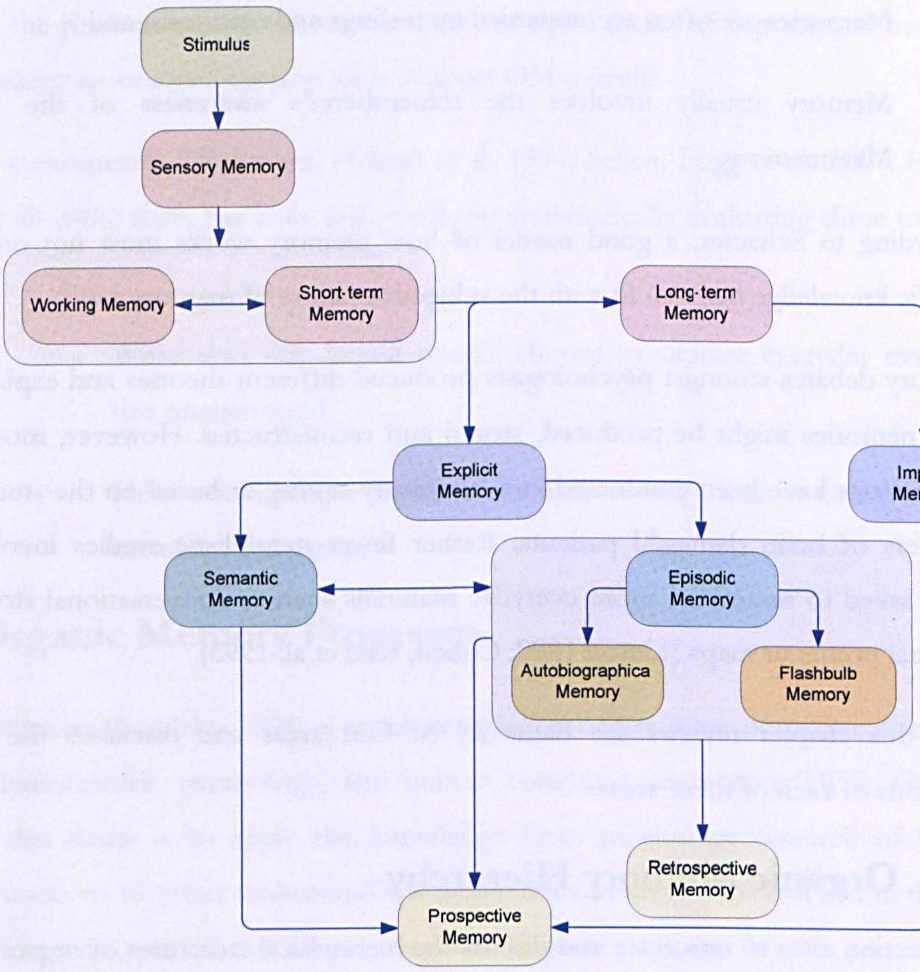


Figure 2.1: The Hierarchy of Organic Memory

Overall Organic Memory consists of various types of memories working in synergy. When a stimulus is presented, sensory memory evaluates which of the senses it is coming from: vision, hearing or touch. When sensory memory identifies the information, this information moves to working/short-term memory. With rehearsal or other reinforcement, this information then becomes part of long-term memory. Based on the type of information, it can be classified as Explicit – memory for facts (e.g. London is the capital on the UK), or Implicit – remembering activities such as driving a car.

Furthermore, Explicit memory in turn can be either Semantic – memory of general knowledge about the world or Episodic – memory of personal or world events. Episodic memory in turn can either be Autobiographical – personal events or Flashbulb – emotionally significant public events. A final distinction is between Prospective memory – memory for the future and remembering to do things, and Retrospective memory – memory for the past and remembering cues to past events. The main focus of this thesis is on Episodic memory in the context of conversations, learning and autobiographical memory.

The following subsection describes and discusses these individual memory areas presented in OM hierarchy in Figure 2.1.

2.4.1 Sensory Memory

Sensory memories are acquired through sensory organs such as vision, hearing or touch. This memory lasts milliseconds after which information propagates to short-term memory. Sensory memory consists of Visual, Auditory and Haptic memory. In Visual memory, also known as Iconic memory, (Sperling 1960) information is obtained through vision. In Auditory memory also know as Echoic memory (Cowan 1984) information is obtained through hearing. Haptic, or memory for touch, helps to interpret memory obtained through the touch and internal muscle tensions (Sperling 1960). Any of those senses can be activated individually or work in synergy with each other to interpret stimuli coming from these multiple senses. After the initial interpretation, the information propagates to short-term and working memory.

2.4.2 Short-term and Working Memory

Alongside short-term memory, there is the functioning of a working memory. This does not replace the short-term memory, but rather complements it for specific short-term memory tasks, for instance, small algebraic calculations. Short-term memory has the capacity to keep information for short periods of time and it can only contain "seven, plus or minus two"' slots of information (Miller 1956).

Short-term memory can hold either of the following:

- recently processed information from Sensory memory; or
- results of recently processed information, mainly from working-memory; or
- memories recently retrieved from long-term memory.

Baddeley & Hitch (Baddeley, Hitch et al. 1974) proposed a working-memory model with two short-term storage mechanisms: a phonological loop and a visuospatial sketchpad. The phonological loop processes *verbal* information whereas the visuospacial sketchpad works on processing visual and special short-term materials, illustrated in Figure 2.2.

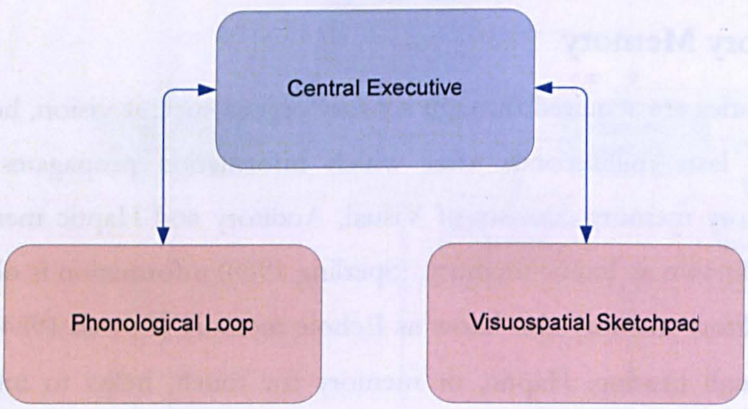


Figure 2.2: Working-memory model proposed by (Baddeley, Hitch et al. 1974).

Once information is processed in short-term memory, it can propagate to long-term memory.

2.4.3 Long-term Memory

After initial processing, information enters Long-term memory. It can stay there as long as a lifetime. The following situations contribute to long-term storage (Schacter 1997):

- Information which captured *attention*, thus received more deep processing;
- Information which is *rehearsed;* and

- Information with meaningful *associations.*

It is still uncertain about the exact biological processes that shift memories from short-term to long-term repositories and the actual capacity of the Long-term memory. Information is generally regarded as being encoded semantically in Long-term memory (Baddeley 1982).

The main focus of this thesis is primarily on Long-term memory. The following sections discuss individual Long-term memory types and their functions in more detail.

2.4.4 Explicit and Implicit Memory

Long-term memory consists of Explicit memory (Tulving and Schacter 1990) and Implicit memory (Schacter 1987). Explicit memory relies on conscious recall or otherwise requires the conscious storage and retrieval and reconstruction of information. Facts and experiences are stored in Explicit memory. Explicit memory can be further sub-categorised into Semantic and Episodic memory illustrated in Figure 2.1.

2.4.5 Semantic and Episodic Memory

- **Semantic** memory - concerned with facts independent of time and place context. Those facts generally relate to knowledge of the world e.g. most swans are white (Tulving 1972); and

- **Episodic** memory – concerned with memory for personal experiences, e.g. I went to London last week. It can include memory for specific events, people, places, times or emotional ties (Tulving 2002).

Semantic memory is related to general knowledge. Episodic memory is used for more personal memories that are sometimes linked to individual emotions, sensations and other personal associations with objects, for instance, the place where one attended school, or people you knew at university.

Episodic memory in turn can be further categorised into Autobiographical memory and Flashbulb memory (Sternberg 2006), as illustrated in Figure 2.1. Autobiographical memory is associated with specific episodic memories that shape one's life (Rubin 1988; Conway and Pleydell Pearce 2000; Rubin, Schrauf et al. 2003). These memories are similar to episodic memories but are of greater personal significance. Flashbulb memories are also of episodic memory type, but for public, highly emotional events (Brown and Kulik 1977), e.g. September 11th.

2.4.5.1 Semantic Memory

Semantic memory deals with general knowledge about the world rather than specific knowledge about personal experiences. The size of semantic memory seems large but a more important question relates to the storage and organisation of semantic memories.

Much human knowledge is inconsistent, incomplete, uncertain and vague. According to (Nickerson 1977) it is important to draw a distinction between fuzzy relational knowledge about events (X occurred before Y) and absolute knowledge (such as quantitative facts). Furthermore, since knowledge of the world for most people is inexact and vague, this level is generally enough for making decisions and actions in everyday life.

A great deal of knowledge is constantly demanded of us, some of which is known but impossible to retrieve, as shown by (Conway 1990; Koriat 2000) in their feeling of knowing effect e.g. things which we know to be true, but cannot remember. For instance, knowing that I was born in Lithuania, but I cannot remember this. Thus an important question concerns how people function with a memory system that appears to have so much vagueness.

2.4.5.2 Episodic Memory

"Episodic memory ... makes it possible for a person to be consciously aware of an earlier experience in a certain situation at a certain time ... the act of remembering a personally experienced event, that is, consciously recollecting it, is characterized by a distinctive, unique awareness of re-experiencing here and now something that happened before, at another time and place" (Tulving 2002).

Episodic memory focuses on recalling events. One of the most popular episodic memory theories is *schema theory* that has been extensively studied and used in everyday applications. Schema theory was first introduced by (Bartlett 1932) who claimed that a story is 'assimilated' to pre-stored schemas based on previous experience. Thus a schema is defined as a way of using past experience to deal with and understand new experience. Schema theory affects everyday memory in terms of memory selection and storage, abstraction, interaction and interpretation, normalisation and finally retrieval. However there are some ambiguities casting doubt on the schema activation time. When does a schema get activated - at encoding or at the reconstruction stage? Anderson *et al* (Anderson and Pichert 1978) conducted an experiment investigating which level schemas operate at. They found that schemas have some effect on retrieval, as well as on encoding, as their results revealed that a new schema given only at the retrieval stage produced additional recall. Other critics (Oldfield and Zangwill 1944) claim that schema theory is inconsistent with the accurate recall of certain events, in particular anomalous events are remembered that do not fit one schema.

Schema theory has been proved to be useful in explaining memory for scenes and events (Brewer and Tenyens 1981). In their experiments they discovered that participants were better able to recall a greater number of items with higher schema-expectancy ratings and fewer items were recalled with low schema ratings.

The popularity of schema theory led to further research in the field which resulted in some modifications to this theory. (Graesser and Nakamura 1982) proposed the *schema-plus-tag-model* which suggested that a memory representation of a specific event includes both the general schema and distinctive tags representing anomalous or unexpected aspects of the event.

Another modification to schema theory was proposed by (Schank 1982) introducing a hierarchical arrangement of memory representations called *memory organisation pockets* (MOPS). At the lowest level there are more specific representations of events. At the highest level those representations become more general and schema like.

Cohen *et al* (Cohen, Kiss et al. 1993) suggests the way that information is treated in memory falls into three possible categories relevant to the build-up of semantic and episodic memories:

1) Memories that are rapidly forgotten;

2) An episode might be recorded in semantic memory, contributing to the creation of a schema, where particular details about an event are lost, but the general scheme remains; and

3) An event may remain in episodic memory and can be remembered in specific detail, e.g. flashbulb memories.

2.4.5.3 The Relationship between Episodic and Semantic Memory

According to (Tulving 1972), real life semantic and episodic memories are interrelated. Tulving suggested that episodic memory is embedded within semantic memory. Everyday memory involves both kinds of memory interacting with each other. The distinction between episodic and semantic memory is closely linked to the schema theory, but there are a few differences illustrated in table 2.1 below.

Function	Episodic Memory – Autobiographical	Semantic Memory
Type of information represented	Specific events objects places and people - context bound	General knowledge and facts about events and objects-context-free
Type of organisation in memory	Chronological (by time of occurrence) or special (by place of occurrence)	In schemas (packets of general knowledge relating to the same topic)
Source of information	Perception personal experiences life events	Abstraction from repeated experiences generalisations learnt from others
Focus	Subjective reality (the self) has personal significance	Objective reality (the world) no personal significance

Table 2.1: Episodic Memory vs. Semantic Memory, adapted from (Cohen, Kiss et al. 1993) p.15.

According to (Linton 1982) over time, repetitive episodic memories decrease in strength while the Semantic base of those events increases. This means that some memories that started off as episodic memories became Semantic memories over time and with appropriate repetition. In other words episodic memories become more generic, until they become more like general rules.

A diary study investigating episodic memories was carried out by (Linton 1982) who tried to explain the loss of episodic memories via repetition. She showed that over time there was a decrease in the significance of a repeated event, thus making it less distinct and memorable. Over time repetitive episodic memories decreased while the semantic memory base of those events increased, which supports Cohen's theory illustrated in Figure 2.3.

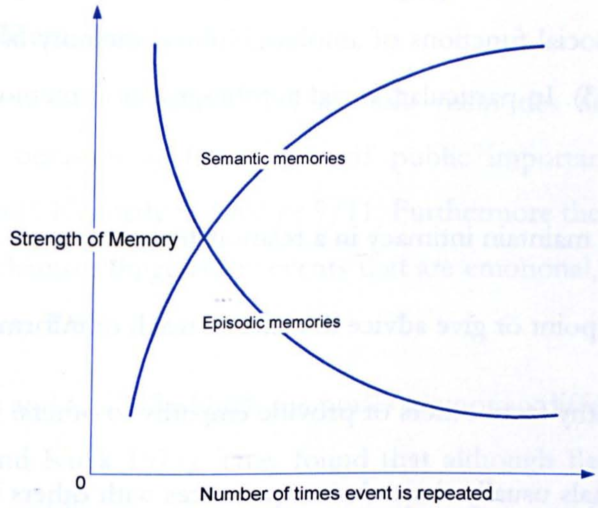


Figure 2.3: The effect of the number of times an event is repeated for Semantic and Episodic memory, adapted from (Cohen, Kiss et al. 1993), p.53.

2.4.5.4 Autobiographical Memory: memory for personal experiences

Autobiographical memory is a subset of episodic memory, but for personal biographical events. It consists of specific life events that have self-reference or personal significance. According to (Cohen, Kiss et al. 1993) some autobiographical memories can be categorised as *declarative* which means they simply record a particular fact such as being a student or owning a house. However, many autobiographical memories are *experiential* or *procedural*. For instance when recalling an autobiographical memory of going on a holiday, one may relive the actual experience rather than just the fact of going on holiday. Such memories might be accompanied further with specific images and emotions.

Three autobiographical memory functions have been proposed (Cohen 1998):

1) Self;

2) Social; and

3) Directive function.

These functions have been further studied by (Bluck, Herbermas et al. 2002) who found evidence suggesting that autobiographical memory is used in everyday life. Research that further investigates social functions of autobiographical memory has been carried out by (Alea and Bluck 2003). In particular, social autobiographical memory functions are used to:

- develop or maintain intimacy in a relationship;
- illustrate a point or give advice in order to teach or inform others; and
- elicit empathy from others or provide empathy to others.

In addition, individuals usually share their experiences with others in order to fulfil their social goals. Thus, social functions are crucial for autobiographical memory due to its ubiquity in everyday life. Social psychology research suggests that people are generally more critical of themselves in the past, than their current selves (Wilson and Ross 2003). One explanation of such findings is that people are not actually traitors of their past selves, but merely impartial observers of former selves. Perhaps people see a past self as they really were and view their present self too favourably.

Some autobiographical events can be deliberately forgotten, especially unpleasant and disturbing events (McNally, Clancy et al. 2001). Studies by (Joslyn and Oakes 2005) further support the finding that intentional forgetting is a plausible explanation for the absence of certain autobiographical memories. Other research into tunnel memories (enhanced memory for the central details of an event) for autobiographical events suggest that tunnel memories are limited to emotionally negative memories (Berntsen 2002).

According to (Conway and Pleydell Pearce 2000), there are two key differences between autobiographical and episodic memories:

1) Trivial events are more likely to be found in episodic than in autobiographical memory.

2) There is always a feeling of re-experiencing with episodic memory but not necessarily with autobiographical memory.

2.4.5.5 Flashbulb Memories

According to (Brown and Kulik 1977) Flashbulb memories are vivid and detailed recollections of an occasion such as news of public importance, for instance the assassination of John F. Kennedy in 1963 or 9/11. Furthermore there is a suggestion that there is a neural mechanism triggered by events that are emotional, surprising and highly consequential (Conway 1995).

According to Brown and Kulik flashbulb memories are not so different from other vivid memories (Brown and Kulik 1977). They found that although flashbulb memories are defined as relating to public importance, some private memories of personally important events are very similar.

The existence and nature of flashbulb memories have been the subject of intense debate. According to one view (Brown and Kulik 1977; Conway, Anderson et al. 1994) flashbulb memories are distinctive types of memory created by a special encoding mechanism. However, other authors stress the importance of reconstructive post-coding factors (Wright and Gaskell 1992; Wright 1993) claiming that as with any other ordinary memory, flashbulb memories get distorted and biased over time.

Furthermore, (Brown, Rips et al. 1985) found that people seldom have a precise memory of the dates of public events, so that these dates are estimated rather than remembered (Cohen 1996). This is based on the amount of information that is available about that public event. Other research suggests that people date personal events by relating them to temporal landmarks (Shum 1998). Thus if one remembers a public event, one is more likely to remember another personal event that happened at that time. This suggests that there might be a direct link between autobiographical and flashbulb memories.

2.4.6 Prospective and Retrospective Memory

Explicit and implicit memories are generally used for two purposes: remembering the past or the future respectively:

- **Retrospective** memory – the memory of past events; and
- **Prospective** memory - 'remembering to remember' (Winograd 1988), that is, the memory of future events, plans.

Retrospective and prospective memories are linked to episodic as well as semantic memories. Prospective memory (Sellen, Louie et al. 1997), is highly reliant on the deliberate use of cues. These cues can be of the following nature:

- Event based - where a memory of a specific past event acts as a cue for some other future event; or
- Time based - where events are remembered by a time cue, for instance, to go to the doctors at 4pm today.

2.4.6.1 Difference between Prospective and Retrospective Memory

Prospective memory is used to remember to do things in the future. Retrospective memory involves remembering events experienced in the past. However prospective and retrospective memories are not completely distinct. An event or a plan that is stored in prospective memory includes some knowledge acquired from retrospective memory. A plan to buy fruit on the way home includes retrospective knowledge about which shops sell fruit and which fruit I like or dislike. However retrospective and prospective memories do not always correlate as shown in studies by (Baddeley 1998). He showed that the phenomenon known as the 'absentminded professor effect' reveals that both of these types of memory are functionally distinct. The distinguishing feature of prospective memory is that it involves a future oriented time element.

Each memory requires some form of internal or external reminders. An internal reminder could consist of mentally linking a prospective task to some constant event or a repetitive habit. Alternatively one could choose to rely on an external reminder such as a diary, a personal digital organiser, a simple task list or any other form of prosthesis. These all provide a trail of time based and event based cueing mechanisms.

2.4.7 Metamemory

Metamemory is a central process that involves awareness of when we forget, as well as about the efficiency of different memory strategies and cuing processes. It is crucial for the study of PM: PM involves the intentional deployment of tools to meet perceived memory goals. A better understanding of Metamemory should allow us to understand how and when such tools are used and how they should be designed.

Generally humans can determine what they will be able to remember and what they will not. This knowledge has been described as Metamemory and was explored by Lachman, Lachman and Thronesberry (Lachman, Lachman et al. 1979). They demonstrated that Metamemory is generally accurate (i.e. that people correctly predict what they will remember), and that most people can direct their memory search effectively and productively. These results proof that people really do 'know their own memories'.

According to (Cavanaugh 1988) there are three types of Metamemory:

1) Systemic awareness;

2) Epistemic awareness; and

3) On-line awareness.

These three types of Memory are all interrelated. Systemic awareness consists of knowing what kind of encoding and retrieval produces the best results, as well as what kind of things are difficult or easy to remember. For instance, it is hard to memorise long words that do not have any meaning. Epistemic awareness consists of knowing what we know and being able to make judgements about the accuracy of what we know. For instance, knowing that there is a double "b" and a double "o" in the word "gobbledegook" and being able to apply grammar rules to determine the accurate location of those double letters. On-line awareness consists of knowing about ongoing memory processes and being able to monitor memory functions, such as whether one is correctly remembering someone's name when one sees them. Thus, when trying to recall a particular fact, **epistemic awareness** will be needed to identify what relevant information is potentially accessible, **systemic awareness** may guide the selection of search strategies directing the search, and **on-line awareness** might be involved in keeping track of progress during the search.

Online awareness of Metamemory allows us to make a decision about the accuracy of the memory as it is being retrieved. Research investigating the relationship between confidence and accuracy in Metamemory has been carried out by (Brewer and Sampaio 2006) who showed that Metamemory confidence in recently completed memory performances plays a major role in the use of memory in everyday mental life. The way that Metamemory helps is by guiding memory searchers, e.g. if one has misplaced the keys, Metamemory will help memory by guiding it through the finding process (e.g. when did you last see the keys, can you visualise where they were, is this definitely the last place where you saw them etc.). Furthermore, Brewer and Sampaio claim that whenever individuals have to act on a memory operation that has just been completed, the use of those memories is mediated by their confidence in the accuracy of the memory product. These findings are crucial in our justification for the concept of Memory Prostheses. Appropriate memory prostheses can influence perceived accuracy, increasing user confidence that prosthetically retrieved memories are accurate regardless of the confidence state at the OM acquisition stage. They also motivate the initial use of a prosthesis. If someone is aware that particular information is likely to be forgotten using OM they are more likely to take strategic action to remember e.g. by using Prosthetic Memory. Similarly when trying to retrieve information, if they are aware of the shortcomings of their Organic Memory and they may rely on an external record for retrieval. Therefore the development of new PM tools mean that information, which up to now has been hard to remember with Organic Memory, can be recalled by the use of

Memory Prosthesis. This should increase user confidence and accuracy of information retrieval.

Furthermore, studies have been conducted into a models explaining how Metamemory processes are used to regulate memory accuracy and quantity performance under free-report conditions (Koriat, Goldsmith et al. 2000). According to the model "when recounting past events, people monitor the likelihood that each item of information that comes to mind is correct. They then apply a control threshold to the monitoring output for the item with the highest subjective probability of being correct. The item will be reported if its assessed probability passes the threshold, and will be withheld otherwise". The control threshold is set on the basis of implicit or explicit payoffs, in other words, the relative utility of providing complete versus accurate information. The stronger the motivations for accuracy, the more selective people are in their reporting and a higher level of accuracy is attained. Thus, individuals can achieve a high level of memory accuracy by withholding a large number of correct answers.

There are a number of phenomena associated with Metamemory. The Tip of the Tongue (TOT) phenomenon involves epistemic Metamemory as shown in studies by Brown (Brown 1991). Feeling of Knowing (FOK) (Wellman 1977) is distinguished from the TOT state, as relating to a whole range of knowledge states: from being sure that one does not know something, to being confident that one can recall it (if given enough time and/or suitable cues).

It is obvious that Metamemory is directly relevant to the use of Prosthetic Memory. Metamemory determines whether memories have been forgotten or not. In the case when memories are forgotten, the decision to use Prosthetic Memory might be prompted by Metamemory processes. But this raises further research questions about how people retrieve forgotten memories from their Organic Memories, or decide to use PM when they were not be able to do so.

2.5 Memory Encodings

In Chapters 6 and 7 I compare pictorial versus verbal cues for remembering complex learning materials. Here I review the different encoding processes for pictures and words.

2.5.1 Memory for Pictures: Visual Encoding

One of the most popular theories is dual coding theory (Paivio 1986) which states that pictures are better remembered because they are encoded with two specific codes (pictorial and verbal) while words are only encoded with single code (verbal). During retrieval, pictures have an advantage because of this extra coding. Paivio also derived two different types of representational units: "imagens" for mental images and "logogens" for verbal entities, which he describes as being similar to "chunks" by (Miller 1956). Dual coding theory identifies three types of processing:

1) *representational* (a direct activation of verbal and non-verbal representations);
2) *referential* (activation of verbal systems via non-verbal interaction and vice-versa); and
3) *associative* (activation of representations takes place within the same verbal or nonverbal system).

According to Paivio, a task may require one or a combination of those three types of processing at once, as illustrated in Figure 2.4.

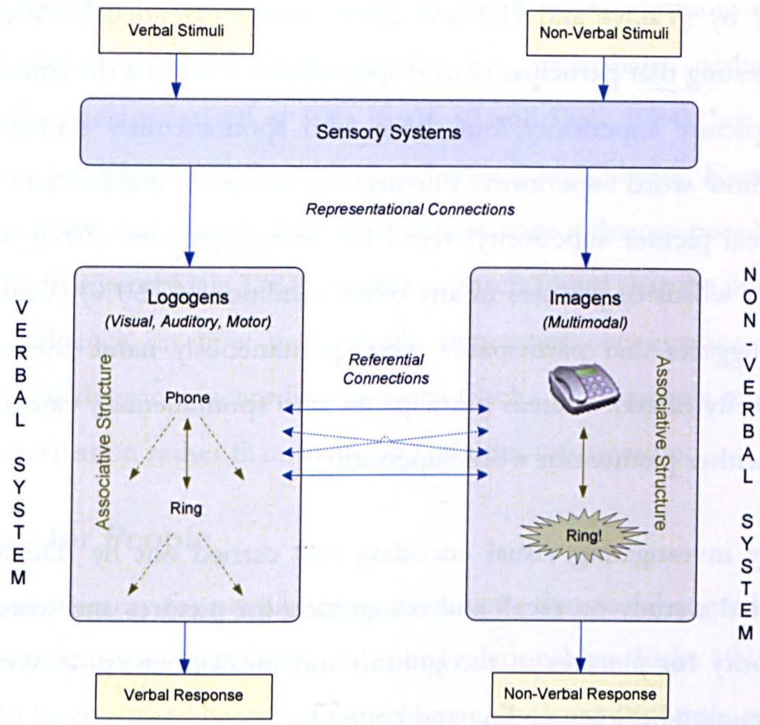


Figure 2.4: Verbal and Non-verbal encoding, adapted from (Paivio 2006) p.143.

Another theory of picture processing is discussed by (Weldon and Roediger 1987) who also observed that pictures and words elicit different types of information at encoding, but the type of processing retrieval is also important. The theory states that items that are encoded with conceptual processing will have a greater chance of being re-accessed at the retrieval stage. This theory also suggests that pictures are encoded with more conceptual processing.

However another study conducted by (Weldon and Coyote 1996) reported results from implicit and explicit tasks which did not fully support the processing theory of picture superiority proposed by (Weldon and Roediger 1987). The reason for the lack of picture superiority in conceptual priming in Weldon and Coyote's study was due to open-ended encoding instructions. Participants were simply asked to pay attention to the pictures and words.

Another study by (Vaidya and Gabrieli 2000) further extended Weldon and Coyote's results in suggesting that participants who spontaneously named the stimuli in their study would show picture superiority, but those who spontaneously semantically classified them would show word superiority. Further extensions of Weldon and Coyote's cued-recall test reveal picture superiority: recall for named pictures (70%) was significantly higher than for words or pictures in any other condition (45-50%) (Vaidya and Gabrieli 2000). This suggests that participants who spontaneously name the pictures show a picture superiority effect, whereas participants who spontaneously categorise the stimuli would show neither picture nor word superiority.

Another study investigating visual encoding was carried out by (Durso and Johnson 1980) conducted a study on recall and recognition for pictures and words. Overall they found superiority for pictures in recognition and recall (Pictures > Words). They also found an interaction between coding and content:

a) for words: verbal cues least effective; and
b) for pictures: imaginal cues least effective (two codes > one code).

However, the visual encoding must be dependent on how relevant the pictures and words are to the participants. Chapters 6 and 7 of this thesis examines the difference between textual and pictorial annotations in a large classroom setting showing that the greater granularity of textual annotations may provide richer cues to the underlying information. Chapter 8 focuses on autobiographical pictures combined with textual and audio annotations highlighting the link between both forms of encoding.

2.5.2 Memory for Spoken Conversation: Echoic Encoding

How well we remember conversations are a crucial part of our functioning as humans. Speaking is an easy form of interaction which demands extensive just-in-time processing and remembering. We assume that when we hear things we are able to remember them without noting things down or transferring this information onto some external memory device for later retrieval. There has been some research into how people remember their everyday conversations. This was focused on recall vs recognition when retrieving conversational information. Average results for immediate verbatim recall of conversations has been estimated at 10% (Stafford and Daly 1984), but for recognition of conversations, participants performed much better (MacWhinney, Keenan et al. 1982). It has also been noted that anxiety and social competence influence people's memory for conversation (Stafford and Daly 1984). Other research found that the importance of the topic had an influence on how well people remember conversations (Conway and Bekerian 1987). Furthermore, according to Sachs (Sachs 1967) people tend to remember the gist of a conversation rather than focus on verbatim information.

2.5.3 Memory for People

Most research on memory for people seems to indicate that social cognition is organised according to the features of individuals and social relationships (Fiske 1995). Memory for people seems to be organised based on the schemas, goals, and motives of the perceiver. However Fiske suggests that social memory tends to be organised in terms of relationships: how people relate to others. For example, if people fail to remember a person based on a specific association, they are very likely to recall another person who has similar relationship ties. This thesis reports results on autobiographical memory for people in Chapter 8.

2.5.4 Memory for Dates: Temporal Encoding

According to Rubin (Rubin 1988) one part of autobiographical memory is concerned with memories for dates. He checked a sample of events that participants recalled and dated against their own diary records and found that 74% were correct to within a month.

Furthermore, it seems that people predominantly organise their autobiographical memories into temporal 'life chapters' with clear end points (Thomsen and Berntsen 2005). For instance, when people were asked to recall a specific time from their college years, such as their first term at college, people remember more episodes from the beginning and the end of a time period rather than the middle.

2.5.5 Memory for Places: Spatial Encoding

According to Cohen, people and places are better remembered than temporal events (Cohen, Kiss et al. 1993). Further research investigating memories for places (Knez 2006) suggests that early childhood places and holiday places are most memorable.

2.5.6 Emotional Memory

We naturally remember events that we have emotional associations with. These are sometimes flashbulb events of a public, but emotionally significant nature, be it negative or positive. Emotional personal events also have a strong impact on our autobiographical memory. We pay more attention to those events which have emotional significance to us (Easterbrook 1959). However a number of studies showed that memories for neutral stimuli decrease, whereas memories for emotionally aroused stimuli remain the same or increase (Baddeley 1982). Other studies (Heuer and Reisberg 1990) have also discovered that memory for emotionally significant information tends to be greater after longer Testing Sections, than after immediate recall. This is consistent with theories (Kleinsmith and Kaplan 1963) that emotionally arousing memories tend to generate a relatively permanent trace whereas less emotionally significant memories tend to be more subjective to decay. Generally emotional memories are thought of and remembered more often than less emotional memories.

Chapter 8 explores emotional associations in autobiographical memory.

2.5.7 Physical Mementos and Material Culture

Finally, research in material culture has looked into the sociology and psychology of mementos, i.e. objects that people acquire in order to deliberately remind themselves about their past. Among numerous physical objects – mementos - people encounter during their lifetime, only a few become reminders of the past events, places or people. There have been anthropological studies investigating the material culture of mementos in the life of different cultures (Kopytoff 1986). Material objects symbolise experiences in a tangible and evocative manner (Turkle 2007). Physical mementos seem to elicit more emotion than digital memento archives (Petrelli, Whittaker et al. 2008). Petrelli *et al.'s* findings also show that people often choose everyday objects as their mementos instead of limiting themselves to pictorial representations of people or events (such as photos). Mementos are ubiquitous, but their individual nature differs greatly. Artwork in public spaces evoke conversation, children's scribbles and everyday mundane objects provide a sense of comfort and stability, while long forgotten memorabilia hidden away in a box in an attic lead to strong immersive experiences (Petrelli, Whittaker et al. 2008).

Physical mementos are based *around* active capture, which is an effective cuing memory technique with physical objects of personal nature (Petrelli, van den Hoven et al. 2009).

2.5.8 Intentional and Unintentional Capture and Retrieval

Schacter and Buckner describe *intentional* or voluntary aspects of retrieval – where people intentionally or voluntary try to recall a recent or past experience, also known as explicit retrieval. They contrast *intentional* memory with *unintentional* or involuntary retrieval of previous experiences where there is no deliberate, effortful attempt to think back to the past; instead one is spontaneously reminded of a past event that is accompanied by a conscious recollection experience, also know as implicit retrieval (Schacter and Buckner 1998), illustrated in Table 2.2.

I would like to extend further this notion of *intentional* and *unintentional* remembering to the field of everyday information capture with Prosthetic Memory (PM). For example, I may deliberately commit a song to memory by rehearsing it. *Intentional* capture, however, this requires a conscious strategic decision to remember information being presented. This strategic decision may then involve the use of PM device. In *Intentional* remembering, Metamemory therefore has a crucial role in judging which pieces of information needed to be deliberately remembered, where they are stored, and how they are accessed.

Tool/Focus	Capture	Retrieval
Intentional – PM	√	√
Unintentional - Lifelogging	X	√

Table 2.2: Differences in capture and retrieval for PM and Lifelogging tools.

2.5.9 Processing Strategies

The importance of attention has long been highlighted by James (James 1890) and numerous other psychologists (Watson 1919) of various persuasions reporting that people remember information to which they pay more attention.

Attention and the level of processing (LOP) at the information encoding stage both have a profound influence on recall at longer Testing Sections. They work closely with short-term or working memory selecting relevant information for storage, while applying perception to the received information. Results show performance is significantly higher for information that is attended to at the encoding stage than information that was ignored, and improves dramatically with increase in the depth of processing (Craik 2002).

2.6 Review of Prosthetic Memory and Lifelogging Tools

2.6.1 The Concept of Prosthetic Memory

Prosthetic Memory (PM) is defined as a digital or analogue tool to support or augment Organic Memory (OM). People often plan how they want to remember events, so they bring cameras, camcorders, Dictaphones, mobile phones or more often simple pen and paper. With technologies such as cameras, people often create self-referential records by posing themselves for the picture. With such experiential capture, people still constantly make conscious decisions to prepare to overcome natural forgetting, e.g. taking a picture of something they want to remember in the future or writing down information they want to recall later, such as a name or phone number. When capturing with PM people make *strategic* and *active* decisions about what information should be captured. This *intentional* capture enables users to select when and what to capture.

2.6.2 The Concept of Lifelogging

Lifelogging in contrast does not require active decision making about what to capture, except for the actual act of wearing the lifelogging device itself. This form of capture is *unintentional* and eliminates the process of any conscious decision making over when and what to capture. Some examples of Lifelogging tools are: SenseCam (Hodges, Williams et al. 2006; Sellen, Fogg et al. 2007), StartleCam (Healey and Picard 1998) and Video Summary System (Aizawa, Tancharoen et al. 2004) which are wearable cameras that take pictures at frequent intervals. Lifelogging tools often capture the surrounding environment as well as any interactions with other people and objects. More often than not, the actual Lifelogger gets excluded from the information logs. Sometimes Lifelogging tools can infer the presence of the Lifelogger by displaying familiar things or environment. Lifeloggers are less likely to pose or be able to create active self reference. However, they can look at the world as an observer.

2.6.3 Prosthetic Memory and Lifelogging Vision

Since Bush's Mermex vision (Bush 1945) the main focus of PM and Lifelogging has been on the development of tools to support OM. There has been less attention on how, when or why people would use these PM and Lifelogging tools.

Although a PM can contain a large and sometimes even exhaustive (e.g. with Lifelogging) amount of information, it can still fail to solve the users' memory problem in the sense that such large archive may be hard to access. Finding the right information requires the user to identify the right type of cue to trigger a particular memory. OM uses the function of Metamemory (Brown 1987), that propagates through a set of internal or external memory cues to prompt the retrieval of a particular memory. Once that memory is recalled there is a continuous Metamemory process assessing the relevance, accuracy and reliability of the information retrieved.

Now let us consider the case when people not only have access to OM, but also to PM tools. Figure 2.5 shows that when person receives a demand for information. This information may be either in that person's OM or their PM. The first thing that Metamemory does is evaluates whether this information can be retrieved from OM (I assume there is a preference for OM, as it is usually more efficient). Efficiency is an important determinant for PM use at retrieval. PM tools that store vast amounts of data might be very inefficient and tedious to access. Thus, if this item of information cannot be retrieved from OM alone, Metamemory decides in which PM it might reside. After interrogating, browsing or otherwise searching PM, and finally finding same trace of the information required, Metamemory then evaluates whether this information is the one in demand. This process can be repeated many times until a point of successful retrieval is reached or the search is abandoned. Metamemory takes the role of the intermediate agent constantly evaluating and communicating between OM and PM.

These behaviours are primarily driven by Metamemory, as shown in Figure 2.5.

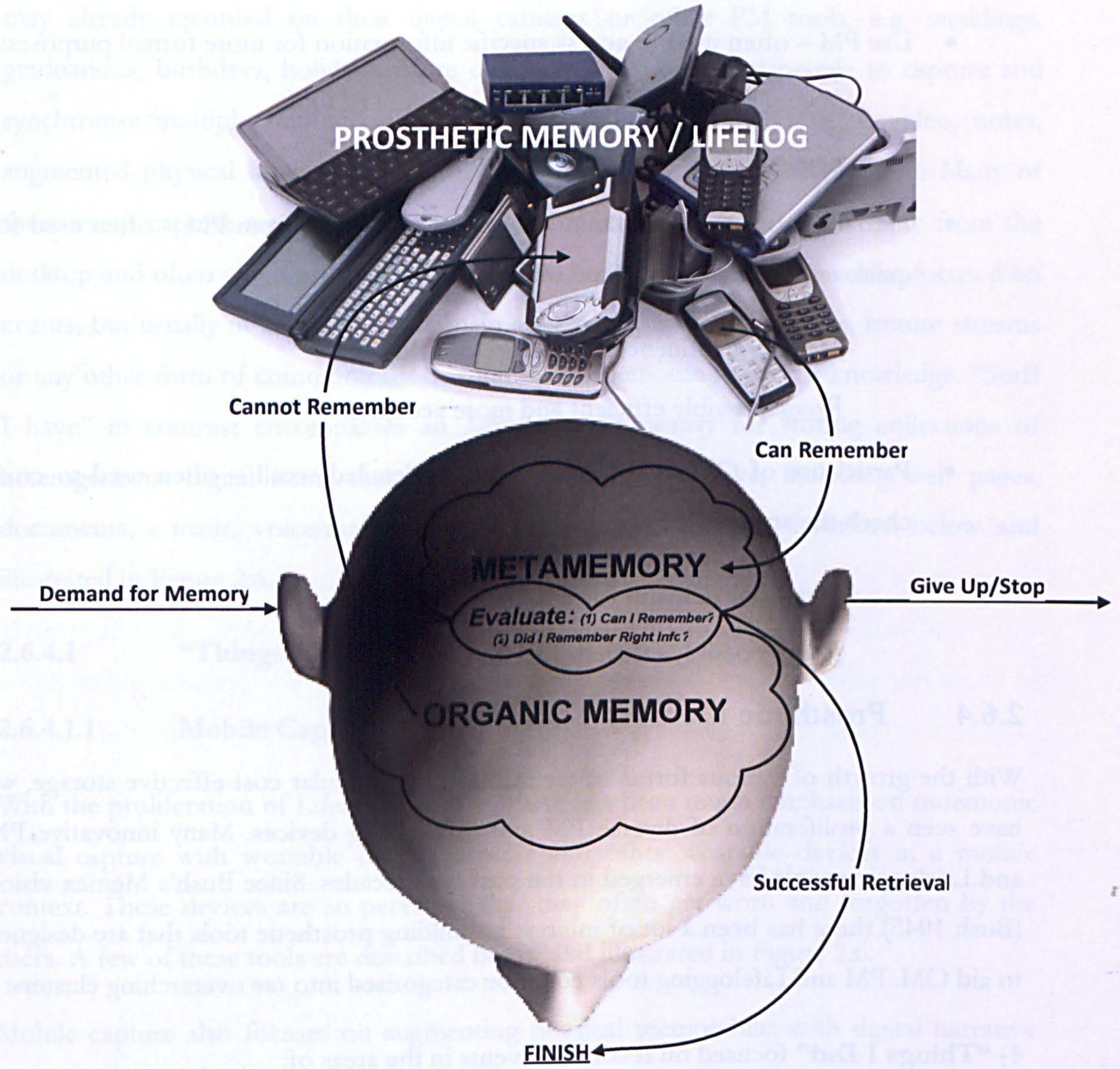


Figure 2.5: Information flow between Organic Memory and Prosthetic Memory with active use of Metamemory.

Based on these continuous interactions between OM and PM, mediated by Metamemory I derived a set of trade-offs that apply when using both forms of memory.

- Use OM alone – often used to find less important gist information:

Motivation: OM is efficient;

Result: not necessarily accurate;

- Use PM – often used to access specific information for more formal purposes:

  Motivation: influenced by potential failure of OM;

  Result: inefficient, but more accurate;

- Partial use of OM with reminders or lookup/cueing from PM – often used for quick reference purposes and reminders:

  Motivation: influenced by failure of OM;

  Result: possibly efficient and more accurate;

- Partial use of OM for gist with PM for detailed recall – often used to cross check the accuracy of specific information:

  Motivation: demand for detailed retrieval;

  Result: possibly efficient and more accurate;

2.6.4 Prosthetic Memory Tool Review

With the growth of various forms of technology, in particular cost-effective storage, we have seen a proliferation of diverse PM and Lifelogging devices. Many innovative PM and Lifelogging tools have emerged in the past few decades. Since Bush's Memex vision (Bush 1945) there has been a lot of interest in building prosthetic tools that are designed to aid OM. PM and Lifelogging tools could be categorised into *two* overarching clusters:

1) **"Things I Did"** focused on recording events in the areas of:

   1) *Mobile Capture*; and

   2) *Domain Specific Capture*;

2) **"Stuff I Have"** oriented towards the organised archive of digital items in the area of:

   3) *Infrastructure Repositories*;

"Things I did" generally encompasses autobiographical events in people's lives that are of important and personal nature. Often these are special events or occasions that people may already recorded on their digital cameras or other PM tools, e.g. weddings, graduations, birthdays, holidays. More complex PM tools allow people to capture and synchronise multiple multimedia objects e.g. combining audio, images, video, notes, augmented physical objects and sensor data to make a rich meeting record. Many of these event capture tools tend to be in the domain of *Mobile Capture* - away from the desktop and often when users are on the move. *Domain Specific Capture* is also focused on events, but usually in a specialised domain area: e.g. speech in meetings, lecture streams or any other form of communication requiring a mnemonic record of knowledge. "Stuff I have" in contrast encompasses an *Infrastructure Repository* for storing collections of heterogeneous digital items that users encounter and archive, including web pages, documents, e-mails, voice-mail and photos. These clusters are described below and illustrated in Figure 2.6.

2.6.4.1 "Things I Did" Tools

2.6.4.1.1 Mobile Capture

With the proliferation of Lifelogging tools, there has been much emphasis on mnemonic visual capture with wearable digital cameras and other wearable devices in a mobile context. These devices are so pervasive that they often get worn and forgotten by the users. A few of these tools are described below and illustrated in Figure 2.6.

Mobile capture also focuses on augmenting physical memorabilia with digital narrative and associations.

<u>Mnemonic Visual Capture</u>
SenseCam (Hodges, Williams et al. 2006) is a wearable Lifelogging device that allows either sensor based or temporal based automated image capture. Images are taken when there are changes in light intensity, posture, or after a specified time interval. From user evaluations it was evident that *passively* captured images with SenseCam were an effective

cue for helping to retrieve past events, when compared to active capture, which was less effective (Sellen, Fogg et al. 2007). SenseCam has also been a significant memory aid for patients suffering from various memory deficiencies (Berry, Kapur et al. 2006).

StartleCam is a predecessor to *SenseCam* containing a skin conductivity sensor trigger, which is sampled by an analogue to digital converter attached to a wearable computer. Images are captured by a permanently connected camera. When the computer detects a response in the user, i.e. a change in skin conductivity, the buffer of images is transmitted wirelessly over the Internet to a remote server. The user may then view the recordings to trigger their memory. Using such startle detection filter, this tool was demonstrated to work with a few wearers (Healey and Picard 1998).

Steve Manns's *EyeTap* is another personal experience capturing system – where the user wears digital eyeglasses that cause the eye to function as if it were both a camera and a display. Eyetap captures videos as they were originally seen by the user. It provides continuous visual Lifelog which can be shared with others. There are clear benefits in using these technologies for surveillance and archiving personal life, but there are no user studies investigating these effects (Mann, Fung et al. 2005).

In additional to mnemonic video capture, *Memory Glasses* (DeVaul, Pentland et al. 2003) aims to provide just-in-time memory support. This tool uses the wearer's context as a relevant reminder of the situation and sends the user low-attention subliminal cues in the form of images or text. User studies have shown that effective support for OM requires situation-appropriate information, which was provided by *Memory Glasses* in this case.

eyeBlog (Dickie, Vertegaal et al. 2004) is another automatic personal video recording and publishing system. A wearable glasses camera captures videos that are uploaded in a chronologically delineated blog. This tool can capture face to face communications as well as record other visuals based on some pre-set pattern, e.g. glyphs. This allows users to control what is automatically captured, but also edit it later once it's been uploaded onto the blog.

Mnemonic Memento-Event Based Capture

Ubiquitous Memories (Kawamura, Fukuhara et al. 2007) is an object-oriented human memory augmentation system that enables users to associate experience data with a physical object. User studies have shown that this tool supports memorisation and recollection of past events.

The Living Memory Box (Stevens, Abowd et al. 2003) multimedia appliance can be seen as both an archival and narration device, allowing families to bring artefacts and tell stories about these mementos. This tool captures stills, video and audio about physical objects placed in the box and gathers metadata about them, e.g. date, time and place. User feedback about the initial prototype was positive; however users felt they wanted to be able to annotate their own mementos instead of having automatic annotation. They wanted to be encouraged by the system to tell stories about their mementos and they also wanted to be able to include multiple voices in their speech-based narratives.

Frohlich and Murphy's *Memory Box* is a wooden box that contains physical mementos. When these mementos are removed from the box, an individual story or an ambient sound associated to that memento is played back automatically. Overall users responded positively to the idea of attaching stories to the souvenirs, as long as this process was simple and was part of how these objects were handled (Frohlich and Murphy 2000).

2.6.4.2 Domain Specific Capture

This type of intentional capture technology focuses on recording domain specific information, e.g. meetings, lectures, and other domains where there is a great demand on OM emphasising on verbatim recall.

Meeting Capture

Great memory demands are present in specific domains. Thus people often use PM tools for these situations. However despite extensive development of "meeting capture" tools people still tend to rely on simple pen and paper to capture important information in meetings. A few of these tools are described below and illustrated in Figure 2.6.

Whittaker *et al.*'s *Filochat* system combines an audio recorder with a tablet PC for taking notes as a means of constructing a meeting record. Notes are temporally co- indexed with audio, so that clicking on a note accesses the speech that occurred when the note was taken. The tablet PC provides a notebook base that allows users to store several pages of notes and organise them into sections. Users can then use the notes they have taken to jump to the relevant portion of the conversation. The system proved successful both in field and lab experiments (Whittaker, Hyland et al. 1994).

Another audio-based PM tool - *NotePals* is a lightweight note sharing system that gives users easy access to each other's experiences through their personal notes developed by Davis *el al* (Davis, Landay et al. 1999). The system allows notes taken by users in any context to be uploaded to a shared repository. Users view these notes with browsers that allow them to retrieve all the notes taken in a given context or to access notes from other related meetings or documents. This is possible because *NotePals* records the context in which each note is created (e.g., its author, subject, and creation time). The system is "lightweight" because it fits easily into group members' regular notetaking practices, and uses informal, ink-based user interfaces that run on portable hardware. This PM proved to be successful for sharing notes between users.

Stifelman's *et al.'s Audio Notebook* is another audio based PM combining the familiarity of taking notes on physical paper and pen with the advantages of an audio recording. This device augments an ordinary paper notebook, synchronising the user's handwritten notes with a digital audio recording. As with Filochat (Whittaker *et al.*), the user's natural activities of writing and page turning, implicitly indexes into audio. Users can explore their notes using spatial layout or use time-based navigation to access the audio recordings. Small field studies showed that the interaction techniques enabled a range of usage styles, from detailed review to high speed skimming of the audio (Stifelman, Arons et al. 2001). However there was no objective evaluation how well this system helped with information retrieval.

Hayes *et al.'s Personal Audio Loop* (PAL) is another audio based continuous, near-term audio buffering system running on a mobile phone (Hayes, Patel et al. 2004). Audio is constantly buffered but then deleted unless user decides they want to keep it. However from the lab study and diary studies users said that they would like to use this PM tool more and they liked it being on a mobile phone and portable. However, study results did not provide any objective benefits of using this PM to aid OM.

Another tool - *Dynomite* (Wilcox, Schilit et al. 1997) is a pen-based system that allowed users to take notes associated with meetings. A critical feature of the system was that users could classify their notes into different types (e.g. 'todo', phone number, name, date, URL, etc.), which allowed users to create different views onto their notes (e.g. all my notes about 'todos' for the last month). Note-taking behaviour was also used to control audio recordings so that although all audio was recorded, only portions that were accompanied by note-taking activity were saved – on the grounds that these were more likely to be of importance. There were no controlled lab or field user evaluations carried out with this system.

Vemuri *et al.'s iRemember* is an audio based wearable system (Vemuri, Schmandt et al. 2006). This system facilitates capture of life experiences in a variety of everyday situations. This PDA based tool captures and transcribes audio in real time. The transcribed information is presented to the user in a searchable interface. However user evaluation was limited to one author recording everyday conversations with colleagues for 2 years. This limited user evaluation results suggests a preference for visual memory to trigger search.

There are also commercial products built to capture meetings in a professional capacity. *Quindi* is one of such tools (Rosenschein 2004). This PC based application captures, audio, video, presentation slides, and typed notes – fully indexed and ready for search and navigation. Since this is a commercial product, there are no published user evaluations.

Other research meeting systems include AMI's speech browsers (Tucker and Whittaker 2006), numerous speech recognition engines developed at Carnegie Mellon University (2009), speech access and processing techniques developed at the University of Toronto (Munteanu, Penn et al. 2006; Munteanu, Baecker et al. 2008).

Information Capture for Learning

Another area that places a great demand on memory is the area of learning. Theories of learning are too specialised to be included in this thesis; instead our focus is on PM tools that can support OM in a learning environment.

Brotherton *et al's eClass* (Brotherton and Abowd 2004) supports multimedia capture and supplements the regular learning experience with digital video recordings of lectures, slides, digital whiteboard activity, and personal digital notes. User evaluation surveys showed students felt that access to the *eClass* recordings allowed them to participate more effectively in classes. More quantitative learning benefits were not so clear however. *eClass* users performed no better on assignments than those using regular teaching materials.

Similar learning PM tools include Ratto *et al.'s Active Class* (Ratto, Shapiro et al. 2003), Berque *et al.'s Debbie* (Berque, Johnson et al. 2001), Anderson *et al.'s Classroom Presenter* (Anderson, Anderson et al. 2004; Wilkerson, Griswold et al. 2005), and Berque *et al.'s DyKnow* (Berque, Bonebright et al. 2004). A more traditional approach supporting handwritten annotations, Harvel *et al.'s NoteNexus* (Harvel, Scheibe et al. 2005), showed increased access for materials directly related to assessments.

2.6.4.3 "Stuff I Have" Tools

2.6.4.3.1 Infrastructure Repositories

Tools that store Infrastructure Repository data focus more on providing a knowledge base of significant digital or physical items of information. These tools act more as a reference to the underlying life events rather than record of those events. People can just tap into them as and when required for constantly re-occurring memory tasks. There are numerous PM tools that support this type of remembering and archiving. A few of the key tools are outlined below, and illustrated in Figure 2.6.

MyLifeBits (Gemmell, Bell et al. 2006) is a 'personal database for everything', which is intended to store all information. It focuses on archiving any form of communication or digital memento, e.g. office documents, email, digital photos or scanned images. The tool also allows people to store context information, e.g. GPS import and map displays. It also stores yet other information that might be relevant to users, e.g. radio capture, TV capture, telephone capture, as well as "flashbulb" news extracts. All stored data is presented to the user in a temporally ordered archive.

Karger and Quan's *Haystack* (Karger and Quan 2004) is a general-purpose information management environment designed to integrate multiple data sources. It treats any data collection the same, regardless whether it is a bookmark hierarchy, a photo album or an e-mail. User studies in *Haystack* suggested that users benefited significantly from this added cross-domain flexibility.

Dumais *et al.'s Stuff I've Seen* (Dumais, Cutrell et al. 2003) (SIS) is a tool facilitating information re-use. The system provides a unified index of all information that a person has seen, whether it was an e-mail, web page, document or other. The system also creates rich contextual cues for encountered information that can be used when retrieving this information. SIS has been thoroughly evaluated with the users, where recent data and people related data were found to be the most accessed.

Presto (Dourish, Edwards et al. 1999) is another document management system, which instead of organising documents in a traditional hierarchy, provides interaction through user-level document attributes, such as "word file", "published paper", "shared with Jim" or "currently in progress". These are similar to the tags applied in many Web 2.0 systems such as Flickr, del.ici.ous and CiteUlike. These attributes capture different roles that a document might have, which allows users to rapidly re-organise their documents based on the task at hand.

Phlat (Cutrell, Robbins et al. 2006) is a desktop document annotation and location tool. It attempts to integrate both, search and browsing functionality through a variety of associative and contextual cues. Phlat supports a unified tagging scheme for organising personal content across various storage systems and types of files, e.g. documents, e-mails and appointments. Phlat has been thoroughly evaluated, with results showing that user generated metadata becomes increasingly critical for how users organise their content. Tagging was preferred over filing in the longer term.

Remembrance Agent (Rhodes 2000) is a contextually aware associative memory tool for just-in-time information capture from the web. It automatically caches web pages as they are loaded, adding hyperlinks to personal files. It pre-indexes e-mail archives, documents, and other files based on keyword co-occurrence. User feedback after using *Margin Notes* was generally positive.

Freeman and Gelernter's *Lifestreams* (Freeman and Gelernter 1996) is a tool that dynamically organises users' personal workspace. The tool constructs a time-ordered stream of documents. This stream is used to organise, monitor and summarise information for the user. A small user evaluation showed that the system was popular amongst users.

Forget-me-not (Lamming and Flynn 1994) is a system that collects information about selected aspects of the user's activities and organises this data into personal biographies. It can dynamically identify location and interactions with other users, which get automatically logged in the system. It can later prompt users about their activities. However, due to a complex infrastructure set up, user testing was limited to one specific location.

Finally, Rekimoto's *TimeScape* (Rekimoto 1999) is a tool for spatially arranging information on the desktop. It supports "time-travel" to the past or the future of the desktop. The ability to spacially arrange information and chronologically navigate through it allows users to archive electronic information without having to worry about document folders or other file classifications.

2.7 PM tool and Lifelogging Overview

This PM tool review has helped to identify scientific gaps and understand research questions that motivate this thesis, in particular:

1) **Little attention to users -** most of the technology has been developed without real attention to users and their memory tasks. There has also been relatively little serious PM tool evaluation.

2) **Much technology but few principles -** many PM tools that have been developed, however most of these do not apply any principles from psychology of OM, in particular failing to specify what *types* of memory they are intended to support. Furthermore, with the focus being mainly on technology, understanding *how* these PM tools interact with OM has not been addressed, thus creating a absence of design principles that could strategically help in the development of effective PM tools;

3) **Too few serious Long-term usage studies** – from the PM review it became evident that where tools have been evaluated, this has tended to be rather short-term, often with only a few participants, and sometimes not evaluated at all. Therefore there is a need for more long-term evaluations. The advantages of long-term evaluations can not only help highlight the benefits of new-generation PM tools, but also develop new usage principles in the real world; and

4) **Few tools for the capture of Autobiographical memories** – with the proliferation of technologies such as multimedia mobile phones, digital cameras, recorders, and camcorders people are actively capturing important moments in their autobiographical lives. However, from the PM and Lifelogging tool review it became evident that there are only a few tools e.g. Memory box (Frohlich and Murphy 2000) that support upload and organisation of autobiographical memory associations. Therefore, there is a need for PM and Lifelogging tools that support the seamless capture, organisation and access of personal memorabilia.

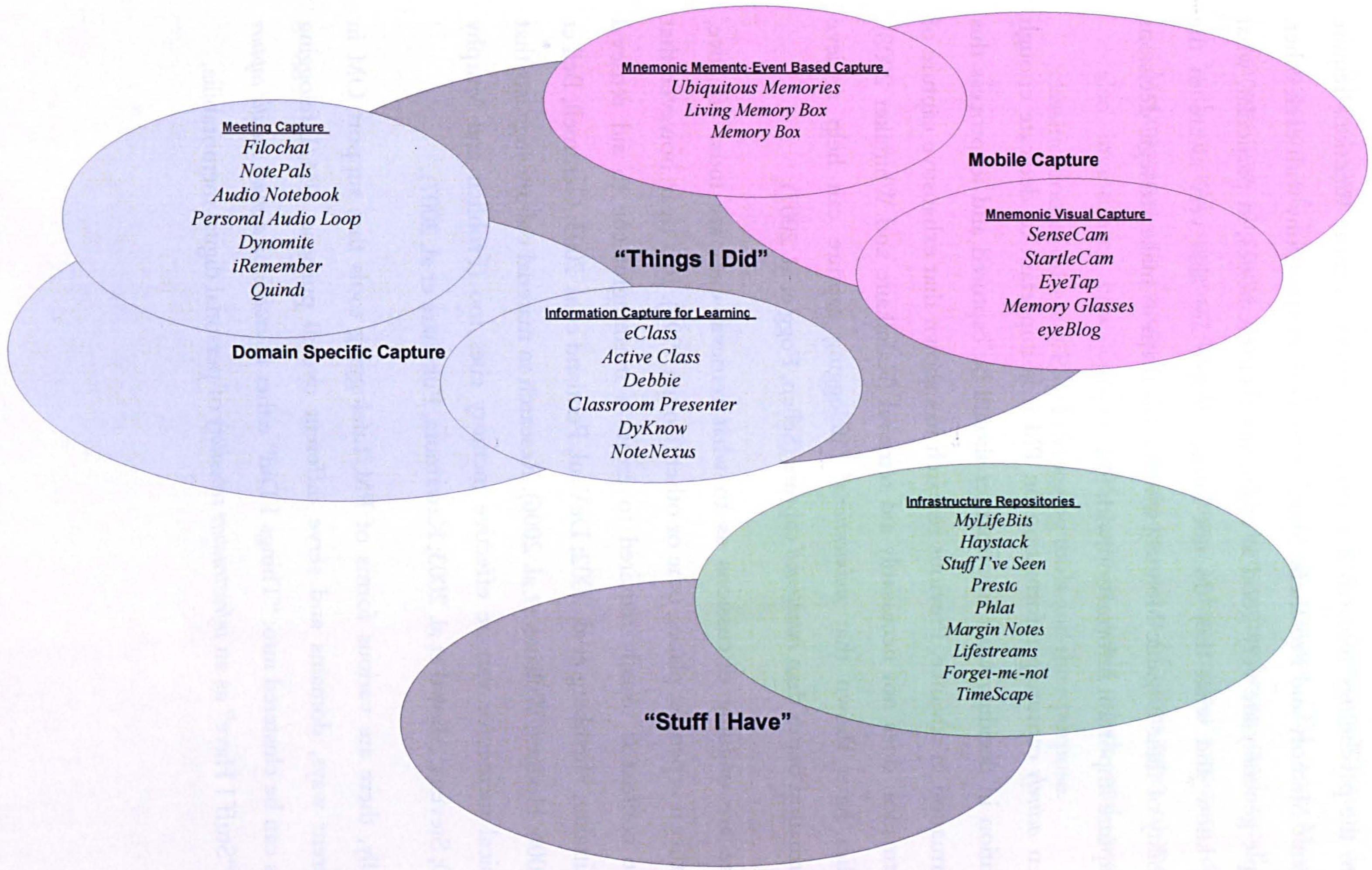


Figure 2.6: Prosthetic Memory Tool Clusters.

2.8 Conclusions

From the psychology research it is evident that Metamemory functions require Organic Memory and Prosthetic Memory to work in synchrony with each other. People generally are very good at evaluating their memories, in particular, what they know and what they do not know. People are also very aware of the fallibility of their Organic Memory which leads them to make strategic decisions to capture important information with PM.

When *actively* capturing information on PM it is important to dedicate enough attention in deciding what information should be captured and to process this information in advance. Previous research has shown that exhaustive capture of information does not necessarily aid retrieval (Kalnikaite and Whittaker 2008). Studies have shown that *unintentional* Lifelogging capture can help retrieve information better than *intentional* capture (Sellen, Fogg et al. 2007).

There are ongoing discussions as to what retrieval cues are most effective, whether it is people, places, time or other (Cohen 1996). It is clear, however, that some context is usually required to create a meaningful cue to aid retrieval (Whittaker, Hirshberg et al. 2002; DeVaul, Pentland et al. 2003; Gemmell, Bell et al. 2006; Hodges, Williams et al. 2006). Research in material culture suggests that physical mementos can be effective memory cues too (Frohlich and Murphy 2000; Stevens, Abowd et al. 2003; Kawamura, Fukuhara et al. 2007).

Finally, there are various forms of PM/Lifelogging tools that support OM in different ways, domains and serve different overall purpose. PM/Lifelogging tools can be clustered into: "Things I Did" either in *mobile* or *domain specific capture* and "Stuff I Have" as an *infrastructure repository* of personal digital memorabilia.

2.9 Chapter Summary

This chapter summarises the relevant psychology literature affecting design and function of PM and Lifelogging tools. It reviews OM and Metamemory interactive functions that lead to effective remembering and recall with or without PM support. It also introduces a hierarchical structure describing Metamemory as a function overlooking processes of other types of memory. It also investigates how current OM support tools work and proposes a classification of different PM/Lifelogging tools and their purpose.

Chapter 3

Methodology

3.1 Introduction

Human Computer Interaction (HCI) focuses on understanding how systems can be built to suit human needs for specific requirements through integration of disciplines such as psychology and sociology of human behaviour and how it applies to computer usage (Sharp, Rogers et al. 2007).

HCI covers a broad range of areas, including pervasive, ubiquitous and tangible interaction technologies (Greenfield 2006), which take the HCI post-desktop by integrating information processing into everyday activities, grounded in psychological theory. There is much debate about HCI methods in general, although there has been a consensus in recent times about using the Socio-technical approach (Sharp, Rogers et al. 2007) for building and evaluating new tools and systems.

This chapter looks at methods that can be used to inform the design and evaluation of new Prosthetic Memory (PM) tools and how they have been applied in the studies presented in this thesis. This chapter further discusses affective design principles (McCarthy and Wright 2004) focusing on user-centric interaction design methods. In particular it summarises the key evaluation and data analysis techniques and how these elements have been applied in the design and evaluation of PM tools in this thesis.

3.2 User-Centric Interaction Design

Design involves trying to achieve some set of user-centred goals given a set of constraints (Dix, Finley et al. 2003; Sharp, Rogers et al. 2007). The goals often involve asking who the system is meant for, and why it is needed. Constraints could be the infrastructure, materials and platforms that the system has to work with. The key goals to *interaction design* are not only to understand the limitations, capacity of tools and platforms required, but also *people* in terms of their psychological and social aspects - their need to interact with these tools while simultaneously allowing for human error.

The PM tool design of this thesis has been accomplished through the use of the user-centric design model, illustrated in Fig 3.1. The basic stages consisted of the following (a) understanding "what is wanted" by the user (user requirements); (b) designing and if necessary re-designing the tool; (c) building an actual prototype of the tool; and (d) evaluating and analysing the final result.

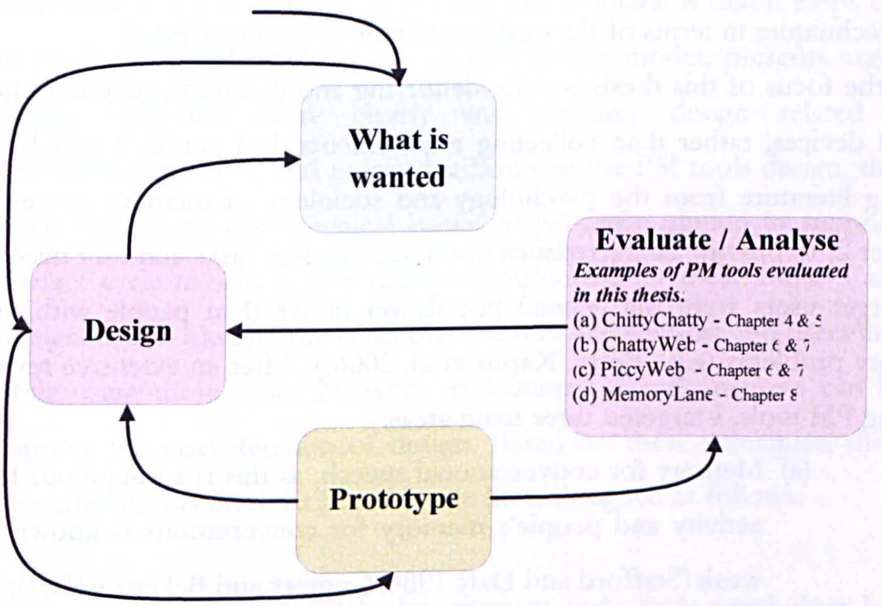


Figure 3.1: Basic User-Centric Interaction Design model, adapted from (Sharp, Rogers et al. 2007).

The following sections outline these stages in more detail, summarising current methods and validating my approach in choosing one method over another for the studies presented in this thesis.

3.2.1 What is wanted?

Gathering user requirements involves finding out what attributes and functionality a PM tool should possess and what infrastructure should be used to support this. These requirements can be gathered using ethnographically informed methods (Hughes, O'Brien et al. 1997). These provide rich data generated through naturalistic user observation while users interact with real PM tools. However, this approach is time consuming and results are often based on the interpretations of the observers. The advantages of this approach are that it provides rich natural data about user behaviours.

Other requirement techniques involve focus groups, interviews or contextual enquiry (Beyer and Holtzblatt 1998). However, there are trade-offs between these techniques in terms of the quality and type of data generated.

Since the focus of this thesis was in identifying and designing generic principles of PM devices, rather than collecting requirements data per se, I relied on the existing literature from the psychology and sociology of memory (reviewed in Chapter 2) to provide characteristics of critical memory tasks and user needs.

My target users were the general population rather than people with known memory problems (e.g. (Berry, Kapur et al. 2006)). After an extensive review of existing PM tools, I targeted three main areas:

(a) Memory for conversational speech, as this is a ubiquitous human activity and people's memory for conversations is known to be weak (Stafford and Daly 1984; Conway and Bekerian 1987);

(b) Memory for complex learning materials. Previous research (Brown 1987; Bransford, Brown et al. 1999; Robertson, Lane et al. 2000; Brotherton and Abowd 2004; Hiltz and Goldman 2005;

Walker and Moore 2005; Munteanu, Baecker et al. 2008) has shown that educational settings place great memory demand on students, requiring them to understand and remember large amounts of complex novel materials;

(c) Autobiographical memory, where there are few tools for this important aspect of memory.

3.2.2 Design and Prototyping

I now turn to the design phase. Generally, there are two types of design: conceptual and physical. Conceptual design is concerned with developing a conceptual model that describes what a tool will do and how it will behave. Physical design is concerned with the details of the design, for instance, screen and menu structures and other graphics (Sharp, Rogers et al. 2007).

Each design has a rationale (Haynes, Carroll et al. 2009) which 'explains' why a tool is the way it is. The key benefit of this rationale is that it helps the designer to communicate throughout the chosen design model, presents arguments for design trade-offs more clearly and captures design related contextual information. In HCI, and more specifically in the PM tools design, this rationale often focuses on psychological issues, in particular aiming to support the task-artefact cycle in which user tasks are supported by these tools (Carroll 2000). Designers also identify the tasks that the tool will support and users are observed while using these tools. Negative evaluations of early designs can be used to improve the next iteration of design. Based on these principles, the PM tools presented in this thesis in general have been designed as follows:

1. Initially, I identified the key memory tasks from psychology literature on memory;

2. I built initial prototype designs based on current practices, e.g. 1) how people take-notes or 2) how people revise coursework and currently use social feedback; or 3) how they create mementos.
3. Where possible, I exploited the use of familiar metaphors e.g. pen and paper note-books, lecture slides, mementos in a house.

My designs incorporated the general principles of constant user feedback and affordances (Norman 1986), to include elements such as: 1) letting users know where they are; 2) letting them know what they can do with the tool; 3) letting them know where they are going and what will happen; and 4) informing users of where they have been or what they have done.
The final design/prototypes emerged through iterative design-evaluation-redesign process involving users.

3.2.2.1 Conceptual Design

According to (Sharp, Rogers et al. 2007), a conceptual model is "a description of the proposed system in terms of a set of integrated ideas and concepts about what it would do, behave, and look like, that will be understandable by users in the manner intended." Often the basis of designing a conceptual model requires the creation of a set of user tasks to be supported by this tool. There are three ways to think about in a conceptual model (Sharp, Rogers et al. 2007):

1. *Which interaction mode would best support the users' activities?*
   For the PM tool conceptual design model, I used a combination of modes. For example, in Chapters 4 and 5 I designed a handwritten voice annotation PM tool, where one of the user tasks was to take notes while recording audio. In this case, to capture audio is more like a conversation where users state the name of the meeting and the recording starts.

Continuous timer feedback indicates to the users that the audio is being recorded as they take notes.

2. *Is there a suitable interface metaphor to help users understand the tool?*

   Users like familiar metaphors, which I explored in the conceptual model of the PM designs reported in Chapters 4, 5, 6 and 7 by using familiar hand written annotations and the learning tools use both notes and slides as familiar metaphors. Similarly the conceptual model for the PM tool design in Chapter 8 explores familiar metaphors of a concept of a house.

3. *Which interaction paradigm will the tool follow?*

   Thinking about the users' tasks for the PM tools presented in this thesis, each task has to be individually generated in a way that is convenient to the user setting for which the *pervasive* interaction paradigm would be most appropriate, because these are everyday memory demands. This way, PM tool users will be able to capture and recall information on their PDAs and personal computers (see Chapter 4 and 8).

3.2.2.2 Physical Design

Physical design involves a more refined outlook on the way that the tool will look and behave. It focuses on the interface design and its elements such as screen, icon or menu use.

I also had to take into account hardware and infrastructure requirements and what users would be expecting of these tools. I decided on the necessity of the following infrastructure and hardware characteristics for PM tools:

1. **Mobility.** Since a lot of memory requests tend to take place when people are actually away from their desks, it is essential to have a portable PM tool that can be accessed at any location.
2. **Familiar Metaphor.** New technologies can be overwhelmingly complex. It is essential that PM tools are easy to use and explore familiar

metaphors such as taking handwritten notes (Chapters 4-7) or exploring a model house (Chapter 8).

3.2.2.3 Prototyping

Prototyping is essential for answering design questions and choosing between alternatives. Thus prototypes can serve a variety of purposes, such as testing out technical ideas, clarifying user requirements, or to do user evaluations.

In prototyping, there are often *three* or more iterations. I have applied all these prototyping stages in the prototyping process of the PM tools presented in this thesis as outlined below (Dix, Finley et al. 2003):

1) **Low-fidelity prototype** – paper based sketches of a PM tool interface and functionality. These allow designers to obtain rapid feedback about initial design concepts.
2) **High-fidelity prototype** – a partially or fully functional mock-up, similar to the final PM tool running on a dedicated platform with the support of an appropriate infrastructure.
3) **Final prototype** – after a pilot evaluation, a final version of the PM tool is designed with full functionality.

The advantages of starting with a Low-fidelity prototypes in this way, is that it is easier to highlight possible errors, and it is possible to identify possible errors in information flow or the overall functionality of the PM tool. As this does not require us to build a working system, it is by far the easiest and most cost effective way to prototype a new tool.

Once the Low-fidelity prototyping has been completed with initial user feedback incorporated into the design, I proceed to the next stage of developing a high-fidelity prototype. This stage involves the implementation of the tool on the chosen platform with the chosen infrastructure. Improvements can be added or new design changes can be introduced that incorporate user feedback. Once the

High-fidelity prototypes were finished, I piloted them in real user settings. The feedback received at this stage allowed me to finalise the PM tool, incorporating necessary changes for the *final* prototype of the PM tool. At this stage, the PM tool is ready for the final user testing.

3.2.3 Evaluation Techniques

In order to understand various aspects of a new PM tool and its impact on users, newly invented tools need to be evaluated. In HCI, there are a number of evaluation techniques that primarily focus on testing usability and functionality of new tools. Evaluation should showcase usefulness highlighting the impact it might have on the users. The evaluation of the new tool should aim to assess: design, implementation and how well it fits the conceptual model. Evaluation, therefore, should be considered at all stages in the design life cycle. More specifically, the goals of evaluation are to assesses the extent of the tool's functionality, the effect of the interface on the users, identifying problems, and in particular for PM tools, there needs to be an evaluation of how these tools help specific memory tasks. There are different methods and approaches to evaluating new tools. The following outlines key evaluation approaches in the field of HCI.

One of the more popular approaches has been adopted from psychology: **Cognitive Walkthrough** – this approach evaluates the design in terms of how well it supports the user in learning new tasks (Lewis, Polson et al. 1990). This type of evaluation is usually performed by experts in the field of cognitive psychology, who "walk through" the design in order to identify specific problems in the system using psychological principles. Experts often use pre-prepared forms to guide system analysis, which focus on goals and knowledge e.g. does the design lead the user to generate the correct goals.

Another method for evaluating a tool's usability is **Heuristic evaluation.** This is an engineering method that helps to identify system usability problems. It involves a small set of evaluators who examine the system and judge its compliance with recognised usability principles. It is necessary to involve multiple evaluators in any heuristic evaluation, and according to Nielsen, three to five evaluators are cost-effective since there are diminishing returns for using larger numbers (Nielsen 1994). There are different heuristic methods, the most popular being: Neilsen's 10 Heuristics (Nielsen 1994), focusing on *nine* principles from showing system status to providing users with documentation. Another is Shneiderman's 8 Golden Rules (Shneiderman 1997), with the focus ranging from system consistency to reducing short-term memory load for the user. Finally, Norman's 7 Principles (Norman 1986; Norman 2002) also provide guidelines on evaluating with heuristics. However, there are tradeoffs between Norman's principles, for instance, making the system resemble the real world as an affordance may reduce the set of possible functions the system might carry out.

Another form of evaluation is a **Review Based Evaluation,** where results from the literature are used to support or refute part of a system design. However the drawback of this method is that it requires ensuring that the existing evaluation results are applicable to the new and improved design. It is also possible to use cognitive models such as GOMS (Dix, Finley et al. 2003) that help filter design options in order to predict user performance. However, these methods are most useful for highly constrained domains where tasks are predictable.

Since all the PM tools reported in this thesis focus on practical tasks as well as theoretical principles investigating the benefits of PM tool use as an OM aid, none of the heuristic methods or cognitive walkthroughs was appropriate to use to evaluate these tools. Instead, users were given specific representative memory tasks to perform. The answers and PM tool selections were carefully logged, as well as any additional qualitative data such as comments that users expressed during these evaluations. After carrying out the tasks, I elicited structured feedback from users in surveys. This feedback was then used to improve the next generation of PM tools and derive the design principles and new theoretical approaches.

3.2.4 Evaluating Through User Participation

When building PM tools, it is important that these tools are applicable to the needs and specific memory tasks that people have in their natural environments. Such natural settings introduce a wide variety of potentially confounding variables. For some evaluations it is therefore more appropriate to use settings that are more controlled, such as laboratory settings. Both laboratory and naturalistic approaches are described in detail below.

In a **Laboratory Study** there is a controlled and structured environment where an evaluation can be conducted with as minimal bias as possible. The results are based on controlled data that can provide better grounds for evaluating a set of hypotheses. However, one drawback is that lab studies lack context and tasks have to be carefully selected so they are realistic and representative of the target context. Although laboratory studies provide more control they reduce the naturalness of the environment.

In contrast to lab studies, **Field Studies** also known as studies in the *wild*, are conducted in a natural environment, retaining the everyday context of the subject being studied. They also include *longitudinal* studies where repeated observations occur over a long period of time which is useful for observing how people habituate to a new tool and begin to use it for everyday activities. However, field studies have other disadvantages as users are free to use the tool however they want – which makes generalisation across users hard. They may also use the tool for purposes it was not designed for and even fail to use the tool at all.

Considering the strengths and weaknesses of both of these approaches, all the studies in this thesis applied a combination of both: controlled laboratory based evaluations and naturalistic field trials. This helped to strike a balance between controlled and natural long-term usage. The following chapters describe the procedures in the field trials and laboratory studies.

3.2.5 Processes for Evaluating Tools

Prior to any evaluation, it is essential to have a working prototype. Besides having an actual tool, it is essential to derive a detailed plan for the evaluation process, an *experimental design*. This often involves having additional logging tools that help with data collection through real-time user interaction.

A tool-focused experimental design might be designed to look at *specific aspects* of interactive behaviour in a group of users to test *hypotheses*. This is often done through automatic *logging* of user behaviours while they use the tool. The evaluation consists of a set number of controlled experimental conditions which differ only in the value of some controlled independent variable. Changes in behaviour are attributed to different experimental manipulations.

This is a straightforward set up for quantitatively comparing a set of tools against each other. We might also supplement this with more qualitative measures and have users answer survey questions at the same time as their responses and other behaviours are logged.

Each study reported in this thesis is focused around the specific *hypotheses* it tries to test. The *hypotheses* are well defined in order to generate meaningful results. The evaluation process for each study was set up to run on custom built web sites (see Appendixes A-C), which allowed me to specify representative user tasks and automatically log user answers and other individual behaviours. I found this type of logging to be an essential part of evaluation. I also complement my quantitative results with qualitative user perceptions. Experimental designs are complex with various experimental factors to be considered, described in detail below.

3.2.6 Experimental Factors

Before running user experiments, it is crucial to decide on the key factors that will influence the experimental design. First of all, it is essential to decide on **participants.** This can be a group of users who are representative of the general public or a specific group of people. It is important to have a sufficient sample of participants. As small number of participants can be influenced by individual variability.

In the studies presented in this thesis, my aim was to recruit a large number of diverse participants. For instance, the largest study in this thesis recruited 98 participants (Chapters 6) and the smallest study contained 25 participants (Chapter 7). These numbers are larger than the majority of studies presented in HCI conferences such as CHI.

The next factor to be considered is the **Variables,** which are the items to be manipulated in the study (Greene and D'Oliveira 2006). There are two types of variables: (a) Independent Variables (IV) – these are the factors that the experimenter is trying to manipulate e.g. different types of PMs used in an evaluation study in Chapter 4; and (b) Dependent Variables (DV) – these are the factors that are hypothesised to be affected by the IV with a change in IV, the experimenter tries to see whether this effects a given DV, e.g. from the experiment in Chapter 4, the experimenter tries to see whether using PM (as opposed to OM) affects the accuracy of the answer.

Next we need to be able to predict the outcomes – to generate **Hypotheses.** This is often framed in terms of IV and DV, e.g. how accuracy of answers differs with different PM device properties. We hypothesise the possible results or impact that a PM tool may have in certain conditions. Individual experimental hypotheses are outlined in detail in each study chapter of this thesis.

Finally, an **Experimental Design** has a crucial influence on the data and analysis of the study. There are two types of experimental designs: (a) within groups design – where each participant performs the experiment under each given condition and there is a possibility of transfer of learning. The key benefit is that this approach is less likely to suffer from user variation as each user serves as their own control; and (b) between groups design – where each participant performs only one condition and there is no possibility for transfer of learning and no possibility of control with user. This type of design requires a greater number of users. In the context of HCI, within group designs are often preferable as it may take users some time to habituate to using a new tool, so that a few sessions with a tool may show wide variability.

I therefore applied within group designs in the studies reported in this thesis. Although this was more time consuming, requiring all users to do all the tasks in a counter-balanced order, it allowed me to compare, contrast and map my findings better.

3.2.7 Analysis of Data

Regardless of the nature of the data, it is essential to analyse collected data. Before applying any statistics, it is a good practice to simply look at the data to see if there are any patterns. The choice of statistical techniques depends on the type of data collected, what information is required and what are the hypotheses.

In this thesis I report two essential forms of data approaches: *qualitative* and *quantitative*. These two different data capture methods provide different types of data and require different ways of analyses. *Qualitative* measures provide numeric data which can be analysed using statistical techniques. *Quantitative* is non-numeric and therefore requires different analyses, but can provide important detail that cannot be determined from numbers.

Mixing qualitative and quantitative approaches allowed me to triangulate the data better. I used *qualitative* data as a way of supporting the results acquired with *quantitative* data as revealed in the following chapters. *Qualitative* observations can also generate hypotheses for new studies.

3.2.7.1 Types of Test Analysis

There are different ways that data can be analysed. The psychology literature motivates the use of various statistical tests based on the experimental procedure followed and the data collected (Greene and D'Oliveira 2006). There are three types of tests that can be applied: (a) **Parametric:** can only be used for data with a Normal distribution, which is a very powerful and robust method to analyse quantitative results; (b) **Non-parametric:** analyses do not assume a normal distribution. This type of testing can be more reliable, but less powerful; and (c) **Contingency tables:** are used for classifying data based on some discrete attributes.

In this thesis, I applied all types of tests. The following chapters describe the test choices based on distribution and other attributes. However, before proceeding with statistical analysis, I conducted Kolmogorov-Smirnov goodness-of-fit test (Sheskin 2007) to evaluate the normality of data distributions. Cases where the data contained non-normal distribution i.e. in Chapter 7 I applied non-parametric analytical test. All other Chapters of this thesis contained normally distributed data which lead to use of parametric statistical analysis throughout.

3.2.7.2 Observational Methods

As well as collecting quantitative data, there are other observational methods that can provide rich data. One of the more popular observational approaches is **"think aloud"** (Wright and Monk 1991; Ericsson and Simon 1993). This method involves asking users to talk out loud as they are performing tasks. More specifically, they are asked to describe what they are doing. The advantages of this technique are that it can give useful insights into user strategies (Wright and Monk 1991), and often can show how the system is actually being used. However, it can be subjective, selective and the act of describing can influence the task performance.

Another approach is to use "**Cooperative evaluation**". This method is based on 'think aloud', but in addition the participant can ask the evaluator questions throughout. This approach carries the additional advantages of being less constraining, users are encouraged to criticise the system and participants can ask for clarifications (Monk, Wright et al. 1993).

"**Post-task walkthrough**" approach involves transcripts being played back to participants for comments (Dix, Finley et al. 2003). This can be immediate or delayed. It is particularly useful for identifying reasons for actions. It is often used in cases where 'think aloud' is not possible.

However, other research requires approaches which focus on observations of user behaviours *after* the experiment is complete. One such approach is "**protocol analysis**" (Austin and Delaney 1998). This method facilitates data capture which is analysed later. There are various ways of capturing data for analysis, in particular: (a) the most basic capture is *paper* based capture – simply taking pen and paper notes of user behaviour. It is cheap, but often limited to the speed of writing; (b) a more verbatim approach is to capture *audio.* This is good for 'think aloud' conditions. However it needs special equipment and could be obtrusive and later transcription can be difficult, expensive and require additional skill; (c) similar to audio, there is *video* capture. This provides very accurate and realistic capture. However it generates large amounts of data making it difficult to analyse; (d) a final method is to ask users to record their own data, specifically asking participants to use *diaries, journals or notebooks.* This approach can provide useful insights that are particularly appropriate for longitudinal studies. However, the diary approach can be subjective and time consuming for users; and (e) finally, *logging* captures actual usage of the system in real-time. This is done automatically, which is low cost and gives the evaluator objective data.

The studies presented in this thesis focus on observational methods that use a 'think aloud' approach with a comprehensive protocol recording for later analysis. In particular, I captured user answers on paper based notes and audio. Most of the quantitative data for the studies in this thesis was done via automatic logging of user answers and other interface behaviours. This allowed me the flexibility to capture qualitative data that often is as valuable as quantitative data. However, quantitative data captured through automatic logging provided me with objective measures on PM tool effectiveness.

3.2.7 Query Techniques

When considering various experimental designs, it is essential to consider what other techniques are applicable for effective data collection. An **interview** approach involves an evaluator asking participants analytical questions on a one-to-one basis around pre-prepared topics (Sharp, Rogers et al. 2007). The advantage of interviews is that they can be informal, can vary to suit any context, participants can be encouraged to talk and they help explore issues more fully. Interviews can elicit user views and identify unanticipated problems with technology or work practices. They also provide information about the strategies that users employ. However interviews can be subjective and time consuming. They are also complex to analyse. An alternative approach is to use **questionnaires**, which often consist of more structured questions given to participants (Sharp, Rogers et al. 2007). These often take the form of closed class questions. The advantages of this technique are that it is quick and can reach a large group of people. Also, the result analysis can be more rigorous as there is often more available data. However, questionnaires can be less flexible and less probing. Often they need careful decisions about how the answers should be analysed. There are various styles to questions. The most popular are: general, open-ended, scalar, multiple choice, and ranked style questionnaires.

Throughout this thesis I applied both types of query techniques i.e. interviews and surveys. The advantage of using both is that it allowed me to collect short responses quickly with questionnaires and follow them up with interviews to acquire richer and more diverse user responses. Specifically I interviewed participants to find out how and when they used the tools, as well as to provide detailed design feedback. I used surveys to elicit structured feedback about the PM tools and their benefits/disadvantages.

3.2.8 Other Experimental Design Considerations

When designing user evaluations it is also important to consider the intrusiveness of the system and the users work load. Users are often very busy with their everyday tasks and if we design experiments that consume a lot of user time and effort, this may negatively influence evaluations and participation. Careful consideration has to be made to reduce the cognitive overload and the level of intrusiveness of a PM tool that is being evaluated.

3.3 Conclusions

The key focus of this thesis is to design new PM tools that aid OM and through well structured empirical evaluations, suggest novel theoretical design principles to improve future PMs. The following shows how the studies reported in this thesis fit the above design, prototyping and evaluation methods described in this chapter.

1. In a **short-term controlled laboratory study** investigating memory for conversations in Chapters 4 and 5, I compared different PM tools where I generated and tested specific hypotheses. This study focused on the effect of different PM tools on Organic Memory (OM) and how these

tools can aid OM. For this, it was important to have carefully controlled tasks and time intervals where people used different PM tools and OM. I also wanted to deliberately manipulate the properties of the different PM devices. This approach also allowed for continuous and uninterrupted user behaviour logging and direct user observations. Finally, it allowed for interviews which helped to generate rich high quality user feedback.

While users were carrying out the specific memory tasks set for this study, they were instructed to "think aloud". Through careful "protocol analysis", this approach allowed me to collect more qualitative data from the users, about their strategies and perspectives on using the different tools. At the end of this study, I administered a survey to collect subjective perceptions and also interviewed users about their PM usage strategies and preferences. The survey provided more general feedback, and also structured subjective user perceptions which directly mapped with the objective data that I logged during the controlled session.

At the end of this study, I conducted a **long-term naturalistic field trial** where I investigated how people used a novel PM tool over the longer term in their own environment, as reported in Chapter 4. This was an exploratory study, with a structured survey and an interview at the completion of the study. The long-term exploratory field study method was very effective in highlighting new usage strategies that participants developed when they applied PM tools in aiding their personal everyday memory tasks. This provided insight into real user OM and PM interactions, as well as long term adaptation in real contexts – which cannot be observed in laboratory studies.

2. In Chapters 6 and 7, I report on both a **long-term field trial** and a **short-term controlled lab study** in the domain of education. One objective here was to study real-life deployment of PM tools in settings

that place high demands on memory. In the long-term study, I explored how people used multiple versions of PM tools to help retrieve information from lectures. I had a set of hypotheses that allowed me to explore how people used these PM tools to help them recall specific information from lectures. The reason for choosing a long-term field trial was to provide users with continuous access to multiple PM tools over a longer period of time, which would allow for more insightful observations into when and how people explored such PM tools to help them retrieve important information from complex lecture materials, for real-life demands. I logged every user access to the system, what they looked at and how long they spent at each stage, for each PM tool. As well as logging user behaviours, I also gave users a survey to collect subjective feedback and conducted interviews to acquire more specific answers about individual usage of all versions of the PM tools.

In addition to the long-term study, I carried out a **controlled short-term lab study** (also known as a **"quiz"**), described in detail in Chapter 6, where users were asked specific memory questions. I automatically logged how people used these multiple PM tools and asked users structured questions in a survey to assess their PM usage strategies. The advantage of choosing a controlled quiz set-up was to be able to quantitatively measure the impact different PM tools might have on OM.

This data was then "fed back" into new versions of the PM tools to generate social summaries, as described in Chapter 7. The direct user feedback derived from real user-tool interaction during the long-term study as well as the lab evaluation, allowed me to build a new novel PM tool. I then observed how people would explore socially augmented lecture materials from the same class (reported in Chapter 7). Again the interest was in long term behaviours and so a field trial method was used.

3. In Chapter 8, I conducted a short-term **exploratory field trial** over 3 sessions, where I explored specific hypotheses in the area of autobiographical memory. During this study I collected a lot of qualitative data through surveys and structured interviews and through users actively using a new PM tool to do tasks I suggested. I also collected quantitative data on this usage, which provided a detailed objective understanding of how and why people engaged with the tool. I applied "think aloud" and "protocol analysis" to collect and to analyse the data. The main advantage of applying an exploratory field trial approach was that autobiographical memories are personal and it requires individual preparation to collect, organise and manage such data, making lab methods a dispreferred option. It was essential to spend time with each participant to look at their collected data, observe their annotations and other explorations. The feedback received and data collected were rich and diverse providing qualitative results.

The hypotheses of each study of this chapter are described in detail in the Methods section of each chapter respectively.

3.4 Chapter Summary

This chapter overviews the interactive user-centric design principles applied to the design, build and evaluation processes of new tools in general. Tradeoffs between different experimental designs, data gathering and analysis are discussed in detail. Evaluation methods and primary techniques are also overviewed.

Chapter 4

ChittyChatty: When Do People Use Prosthetic Memory to Augment Organic Memory?

4.1 Introduction

Despite the proliferation of lifelogging tools there remains a lack of understanding of the relationship between such Prosthetic Memory (PM) tools on their impact onto unaided Organic Memory (OM) and how their usage depends on OM. We also know little about how the properties of different PM tools affect usage.

This chapter aims to explore and address some of these mapping issues. The focus is on how people remember conversational speech in an informal setting.

We have numerous verbal conversations with people on a daily basis. We attend meetings and conferences where we listen for streams of important information presented to us verbally. Our organic memories (OM) are not capable of remembering everything that is presented to us. So we supplement fallible OM by using analogue and digital prosthetic memory tools (PMs) such as note books or personal digital assistants (PDAs) to capture important pieces of information. This chapter aims to understand how new forms of prosthetic memory, in particular new generation digital tools can assist people in remembering information they might otherwise forget.

As discussed in chapter 2, in his oft-cited vision of the future, (Bush 1945) proposed Memex, a tool designed to help users remember and index information that they have previously encountered. In the last 15 years this vision has become a reality. Reductions in the cost of digital storage and the emergence of sophisticated recording technologies have led to the development of many different prosthetic memory systems. One type of prosthetic memory tool captures personal experience, whether this is by recording conversations (Whittaker, Hyland et al. 1994; Moran, Palen et al. 1997; Stifelman, Arons et al. 2001) or by recording both sound and visual experiences (DeVaul, Pentland et al. 2003; Dumais, Cutrell et al. 2003; Dickie, Vertegaal et al. 2004), more discussion on these tools is included in Chapter 2. Other prosthetic memory systems serve as general repositories for personal data (Kazman, Hunt et al. 1995; Dumais, Cutrell et al. 2003; Karger and Quan 2004; Gemmell, Bell et al. 2006). Finally there are prospective memory systems designed to help users remember *future* tasks and commitments (Whittaker, Hyland et al. 1994; Rhodes and Maes 2000).

With some notable exceptions (Whittaker and Sidner 1996; Hori and Aizawa 2003; Gemmell, Bell et al. 2006; Sellen, Fogg et al. 2007), however, the majority of this work has focused on the development of proof of concept systems rather than on how and why they are used. One critical issue that has not been systematically explored is the *relationship* between prosthetic memory and natural organic memory. It is obvious that a well-designed prosthetic memory can help users to access information that they may otherwise have forgotten, but there may be reasons why users sometimes prefer to rely on organic memory. For example there is little incentive to use prosthetic memory when one can remember information unaided. There may also be *inefficiencies* associated with using prosthetic memory, as retrieving information from organic memory is extremely rapid. Compare, for example, the ease and rapidity of accessing a familiar contact name from memory with the effort of retrieving it from a poorly organised email archive.

These prosthetic memory trade-offs are illustrated in an exploratory study (Whittaker and Sidner 1996). This study investigated prosthetic memory use for remembering simple conversations. Users had access to 3 types of prosthetic memory tools: (1) pen and paper notes, (2) a dictaphone and (3) Filochat - a device that co-indexed recorded speech and handwritten notes. Although both Filochat and the Dictaphone created accurate verbatim records, users did not always exploit these accurate prosthetic memory devices to retrieve information. Both user confidence in the accuracy of their organic memory, and the efficiency of the prosthetic memory device appeared to influence prosthetic memory usage - although these factors were not systematically varied. For example, users were unlikely to use prosthetic memory when this was laborious, e.g. extracting information from a Dictaphone recording.

Other research also suggests the factors underlying prosthetic memory use are not well understood. A number of sophisticated systems have been developed to help 'meeting capture', i.e. user extraction of key information from a meeting record. However field trials indicate that these digital prosthetic memory tools are not always used in practice to retrieve meeting content (Landay and Kaugmann 1993; Kidd 1994). Often users prefer to rely on their own memory or on paper based prosthetic memory, such as handwritten notes.

4.2 ChittyChatty Interface

This section introduces a new generation prosthetic memory device called ChittyChatty, Figure 4.1. ChittyChatty is similar to other note-taking systems such as (Whittaker, Hyland et al. 1994; Abowd, Harvel et al. 2000; Stifelman, Arons et al. 2001). Like those systems, it supports memory for conversation using temporal co-indexing (Whittaker, Hyland et al. 1994) of handwritten notes and speech (as illustrated in Figure 4.2). The main representation is a blank page where users create notes and/or other visual cues while recording a conversation. Users can follow their normal practice of taking handwritten notes but each pen stroke is temporally co-indexed with the underlying recorded speech. This allows the notes to be used to access the conversation; when users want to re-access recorded speech, they click on a specific note, and the system begins to replay the speech that was being recorded at the moment that note was taken. In this way the notes serve as a visual analogue to the underlying speech, allowing straightforward access to a specific part within that speech. This gives the users a more precise way of accessing a specific part of speech without having to listen to the whole audio again. The notion of coindexing handwritten notes to sequential media for access is similar to the technique used in (Whittaker, Hyland et al. 1994). ChittyChatty runs on any version of Windows Mobile edition on a PDA, making it even more portable than Filochat and easy for taking meeting notes.

Figure 4.1: ChittyChatty Interface.

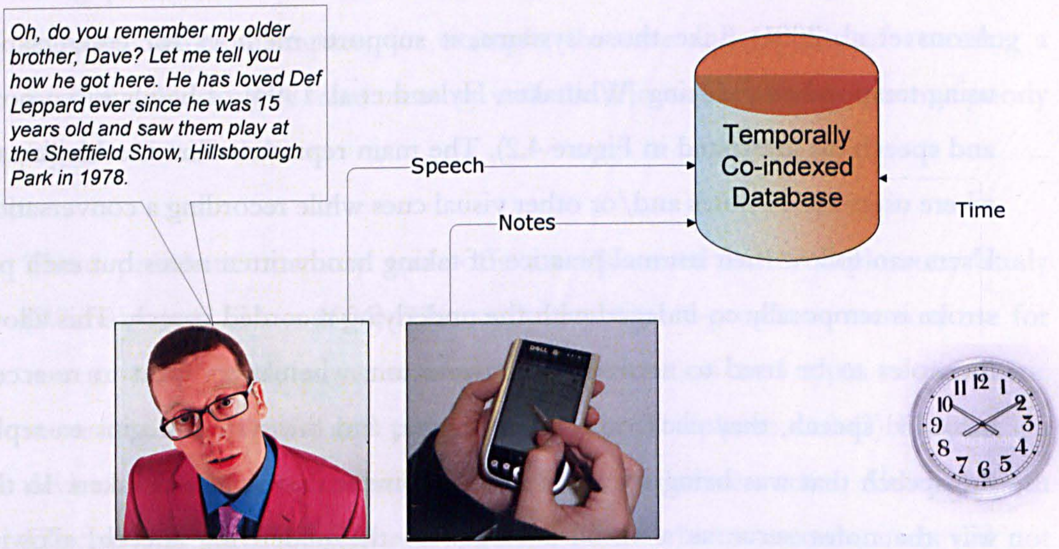


Figure 4.2: Temporal co-indexing of notes and speech.

4.3 Evaluation

To understand various PM properties, a controlled laboratory study was conducted, which investigated the role of 4 factors on prosthetic memory (PM) usage: (a) PM *accuracy*, (b) PM *efficiency*, (c) *user confidence* in the accuracy of their own memory and (d) time from the event to be recalled (*Testing Section*). There are two main research questions. First, what are the *objective costs and benefits* of using PM compared with organic memory (OM). Specifically, what are the trade-offs between the potentially greater accuracy of using PM compared to the efficiency costs incurred in using it? Second, it takes a detailed look at PM *usage patterns*, in particular *when* and *why* PM might be used in preference to OM. These questions are investigated by comparing retrieval for spoken conversations using various PM devices, including pen and paper, a dictaphone and a new generation of Filochat-like device (ChittyChatty). Examination is carried out on how these different PMs are used for retrieving different types of information, over different periods of time, and how this compares with OM usage.

Specific questions addressed in this evaluation study are:

i. What are the *benefits* of using PM compared with OM? Are digital memories more accurate than OM, and are they quicker and more efficient to access?

ii. How does PM use relate to *people's perceptions* of the accuracy of their OM? For example if users are confident they can retrieve information unaided, we cannot expect them to use PM.

iii. Does PM use depend on the *characteristics of the PM device,* in particular how easy it is to retrieve information from that device? Obviously, a greater use of a PM device that makes retrieval straightforward is expected. In contrast, a hard-to-use PM device such as a Dictaphone might force users to fall back on OM, even when they are unsure they can retrieve information unaided.

iv. Does PM use depend on the *type* of information being retrieved? E.g. it might be expected that PM use is more prevalent for complex verbatim information which is hard to recall unaided.

v. Finally does PM use change over time? Do users become more reliant on PM for events that are further in the past? As OM degrades greater use of PM devices might be expected.

This study investigated memory using three different types of prosthesis: Dictaphone (DP), which was a Sony digital voice recorder ICD-P320 with standard functions such as record, play, pause, stop, rewind and forward for normal audio playback; Pen and Paper (PP) and our new generation system called ChittyChatty (CC) which is described in detail above. The goal was to test how these different PMs helped users remember everyday conversations. A series of conversational stories were read aloud to users (Appendix A), asking them later to retrieve information from those stories. Users were given either a PM, to help them remember, or they could choose to rely on unaided OM.

4.4 Differences in Efficiency and Accuracy between PM and OM

Our three different types of PM have different properties. PP notes are a schematic and incomplete record of what was said, whereas DP and CC offer verbatim records. Retrieval efficiency is also different for these PM devices. Extracting information from PP or written CC notes is efficient because the eye can rapidly scan text to identify information. The DP is highly inefficient because of the serial nature of recorded speech, making it hard to identify important information (Whittaker, Hirshberg et al. 2002). CC however should support reasonably efficient access to the underlying speech record, using handwritten notes and other visual cues should allow users to identify regions of speech to access and listen to. Of course, all three PM devices contrast with OM which is efficient to access but fallible as summarised in Table 4.1.

Device / Characteristics	Accuracy	Efficiency
Organic Memory (OM)	x	√
Pen & Paper (PM)	x	√
Dictaphone (DP)	√	x
Chitty Chatty (CC)	√	√

Table 4.1: Prosthetic Memory Tool Properties.

4.5 Time as a Testing Section

The entire evaluation study consisted of 3 Testing Sections. The first Testing Section consisted of an introduction, training on the prosthesis, exposure to the conversational stories and initial memory testing. This took about fifty minutes. The second Testing Section – a week later - involved remembering certain aspects of the stories presented at the first Testing Section and lasted about thirty minutes. The last Testing Section took place a month after the first and again involved retrieval of information presented at the first Testing Section and lasted about thirty minutes.

Before listening to each story users were given a single PM (CC or DP or PP) or be in an OM condition but they also had OM at all times, which they could choose to use instead of the allocated condition. They were instructed to record or remember the story either with the assistance of a PM or OM using memory techniques they would normally use. Their memory was tested at 3 different Testing Sections: same day, 1 week later and 1 month later. On each test occasion users had the same prosthesis support as when they heard the original story. For instance if a user had access to CC at the first Testing Section when listening to the story, the user was given CC again at remaining Testing Sections with the same story.

A critical research question was whether and when people made use of PMs as opposed to relying on unaided OM. So, even when users had access to a PM, I made it clear that they were not compelled to use it, and it was noted when PMs were used in preference to OM.

4.6 Stories and Test Questions

The 3 stories were intended to simulate real-life conversations between two old friends who had just bumped into one another after a period of several years (see Appendix A). The stories contained a mixture of facts and fiction equally distributed within each story. They were written in plain English so they could be easily understood; they did not contain any unfamiliar or unusual terms. Extensive pilot studies were conducted with the stories. User comments indicated that they were enjoyable to listen to, as well as achieving their objective of simulating real-life conversational experiences. The average story time was 3.20 minutes.

An example fragment of one story was the following (for other stories, see Appendix A):

"Oh, do you remember Barrington Vera-Smith from the sixth form? I ran into him the other day and he was telling me he lived in Naseby now, which is just a few miles south of Harrogate, in a beautiful two bedroomed cottage with his wife, Jenny. They've just had their tenth wedding anniversary, which they spent in Cornwall...Barrington and Jenny are nationally recognised dog breeders. Their dogs, Richard and Susan, are their treasures, Barrington told me. I'm sure you've heard of cocker spaniels and I'm certain you know what a poodle is. Well, Richard and Susan are a cross between the two. They are cocker-poos...Barrington was telling me for almost an hour about his dog breeding business and how Richard and Susan are widely respected as professional show dogs. So far they have won three 1^st^ prizes in the toy category and a couple of best in shows. Barrington admitted it was a struggle to get to this standard."

After hearing the story, users were asked two different types of recall question: *verbatim* questions requiring them to remember specific facts from the story, and *gist* questions probing their ability to combine approximate information from different parts of the story. Memory research suggests that gist information is easier to remember unaided, so we should predict greater PM use for verbatim information (see Appendix A).

The above story generates 3 verbatim questions: (1) *"How many first prizes have Richard and Susan won in the toy category?"* (2) *"What breed were Richard and Susan?"* and (3) *"How many prizes for "best in show" did Richard and Susan win?"* One verbatim question was asked the same day, another 1 week later, and the final one a month later. There were also 3 gist questions asked at the same Testing Sections: (1) *"Does Barrington give an impression that dog breading is hard work?"* (2) *"How did the* *speaker feel about talking to Barrington?"* (3) *"Does Barrington feel confident about his knowledge on running a successful business?"*

4.7 Procedure

The whole evaluation study was run using a custom built website (see Appendix A). Users were first given a general description of the experiment, the stories and the different types of questions that they would be asked as part of each session.

Users then were given a brief web-based, hands-on tutorial providing detailed descriptions of each memory prosthesis and procedures for the experiment. They carried out 4 practice tasks, one with each prosthesis, and one with OM. The practice tasks were similar to those used in the experiment. Users were allowed to proceed to the actual experiment only if they felt confident with each PM and the procedure. They also had to complete each practice task successfully before being allowed to proceed.

4.8 Experimental Tasks

Users heard a story with CC, DP, PP or OM depending on the experimental condition. To control for story/Retrieval Method sequence effects, the order in which users received stories were counterbalanced , the PM they used to carry out each task, and the type of question (verbatim/gist) they were asked. Users answered questions on web based forms (see Appendix A).

Before answering each memory question users were asked to evaluate their *confidence* in their ability to answer the question without using PM. The confidence question was asked after the user had read the memory question but *before* they answered it: *"How confident are you that you can remember the answer to this question* *without using your [memory prosthesis name]?"* Responses were generated as 5-point Likert scales.

Users then tried to answer the question. In all conditions the retrieval time was recorded, i.e. how long it took users to answer each question. Whether users relied on OM or PM to answer the question was also noted. With the various PM devices people might choose to use OM if they felt that they could answer the question without recourse to the PM device. For example a user who is confident their memory is accurate may choose not to use a Dictaphone for retrieval because they feel it is inefficient to do so. This chapter reports data where users chose to use their OM ignoring the OM condition. The results from purely OM condition are reported in Chapter 5.

After experience in using each PM, users received a brief survey, asking them to rate retrieval support for the PM or OM they had just experienced (see Appendix A). They were asked questions about (a) accuracy, (b) efficiency, (c) usability, and (d) enjoyability. For example a user who had just completed a session with pen and paper would be asked to rate each of these categories for pen and paper as a retrieval tool. Responses were generated as 5-point Likert scales. In addition users were asked open-ended questions about what users perceived to be the main differences between each PM and OM, and why they preferred one PM or OM to another, on the basis of these categories. Results are discussed in the subjective data section of this paper.

Finally users were asked to complete an order of preference survey: they ranked each PM and their OM in order of preference. In addition users were asked open-ended questions about what they perceived to be the best features of each PM and OM, why they preferred one PM over another or whether they preferred

OM. Spontaneous comments made during the experiment about PMs, OM and the tasks were also recorded.

4.9 Participants

Twenty five users (14 women and 11 men, ranging in age from 23 to 55) took part. Users were volunteers consisting of university researchers, administrative and management staff, as well as other professionals from public and private sectors. Users had no prior knowledge of the project or our experimental hypotheses. None of the users had any experience of using CC or DP, but obviously all had extensive experience of OM as well as PP.

4.10 Variables and Measures

The following data was collected:

- Efficiency i.e. retrieval time for each question;
- Actual retrieval accuracy for each question;
- User retrieval confidence scores for each question;
- Subjective ratings for each PM after each Testing Section with that PM, rating accuracy, efficiency, usability and enjoyability;
- Comparative subjective rankings of all prostheses and organic memory for overall utility;
- Qualitative comparisons of the different PMs and OM.

Retrieval accuracy was scored as follows. Each answer was evaluated against a scoring scheme using pre-specified target answers for each question. Two coders blind to the experimental conditions listened to all the stories twice and constructed a set of target answers consisting of target keywords that needed to be present in each answer. Each answer contained a number of keywords and context.

Criteria for accuracy were strict, with accuracy determined at the keyword and context level. If user answers included all the keywords (or their synonyms) and context, they received a maximum score of 5. Partially correct answers were defined as either (a) containing all keywords, but inaccurate context, or (b) accurate context and incomplete set of keywords. Scores ranged from 0-5 for each answer. Scoring was carried out independently by two judges, and disagreements were referred to a third judge for resolution.

4.11 Hypotheses

H1: Comparing the costs and benefits of PM and OM.

These hypotheses address the objective differences between PM and OM for retrieval accuracy and efficiency.

- H1.1 – *Accuracy*: It was expected for PM to be more accurate overall than OM. This advantage of PM over OM should be greater at longer Testing Sections, as OM decays over time.
- H1.2 – *Efficiency*: Given the ease of accessing OM, it was expected that OM would be more efficient overall than PM.
- H1.3 – *Effect of device type:* It was also expected that there were accuracy and efficiency differences between different types of PM. On the one hand, it could be predicted that CC and DP would be more accurate than PP. On the other, I expected PP to be more efficient given the ease of scanning information from notes.

H2: Usage Hypotheses: identifying the factors that determine PM usage

- H2.1 *Confidence:* It was expected that PM usage could be affected by users' *perception* of the accuracy of their OM, i.e. they are more likely to use PM when they are not confident they will be able to retrieve unaided

- H2.2 *PM type:* It was expected for PM usage to be affected by the *characteristics of the PM device* specifically their accuracy and efficiency. It was expected that there would be less use of PP because this doesn't usually offer a verbatim record. Also for devices that do provide verbatim memories it was expected that there would be greater use of CC than DP because CC provides more efficient access.
- H2.3 *Type of information retrieved:* It was expected that PM usage could be affected by the *type* of information being retrieved. It was expected that PM use would be more prevalent for complex verbatim than gist information.
- H2.4 *Testing Section:* It was expected that PM usage would be affected by Testing Section. As OM degrades at longer Testing Sections (such as a week or a month), it might be expected that there would be greater use of PM devices.

4.12 Results

Before proceeding with parametric statistical analysis, I applied Kolmogorov-Smirnov (Sheskin 2007) to test for normality in data collected. The results of this test have confirmed that data collected is normally distributed.

H1: Comparing the Benefits of PM and OM Usage

First, the relative benefits of all Retrieval Methods in terms of both accuracy and efficiency (i.e. time to retrieve the answer) were compared. Then OM scores from those cases where users had access to a PM but nevertheless chose to rely on OM[1] were obtained.

First overall differences between all Retrieval Methods by combining results for all prostheses (i.e. combining PP, CC and DP scores together) were analysed, and comparing these with OM. A two-way ANOVA was conducted with

[1] All participants contributed data to all means in comparison when participants could have chosen exclusively to use PM or OM.

independent variables being: 1) *Retrieval Method* – i.e. whether a prosthesis was used or whether the user relied on OM alone; and 2) *Testing Section* – i.e. length of time since the user had heard the story told. The respective dependent variables were accuracy and efficiency.

H1.1 Accuracy and efficiency scores are shown in Figure 4.3. Confirming H1.1, there was a significant difference among different Retrieval Methods ($F(1,893)=7.0$ $p<0.008$). This confirmed our observations that PM is more accurate overall than OM. Furthermore, as predicted, there was a significant interaction between accuracy and Testing Sections ($F(2,893)=12.9$, $p<0.0001$). Planned comparisons revealed that PM was more accurate than OM for retrieval after the initial day, i.e. when I combined retrieval scores for 7 and 30 days and compared OM and PM for recall accuracy (Bonferroni, $p < 0.0001$). Figure 4.3 shows how OM accuracy drops 7 and 30 days.

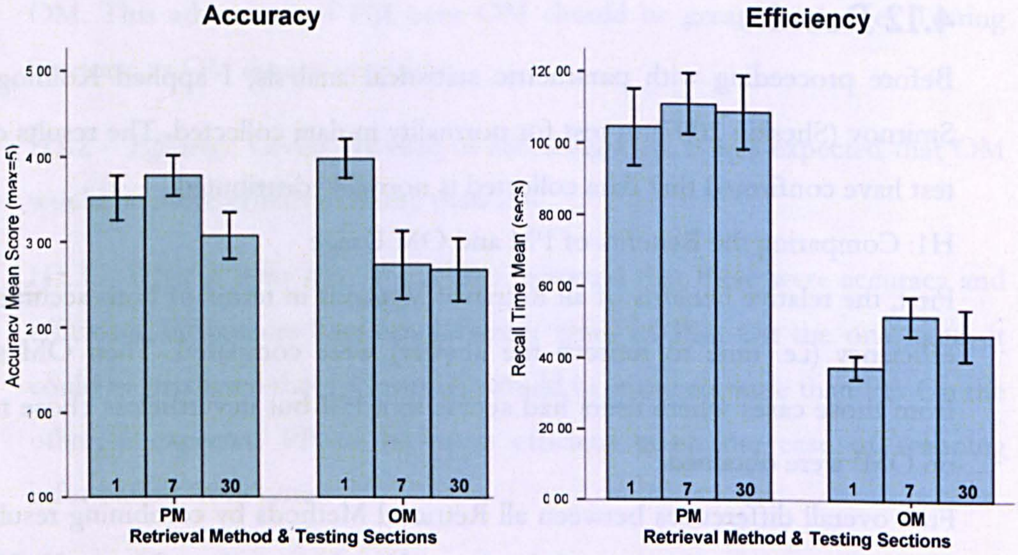


Figure 4.3: Actual Accuracy and Efficiency for different Testing Sections and Retrieval Methods.

H1.2 Confirming H1.2, PM was less efficient; retrieval was slower when people used PM ($F(1,893)=263.7$, $p<0.0001$) as illustrated by Figure 4.3.

H1.3 Next different prosthesis types with OM, again at the 3 different Testing Sections see Figure 4.4 were compared. A two-way ANOVA was conducted with Retrieval Method (PP, CC, DP, OM) and Testing Section (same day, 7 days, 30 days) as independent variables. The dependent variables were accuracy and retrieval time. Figure 4.4 shows accuracy and retrieval efficiency broken down by the specific PM device used, and compares this with OM.

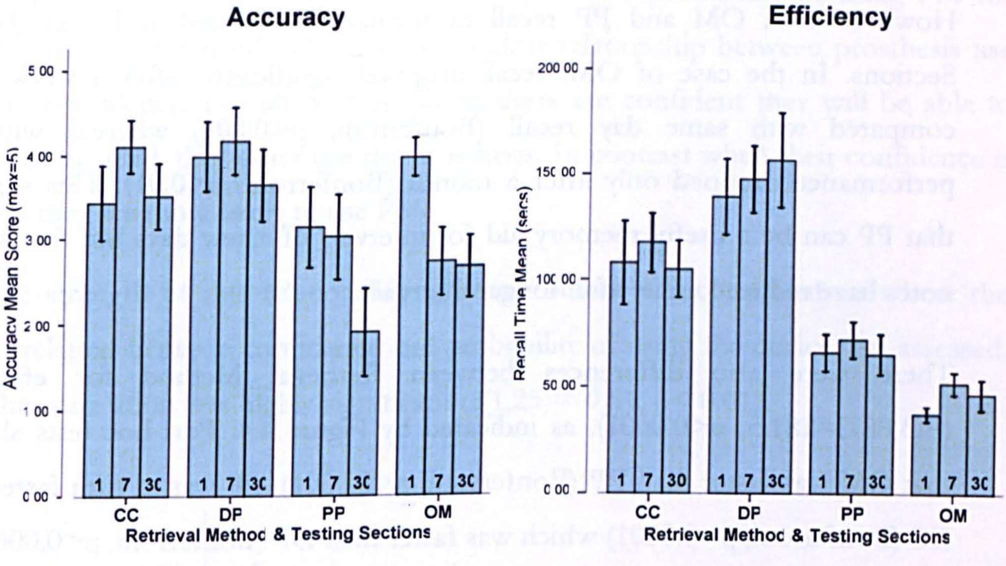


Figure 4.4: Actual Accuracy and Efficiency over different Testing Sections and Retrieval Methods.

As predicted there were differences between the different Retrieval Methods used for recall ($F(3,887)=18.8$, $p<0.0001$). Planned comparisons showed that accuracy using CC and PP was greater than OM (Bonferroni, CC vs. OM, $p<0.05$, DP vs. OM, $p<0.001$). But contrary to predictions, OM performed better than PP (Bonferroni, $p<0.001$). These results can be explained as follows:

CC and DP perform better than OM and PP because they contain a verbatim recording of what was said. However people using PP may perform worse than OM because people only rely on OM when they are very confident that they can recall information unaided.

Additionally, there was evidence of an interaction between Retrieval Method and Testing Section on recall accuracy ($F(6,887)=6.3$, $p< 0.0001$). Further post-hoc tests for each Retrieval Method showed equal recall accuracy at the different Testing Sections for CC and DP (Bonferroni, all $p>0.10$) – presumably because both prostheses supported permanent memory for exactly what was said. However both OM and PP recall performance degraded at longer Testing Sections. In the case of OM, recall dropped significantly after a week when compared with same day recall (Bonferroni, $p<0.001$) whereas with PP performance dropped only after a month (Bonferroni, $p<0.01$). This suggests that PP can be a useful memory aid for intervals of a few days but that simple notes have reduced utility after longer intervals.

There were also differences between Retrieval Method for efficiency ($F(3,887)=181.5$, $p<0.0001$), as indicated by Figure 4.4. Post-hoc tests showed that OM was faster than PP (Bonferroni, $p<0.0001$). PP was in turn faster than CC (Bonferroni, $p<0.0001$) which was faster than DP (Bonferroni, $p<0.0001$).

To sum up, these analyses show that there are costs and benefits to using PM. Using PM usually guarantees better recall accuracy but access via PM takes longer, especially when CC and DP are used. Furthermore there are differences between different types of PM with CC and DP being more accurate than PP at longer Testing Sections although PP is adequate for short Testing Sections.

4.13 Factors Affecting Prosthesis Choice

Our next analysis examined what determined PM usage. For each retrieval question when a PM was available it was noted whether users exploited the PM or relied on OM. A four-way ANOVA was conducted with the independent variables: Confidence (on a scale of 1-5); Retrieval Method (CC, DP or PP); Information Type (Verbatim vs. Gist); and, Testing Section (1, 7, and 30 days). The dependent variable was Probability of PM Usage.

H2.1 was confirmed. The ANOVA showed an effect for confidence ($F(4,810)=50.8$, $p<0.0001$). Figure 4.5 depicts the probability of using PM for different confidence levels. It shows a clear relationship between prosthesis use and confidence. For all devices, when users are confident they will be able to recall unaided, they don't use the prosthesis. In contrast when their confidence is low they are more likely to use PM.

The strength of this relationship is borne out by a second analysis where the correlation between confidence and probability of using the device was assessed. The correlation was highly significant ($r(1,25)=-0.54$, $p< 0.01$).

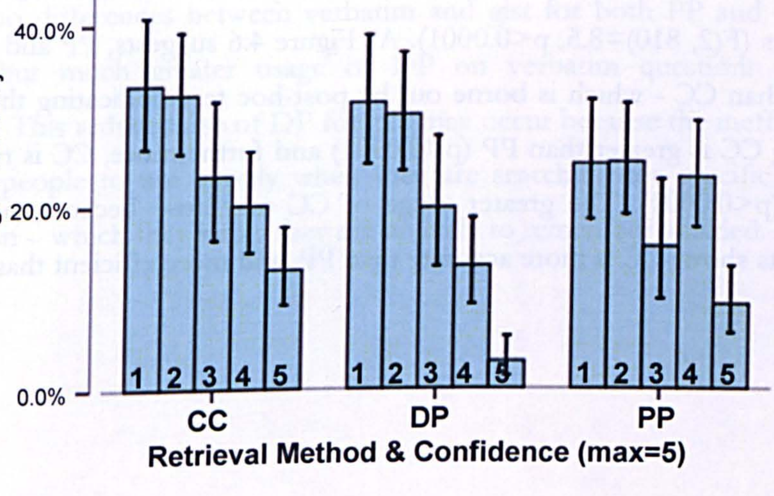


Figure 4.5: Percentage Accuracy of PM per Recall Method.

As Figure 4.5 suggests, there was also an interaction between confidence and recall method for PM usage ($F(8,810)=3.5$, $p<0.001$). There was a significant difference between CC and DP with the CC maintaining greater likelihood of usage when users were very confident (conf=5), according to post-hoc Bonferroni tests ($p<0.001$). There were also significant differences between PP and DP, with PP having a greater likelihood of usage than DP when users are very confident (Bonferroni $p<0.03$). This may be because the efficiency of accessing PP means there is little cost associated with using it even when users are confident.

There were also differences in likelihood of usage of different PMs for lower confidence levels. When users did not feel confident (conf = 1 or 2), post-hoc Bonferroni tests revealed significant differences between CC and PP, with CC showing greater likelihood of usage ($p<0.001$) as well as between DP and PP with DP showing greater likelihood of usage ($p<0.0001$). This may be because people feel that the lack of detail and context in PP makes it futile to employ PP when confidence is low.

H2.2 was confirmed. Users are more likely overall to use some devices than others ($F(2, 810)=8.5$, $p<0.0001$). As Figure 4.6 suggests, PP and DP were used less than CC - which is borne out by post-hoc tests indicating the likelihood of using CC is greater than PP ($p<0.0001$) and furthermore, CC is more used than DP ($p<0.0001$). The greater usage of CC may arise because (as our previous results show) CC is more accurate than PP, and more efficient than DP.

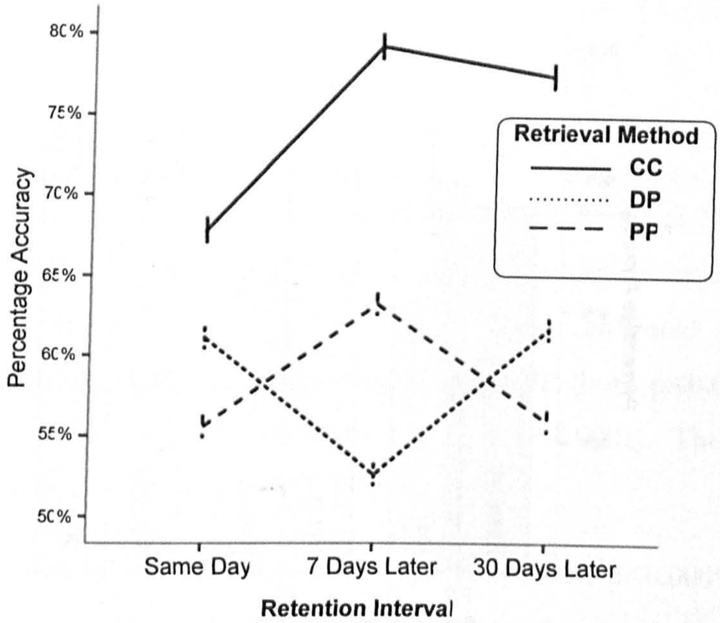


Figure 4.6: Percentage Accuracy of PM Usage at Different Testing Sections.

H2.3 was not confirmed. As Figure 4.7 indicates, there were no overall differences in likelihood of using PM for different question types ($F(1,810)=1.1$, $p>0.10$). However, there was an interaction between device and question type ($F(2,810)=3.8$, $p<0.05$). Figure 4.7 suggests this is because people are less likely to use DP for gist questions. This is borne out by post-hoc Bonferroni tests showing no differences between verbatim and gist for both PP and CC (both $p>0.10$), but much greater usage of DP on verbatim questions than gist ($p<0.001$). This reduced use of DP for gist may occur because the inefficiency of DP leads people to use it only when they are searching for specific verbatim information – which they know they are unlikely to remember unaided.

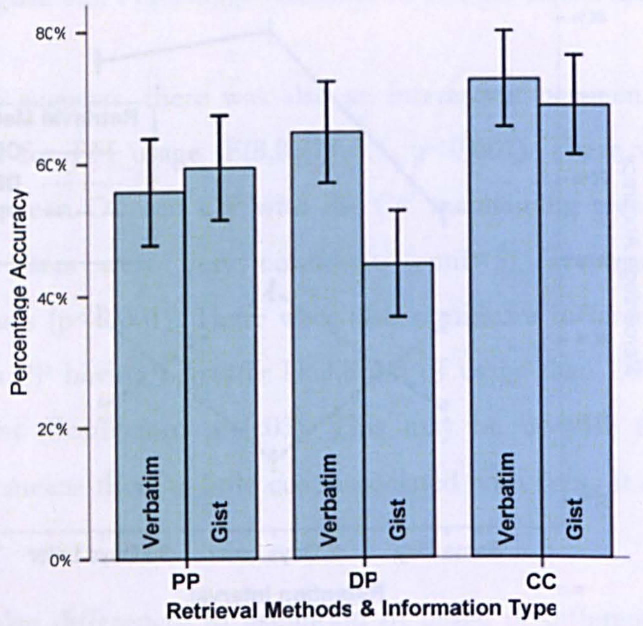


Figure 4.7: Percentage Accuracy of PM Usage for Different Information Types.

H2.4 was not confirmed: Overall PM usage did not depend on Testing Section. ($F(2,810)=0.17$, $p>0.10$). This may have been because people were more inclined to use CC in later sessions but these increases were counterbalanced by reduced use of PP which users felt to be less accurate.

4.14 Subjective Data and User Comments

User evaluations of the accuracy, usability, efficiency and enjoyability of different Retrieval Methods (i.e. for PP, CC, DP and OM) were also analysed. These are shown in Figures 4.8 and 4.9. I carried out two-way ANOVAs with Retrieval Method and Testing Section as independent variables and accuracy, efficiency, usability, enjoyability, and overall rating as dependent variables.

In addition users were asked open-ended questions about what users perceived to be the main differences between each PM and OM, and why they preferred one PM or OM to another on the basis of these categories. Our findings are as follows.

4.14.1 Accuracy

Figure 4.8 shows the subjective accuracy for different Retrieval Methods. The two-way ANOVA results show that Retrieval Method is significant ($F(3, 290)=145.2$, $p<0.0001$). Post-hoc Bonferroni tests confirmed that people thought that CC and DP were more accurate than PP (both $p<0.0001$) and in turn all PM were more accurate than OM (all $p<0.0001$). There were no differences between CC and DP ($p>0.99$).

The Testing Section was also significant ($F(2, 290)=18.1$, $p<0.0001$). Post-hoc Bonferroni tests confirmed that accuracy on the same day was better than 7 days later ($p<0.002$) and 30 days later ($p<0.0001$). 7 days later in turn was better than 30 days later ($p<0.05$).

As expected, our observations confirm that the subjective accuracy provided by the users and the actual accuracy results collected during the experiment are very similar. The strength of this relationship is borne out by a second analysis where I assessed the correlation between actual accuracy and subjective accuracy. The correlation was highly significant ($r (1, 25) = - 0.45$, $p< 0.01$).

User comments confirmed our findings as follows.

CC was viewed as accurate and this perception didn't change over different Testing Sections. "*With [CC on the 30 day session] the entire story recorded as well as having notes means you get all the little details as well as any key facts you make notes of*"

DP was also though to be accurate, but users were aware of its inefficiencies. "*The [DP] was 100% accurate, if you can be bothered faffing with it.*"

In contrast PP was thought to be good at first. "*[PP] was good, but there may have been details which I didn't write down fully*", but users found PP to be less useful at later Testing Sections "*With the [PP] I had missed out some important parts and even the bits I did have written down I couldn't be sure were 100% accurate.*"

The same was true for OM, which users thought was inaccurate at later Testing Sections. "*With [OM] I just can't remember*"

4.14.2 Efficiency

Figure 4.8 illustrates the main subjective efficiency effects for different Retrieval Methods. The two-way ANOVA shows that the retention method was significant ($F(3, 290)=39.4$, $p<0.0001$). Post-hoc Bonferroni tests confirmed that users found the CC most efficient (all $p<0.001$) in comparison to DP, PP and OM. But post-hoc test also showed that PP was more efficient than DP ($p<0.04$) but there was no evidence to suggest PP was more efficient than OM ($p>0.1$).

The Testing Section was also significant ($F(2, 290)=6.2$, $p<0.002$). Post-hoc Bonferroni tests confirmed that OM was most efficiency on the same day. There was no evidence to suggest it degraded 7 days later ($p>0.1$) but there was evidence that it degraded 30 days later ($p<0.002$). The degradation between 7 and 30 day Testing Sections was not significant ($p>0.1$)

Subjective efficiency was further compared to the actual efficiency. The correlation was not significant ($r(1, 25)=0.1$, $p>0.1$).

User comments again suggest that CC was viewed as efficient at all retrieval intervals. "*[CC] was very easy to use by now. Not too many functions to distract.*"

People were aware with DP that they had a complete record of the story but were also dissatisfied with the trouble it took to find information. "*[DP] was...knowing there is a full record...but actually retrieving info is more longwinded than other methods.*"

With PP notes people realised that these became inefficient as time passed. "*...with [PP] it was more frustrating coming back to them now [after a month] as there are lots of gaps.*"

Similarly OM was unsatisfactory after longer Testing Sections. "*With [OM] I felt like I couldn't rely on the little information that I did remember and would have preferred to have a second method to back it up.*"

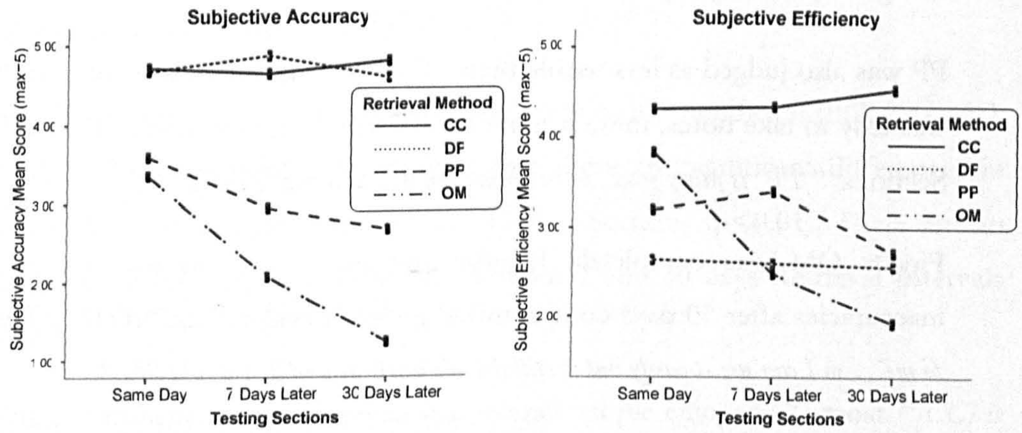


Figure 4.8: Subjective Accuracy and Efficiency.

4.14.3 Usability

As Figure 4.9 illustrates, and two-way ANOVAs confirm, the Retrieval Methods are significant for subjective usability results ($F(3, 290)=19.9$, $p<0.0001$). Post-hoc Bonferroni tests confirmed that users thought that CC was viewed as more usable than DP, PP and OM (all $p<0.0001$). There was no evidence to suggest PP was more usable than DP ($p>0.1$). The superiority of CC in the subjective usability category can be explained by its flexibility and the novelty of this method.

The ANOVA results showed that Testing Section was also significant ($F(2, 290)=4.9$, $p<0.01$). Post-hoc Bonferroni confirmed that the only significance is between same day and 30 days later ($p<0.001$).

User comments confirmed that CC was clearly viewed as most usable and there were many positive comments about ease of use, the interaction of notes and audio. "*[CC] is dead simple, just write and then click to listen. Quite good being able to see what you wrote as the audio is playing.*"

DP was also thought to have simple controls, but to be much less usable because of the effort needed to extract relevant information. "*[DP] is simple but a pain having to rewind and fast forward.*"

PP was also judged as less usable than CC, but for different reasons. Although it was easy to take notes, these weren't all that accurate, especially at longer Testing Sections. "*[PP] is fairly good...but information was of a poor quality.*"

Finally OM was completely familiar and hence easy to use, but again its inaccuracies after 30 days compromised people's evaluations. "*[OM] is hard wired to me... so I can use it easily but it still doesn't work as well as I would like.*"

4.14.3 Enjoyability

The enjoyability results are illustrated in Figure 4.9. Two-way ANOVA tests confirm that the Retrieval Methods for enjoyability are very significant ($F(3, 290)=39.9$, $p<0.0001$). Post-hoc Bonferroni tests confirm that people found using CC most enjoyable in comparison to DP, PP and OM (all $p<0.0001$). However the results also show that comparing PP and DP, users found PP more enjoyable than DP ($p<0.05$). There were no significant differences between the enjoyability levels of PP and OM ($p>0.1$).

ANOVA results also confirm that Testing Sections are significant ($F(2, 290)=6.2$, $p<0.01$). Post-hoc Bonferroni showed that there are significant differences in enjoyability between same and 30 day Testing Sections ($p<0.01$). There are no significant differences in enjoyability between 7 and 30 days Retrieval Intervals ($p>0.1$).

In user comments it was observed that overall people enjoyed CC most. "*[CC] is a new and exciting tool.*" and " *I thought [CC] concept was excellent...great stuff. I want one!*"

Although it is accurate, DP was viewed as inefficient. It frustrated a lot of users and did not provide an enjoyable experience. "*I got bored of [DP]. It was frustrating when I couldn't find a passage of the story...*"

Although PP notes were perceived as enjoyable at first, their appeal degraded when gaps in context and mood were revealed as time passed. "*I thought [PP] notes would be more useful... but I didn't have the right information to help with the questions.*"

Although OM enjoyability was high at first, it also rapidly degraded leaving users to feel that using OM was just another un-enjoyable chore. "*After 30 days it was a struggle to recall things [with OM] therefore not the most enjoyable.*"

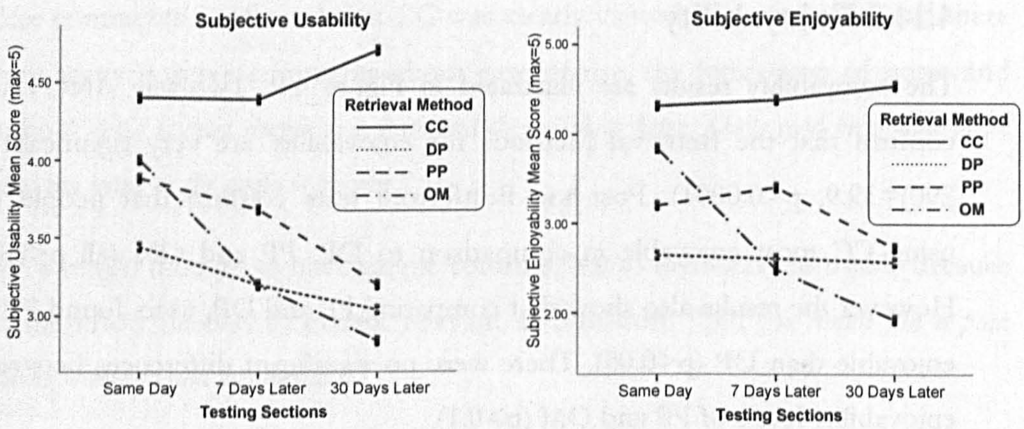


Figure 4.9: Subjective Usability and Enjoyability.

4.14.4 Overall Ratings

Finally the overall ratings that users gave at different Testing Sections were evaluated, see Figure 4.10. The two-way ANOVA showed a significant difference between Retrieval Methods ($F(3, 388)=105.2$, $p<0.0001$). Post-hoc Bonferroni tests confirmed that CC received higher ratings (all $p<0.0001$) than other Retrieval Methods. Furthermore post-hoc tests showed that PP was more highly rated than DP ($p<0.001$) and DP was more highly rated than OM ($p<0.05$).

However ANOVAs showed that there is no significant difference between Testing Sections ($F(2, 288)=0.006$, $p>0.99$).

Overall user comments again supported the findings that CC was hugely preferred. Not only was it fun and usable, but it was accurate as well - giving users a strong sense of control. "*[CC] was very easy…and makes you feel like you are in control!*" and "*[CC] was easy to use, easy to find information. Two sources of info available very easy to access.*"

DP was regarded as accurate, but just too laborious. "*The [DP] is accurate but it is dependent on me being able to find the answers.*"

PP was also negatively viewed. Although notes were relatively easy to generate, their utility was reduced in later sessions, as people relied on their OM for interpreting their notes. "*My notes [PP] were a mess, and were contextual with my own memory, which itself had faded.*" PP notes also excluded affective and emotional information that people felt was useful to recall. "*[PP] doesn't capture nuances of mood and emotion like audio recording does.*"

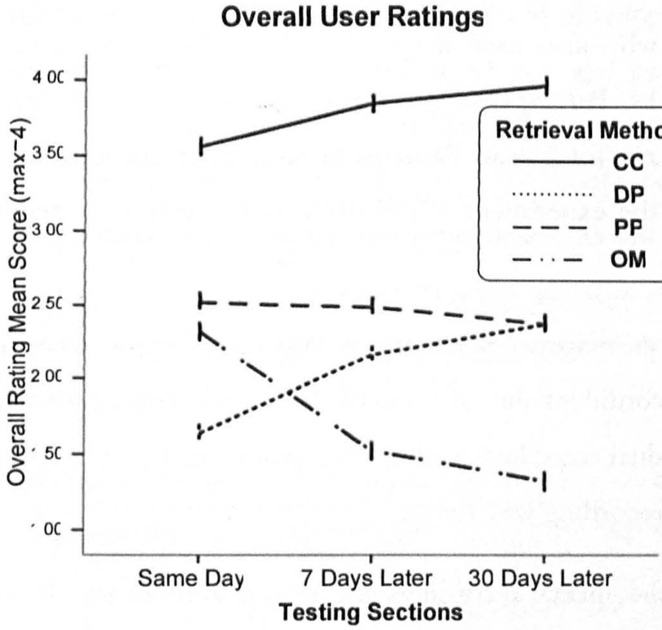


Figure 4.10: Overall User Ratings.

However, even with such negative feedback, PP was rated higher than DP overall, which suggests that people prefer efficiency to accuracy.

OM was thought of as familiar, and reasonably effortless, but inaccurate at longer Testing Sections. "*[OM] is good because you can just concentrate on remembering, but not very reliable, if you lost concentration then you could miss important points.*"

4.15 Long-term ChittyChatty Trials

To understand how ChittyChatty could help people with their memory tasks in real life situations, I studied long-term use. 3 participants were recruited. They all were very keen on trying out ChittyChatty in their work environment. Two of the users were researchers working in the field of ethnography and anthropology heavily relying on their notes and interviews to perform well in their everyday research tasks. Another participant was a business manager and a part-time magistrate, who also used a lot of notes to help with everyday work-related memory tasks. Participants' age ranged between 27 – 55 years. The long-term experiment ran for 1 year. Participants were interviewed at the beginning and at the end of the experiment. Their dated notes were collected at the end of the study.

Long-term participants generated a huge number of recordings. Due to the nature and confidentiality of some of those recordings, participants shared a total of 34 individual recordings with a time span from 1 month to 1 year. The average length of a recording was 45min.

I analysed the speech recordings and textual annotations. It was clear that over time people changed the way that they annotated digital recordings. At the beginning, people used CC devices as if they were traditional pen and paper. They tried to capture a lot of key information as detailed textual annotations. However, after a few months, as participants used ChittyChatty more in their everyday work environments, their note taking behaviours changed. They started taking fewer notes and instead started using new representations. For instance, one participant said when they were attending meetings, after a while, they would draw asterisk instead of writing text. The asterisks were a short-hand way to record the presence of important information in the recording, allowing them to create an index into a complete record to help them access important information later. By using this type of shorthand convention, users switched to taking fewer notes, without loosing access to valued information:

"...I didn't have time to write everything, so I just started drawing asterisks and other symbols to help me remember which part was important...it works for me..."

Figure 4.11 illustrates textual annotations made at talks and meetings during the period of the first few months after participants started using ChittyChatty to record work related information. All four a), b), c) and d) snippets show that at the ChittyChatty notes were similar to traditional pen and paper notes. They were detailed, verbose and fairly complete. In the interviews people said they felt they had to capture as much textual information as they could, because this is the way they did it with their traditional pen and paper notes. As one participant said: *"[with ChittyChatty] I began by trying to capture the whole content in writing, I eventually reverted to writing down only the topics that the lecturer/speaker elaborated on and relying on audio recording for the content..."*

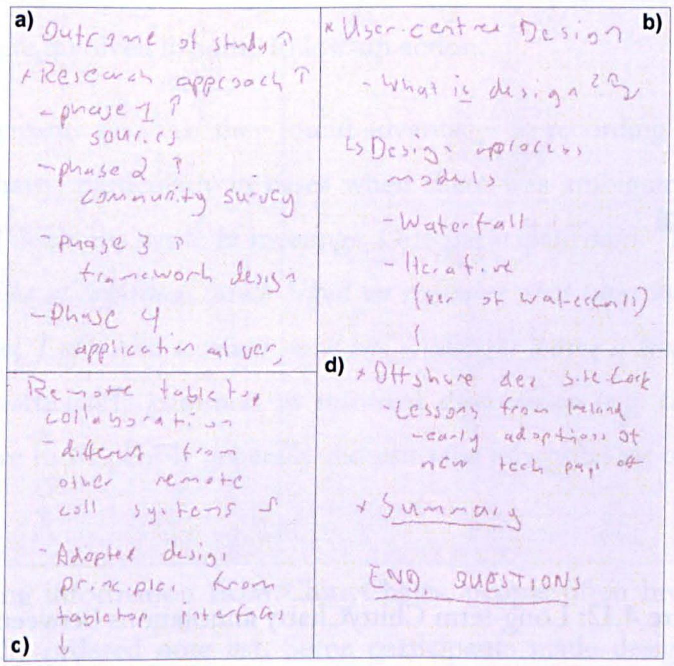


Figure 4.11: Naturalistic ChittyChatty annotations within the first 1-4 months.

After a couple of months, when participants became more accustomed to taking notes on ChittyChatty their note taking behaviour changed. The textual annotations became more sketchy. Figure 4.12 illustrates the shift from systematic note taking, to more schematic note taking. All four sections e), f) g) and h) show the notes of one participant who moved to creating sketches as information representations. They said: *"I took a bit less [notes] with time. But mostly I began to draw and make different kinds of shapes rather than just written words."*

Over time, the overall volume of notes decreased by almost 50% for all participants. This suggests the view that people started taking significantly less notes as they continued using ChittyChatty in their meetings.

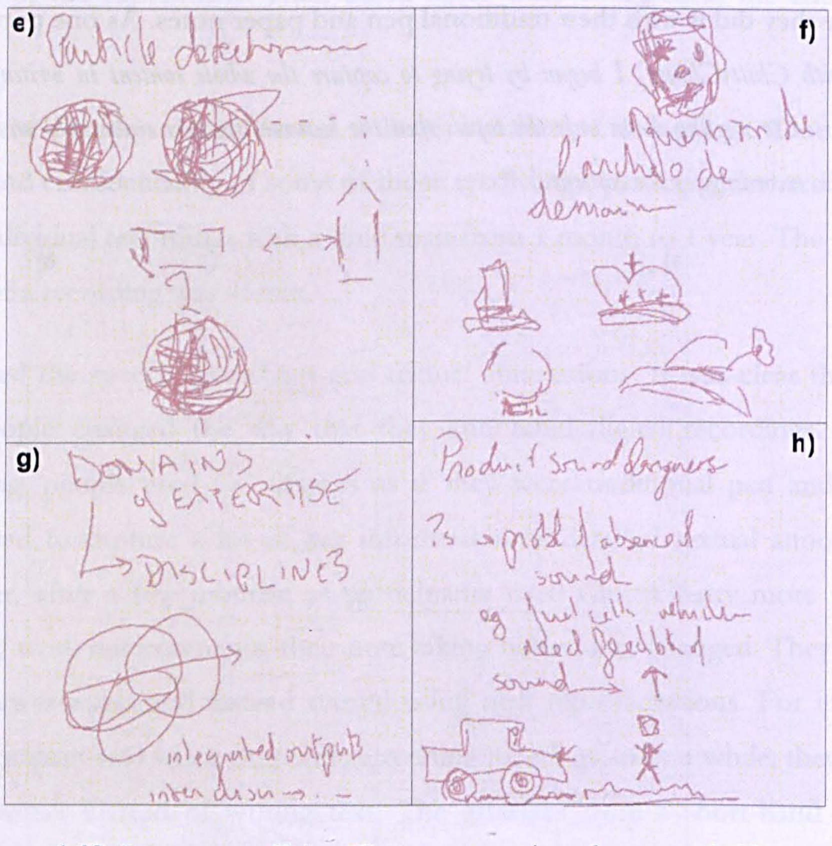


Figure 4.12: Long-term ChittyChatty annotations between 6-12 months.

The interviews indicated that people used ChittyChatty to help them capture *"subjects of conversation during meetings"*. They also used it to capture *"...the way the speakers presented their work, how they structured the thoughts they wished to share and in what order"*. Some participants used ChittyChatty not only for work related information: *"I also used it to record some ambient sounds such as a fire alarm, or a small bit of dialogue during a lunch with a friend"*.

At the beginning of the study, participants tended to look at their notes straight after meeting, just to make sure that they captured important parts of the meeting. However over time, most participants said that they tended not to look at their notes immediately after capture, instead they tended to *"briefly glance"* at their notes just before their next meeting, to remind themselves what happened in the previous meetings.

Participants repeated using ChittyChatty for 3 types of meetings: 1) lengthy discussions which would be hard to capture on paper alone; 2) talks, particularly to capture other participants' questions; and 3) meetings where participants themselves were involved in some follow-up action.

Overall, participants said that they found advantages in recording their meetings with ChittyChatty, particularly in cases when there was ambiguity in what was discussed and decisions made in meetings. One participant said: *"I recorded weekly meetings and talks at conferences, which helped me remember what topics were discussed and what was decided. I also tried to record small bits of dialogue during a lunch with friends"*. Those were particularly common in informal discussions (e.g. brain storming sessions) where most people generally did not take any notes or other forms of recordings.

When retrieving information from ChittyChatty, people often browsed though their temporally ordered note list. Some participants made design suggestions such as being able to see multiple notes at once. All participants expressed a strong interest in using ChittyChatty for a longer time and two participants are continuing to use this tool in their everyday work practices.

Overall, long-term study shows that ChittyChatty can help capture and retrieve information in a work environment. It also shows how people adapted their usage over time. Positive user feedback indicates that the tool was successful as everyday memory prosthesis.

4.16 Conclusions

This study showed complex interrelations between OM and PM that determine PM usage. For example when users are confident they can remember unaided, they are less likely to use PM. However PM use also depends on device properties, with usage being affected by both accuracy and efficiency. Devices such as CC that support both accuracy and efficiency are used more than those that are accurate but inefficient (DP), or efficient but inaccurate (PP). Subjective user ratings exactly reflect this usage data, showing a strong preference for CC over other PMs.

These findings relate to existing research into devices to support effective memory for conversations. Consistent with other studies (Abowd, Harvel et al. 2000), the success of CC provides further evidence for the utility of note-taking and personal visual indexing as an effective way to provide controlled access to speech records. Users were highly positive about CC, finding it highly usable and enjoyable. Indeed several have begun to use it in their everyday work environment.

My findings also allowed me to make sense of the controversy surrounding the utility of PM in the form of pen and paper notes. Some researchers have argued (counter intuitively) – that notes are of little use as memory aids (Kidd 1994; Whittaker and Hirschberg 2003), while others suggest they are useful (Landay and Kaugmann 1993). The study results suggest a resolution to this dispute. It was found that notes may be useful in the short-term, but have little long-term utility, becoming no better than OM after a month.

Turning to design implications, these findings indicate a need for PM designers to focus more on *device efficiency*. Some users chose OM over DP even when they were not confident they could remember unaided - suggesting that people will rely on OM if PM access is inefficient. Similarly, users gave higher ratings for PP than DP - because of PP's efficiency, and despite the fact PP was not an accurate record. Although the best designs need to optimise both efficiency and accuracy, if this is not possible, then it may be interesting to explore PM devices that favour efficiency over accuracy, i.e. those generating fast but approximate results. For example, users trying to find relevant documents from an archive may be more responsive to quick fuzzy matches than inefficient accurate searches. Finally the extensive use of CC and the strong positive reaction to its exploitation of handwritten notes suggest that PM devices need to exploit familiar metaphors and work practices in their interface.

The results also have interesting ramifications for psychological memory theories. These have tended to focus on OM to the exclusion of PM. However, my results show that retrieval is often a combination of OM and PM. For example, the crucial role of OM control processes was observed, such as metamemory - whereby user confidence informs decisions about whether or not to use PM. This study shows that such metamemory processes are generally accurate leading to appropriate decisions, e.g. when people decide to use OM they tend to recall accurately. Finally it was observed how PM and OM interact when a PM device is used. E.g. PP notes are useful memory aids at short Testing Sections. However, once OM begins to decay after a month, the notes themselves are no longer useful – indicating that a combination of notes and OM is needed for successful recall. However, long-term trials suggest that people develop faster and more sketchy ways of capturing important pieces of information in conversations.

In conclusion, this study identifies a number of parameters that will inform the design of future PM devices, as I move from proof of concept systems to determining when, why and how such devices can help us remember.

4.17 Chapter Summary

This chapter attempted to investigate the role of a prosthetic memory tool that could aid OM in remembering important information in a conversation. The study demonstrates a synergy between OM and PM, as people only use PM when they think their OM is poor. Also, users generally have a good perception of the accuracy of their own memory. They were also influenced by the properties of the PM device. They preferred a fuzzy, but efficient search rather than laborious but accurate retrieval.

Long-term user behaviour revealed a change in note-taking strategies that people generally want to save time in capturing while still being able to access important information from conversations.

Chapter 5

Digital Notes: How and Why They Cue Memory

5.1 Introduction

This chapter aims to examine the efficacy of memory strategies in the context of digital and paper-based note-taking. The evaluation study and long-term trials with digital and analogue note-taking techniques in Chapter 4 have shown that people generally take many notes to help them remember everyday information. However the benefits of taking notes are not clear. For example prior research has claimed that (a) notes may not always be useful in promoting later retrieval (Kiewra 1985); (b) taking notes may distract people from effectively processing important information (Piolat, Olive et al. 2005).

We are all aware of the fallibility of our unaided organic memories (OMs). In our everyday lives, we often prepare for future memory using a variety of prosthetic memory devices (PMs). We carry PDAs, notepads, diaries and other writing devices to help us remember information that we may need to recall in the future (Landay and Kaugmann 1993; Bellotti, Dalal et al. 2004). We leave emails in our inboxes or sticky notes and paper files on our desktops when there are outstanding actions associated with these (Malone 1983; Whittaker and Sidner 1996; Whittaker, Swanson et al. 1997; Bellotti, Ducheneaut et al. 2003; Kalnikaite and Whittaker 2007). In the longer term, we create mementos (Petrelli, Whittaker et al. 2008) or take photos to trigger memories of events, people and places

(Frohlich, Kuchinsky et al. 2002; Rodden and Wood 2003; Kirk, Sellen et al. 2006; Sellen, Fogg et al. 2007). Recently there has been much interest in replacing this heterogeneous set of memory devices with so called 'Lifelogging' technologies, as technical developments in capture, storage, and information retrieval now make it possible to record every event we experience (Lamming and Flynn 1994; Dumais, Cutrell et al. 2003; Dickie, Vertegaal et al. 2004; Mann, Fung et al. 2005; Gemmell, Bell et al. 2006; Sellen, Fogg et al. 2007).

However, the effective use of any PM device requires *strategic planning*, as well as *systematic knowledge* about how our memories work. This knowledge is referred to in the psychology literature as *Metamemory* (Lachman and Leff 1989), as described in Chapter 2. Metamemory involves knowing for example *what* we are likely to forget and hence what we need to store prosthetically. There is little point in using complex technology to store information that is easily brought to mind, e.g. familiar names, routes or phone numbers, see Chapter 4. Metamemory also involves knowing *how* to store information; we need to organise it in a way that ensures we can access it when it is needed.

There is, however, evidence that people may not be very effective at metamemory strategies associated with stored digital information. First people may store information *that they don't need later.* People prepare for later re-access to web pages by creating complex bookmark hierarchies, but such structures are often not used. For example, 42% of bookmarks are never re-accessed (Tauscher and Greenberg 1997; Abrams, Baecker et al. 1998): by far the most common strategy for re-finding a webpage is to type its URL (Jones, Bruce et al. 2001). Second, *they may spend time organising information in ways that don't benefit later retrieval.* Studies of email show that some people engage in quite complex filing strategies that are counterproductive – creating folders that contain only one or two messages that make future filing less effective (Whittaker and Sidner 1996).

These results are important because they show that efforts to prepare for retrieval may often be suboptimal.

This chapter examines common strategic memory behaviours, namely note-taking, as a form of prosthetic memory strategy. I investigate whether notes *do indeed facilitate* later retrieval of complex conversations. We are all practiced note-takers and it seems intuitively obvious that notes should help memory. One straightforward way that notes can help is to serve as *prosthetic cues*. Looking at a note can directly trigger memory for information such as a name or phone number that is otherwise hard to remember (Landay and Kaugmann 1993; Whittaker and Sidner 1996). Note-taking may also facilitate *organic memory directly*. The act of concentrating on key information to compose notes may help OM by causing people to *focus* more on incoming information - even if those notes are never consulted (Benton, Kiewra et al. 1993; Kidd 1994; Piolat, Olive et al. 2005).

But the potential benefits of notes are not necessarily guaranteed, and recent research has challenged some of the claims about the utility of notes. As with web bookmarks, we may select the *wrong prosthetic cues* - expending effort noting down information that later turns out to be irrelevant. Various studies have shown that notes do not always successfully cue PM recall especially at longer Testing Sections, see Chapter 4. And notes may turn out to be ineffective if users can no longer remember what they mean (Whittaker, Hyland et al. 1994).

Worse still, taking notes may also *detract* from *organic memory* processes by compromising how we process incoming information. With the demands of complex meetings, we may be so busy trying to record a previous critical point that we miss new important information. This new information may end up being only partly processed and soon forgotten. Thus preparing for future PM retrieval may have *attention costs* that are detrimental to OM.

Finally this chapter investigates what motivates note-taking strategies. Are dedicated note-takers people who are aware of the limitations of their organic memory who are interested in offloading retrieval onto PM devices?

I investigated whether, how and why notes might help memory by comparing retrieval for spoken conversations using 2 PM devices, (a) Pen and Paper (PP), and (b) a new generation *digital* note-taking device (ChittyChatty – CC as described in Chapter 4) where digital notes are indexed to a recording of the meeting. I examine how these different PMs are used for retrieving information over different periods of time, and how this compares with unaided OM usage. Further, I am interested in what motivates note-taking. Are people who are not confident about their memory more likely to take notes?

5.2 Research Questions in Note-Taking

This chapter also examines different note-taking strategies and how these affect memory. I investigated: (a) *Quality* of PM notes; (b) *Quantity* of PM notes; and looked at how these affect (c) *Memory Performance* (whether using OM, PM or a combination of both) and (d) *Efficiency* of retrieval, i.e. the time taken to remember whether this is with OM or PM. I also conducted an analysis of different note-taking styles, comparing digital and analogue notes along various dimensions.

More specifically this chapter investigates the following research questions:

- *Do notes help overall memory* – regardless of whether people use their notes to answer a question or choose to rely on what they can remember? If notes do indeed help memory, *how* do they help? Here I distinguish between benefits of notes on PM and OM.

- *PM cueing*: Does note-taking help PM accuracy by generating cues that people use directly to prompt retrieval? Or are such cues ineffective because people expend effort noting information that turns out not to be useful for retrieval?

- *OM effects*: Here I focus on cases where people take notes, but choose not to use them at retrieval – usually because they believe that they can retrieve information correctly unaided. Does careful note-taking *promote OM* accuracy – by causing people to focus more carefully on what was said and hence remember better using OM? Or do notes *detract* from organic remembering? Are attempts to take exhaustive notes counterproductive because users miss much of what was said making them unable to remember little unaided – reducing OM *accuracy*? These competing hypotheses can be referred to as *OM focusing* and *OM distraction.*

In addition, I was also interested in the *types* of notes taken and how this affects recall

- *Effects of Note-type on recall Accuracy:* Is it better to take larger numbers of notes (i.e. a large *Quantity* of notes) in order to generate more complete PM cues, at the risk of noting irrelevant information? Or should people employ more concise note-taking strategies that try to focus on more critical information (i.e. high *Quality* notes)?

- *Effects of Note-type on recall Efficiency*: Is it more time-consuming to retrieve information from highly detailed notes? Or do these more exhaustive cues make retrieval more efficient?

I also wanted to explore people's *reasons* for note-taking. How do notes relate to people's *evaluations* of their memory capabilities?

- *Reasons for note-taking:* Are people who are very *confident* about their organic memory less likely to take notes?

Finally this chapter investigates the *general differences* between digital and paper based notes. Do people tend to take the same types of notes in both cases, or are there significant differences between them?

5.3 Experimental Method

I investigated memory using 3 different types of prosthesis: (a) an analogue note-taking system - Pen and Paper (PP), (b) a digital note-taking system called ChittyChatty (CC) which is described in Chapter 4, and (c) no prosthesis (NP) where people rely on unaided memory. The goal of this study was to test how these different note-taking devices helped users remember everyday conversations as in Chapter 4. I read a series of conversational stories aloud to users, asking them later to retrieve information about those stories. In the CC and PP conditions, users had a device to help them remember prosthetically by looking at their notes, but they could also choose to rely on unaided OM. In the final NP condition they were not provided with any device support and instead had to rely on OM alone. Unlike in Chapter 4 where I reported on actual OM usage, this chapter reports on OM condition to help better understand note taking behaviours using the additional data collected during the experiment reported in Chapter 4.

Twenty five users took part (14 women and 11 men, aged 23-55). Users were volunteers consisting of university researchers, administrative and management staff, as well as other professionals from public and private sectors. Users had no prior knowledge of the project or our experimental hypotheses. None of the users had prior experience of using CC, but obviously all had extensive experience with both OM and PP.

5.4 Stories and Test Questions

I used the same set of stories as detailed in Chapter 4 (Appendix A). The 3 stories were intended to simulate real-life conversations between two old friends who had just bumped into one another after a period of several years. The stories contained a mixture of true facts and made up fiction equally distributed within each story. I conducted extensive pilots with the stories, to ensure they could be easily understood, and they did not contain any unfamiliar or unusual terms. User comments indicated that they were enjoyable to listen to, as well as achieving their objective of simulating real-life conversational experiences. The average story time was 3.20 minutes (see Appendix A).

An example fragment of one story was the following (for more stories, see Appendix A):

"Oh, do you remember my older brother, Dave? Let me tell you how he got here. He has loved Def Leppard ever since he was 15 years old and saw them play at the Sheffield Show, Hillsborough Park in 1978. The hair, the tight trousers, the heavy guitars, the thunder of the drums and the screaming vocals. He was particularly entranced with their Yorkshire lyrics. To be honest, he was obsessed. They used to rehearse in some old warehouses and he would hang around outside listening to them tune their guitars. He found it entertaining. When their practice sessions were over, they'd catch their bus home and Dave would pretend he was getting the same bus...In 1979 Def Leppard were one of the biggest rock bands in the country, but then a strange thing happened. A journalist for Sounds magazine wrote that the band had "sold out" to America. Dave wasn't sure what that meant. Like just about every other band, they wanted to be successful in America, but so what? It's not like they had cut their hair, but suits on and started singing mushy ballads. Most of their original fans believed this story and when they played a Reading festival in England, they showered them with bottles. It was another 7 years before their home country would ever really accept them again... Anyway success followed again at the end of eighties, followed by the inevitable decline. The albums began to lose their edge and when Steve, the guitarist died, Dave thought they would pack it all in. But they kept going, keeping the tour bus rolling, last night they came home to Sheffield to play the Arena, and as usual Dave was right at the front going crazy. Suddenly, Joe, the singer spotted Dave in the crowd. He'd recognised Dave after all those years, thought obviously he was a bit fatter and his long hair was thinning a lot. To Dave's complete surprise Joe pulled him out of the crowd and introduced him to the whole arena as Def Leppard biggest fan. They nicknamed him "Mad Dave". Dave raised his arms into the air to bash in his glory and then dived forward back into the crowd. Obviously they didn't fancy a fat, balding and middle-aged rocker

landing on their heads. So that's how he got here, Northern General Hospital with crushed ribs, a fractured arm and a broken nose"

After hearing the story I asked users different recall questions. The above story generated 4 questions: (1) "*Which year did Def Leppard became one of the biggest rock bands in the country?*" (2) "*What was Dave's nickname?*" (3) "*How did the local fans feel about Def Leppard's success in America?*" and *(4)* "*How did the crowd respond to Dave diving onto them*".

5.5 Prostheses

ChittyChatty (CC) - Digital Notes

CC is similar to other note-taking systems such as (Whittaker, Hyland et al. 1994; Abowd, Harvel et al. 2000; Stifelman, Arons et al. 2001) and described in detail in Chapter 4.

Pen and Paper (PP) - Analogue Notes

People were given pen and paper to take notes and they were instructed to take notes as they would normally do to remember complex verbal materials.

Organic Memory (OM)

The final condition was OM. In this condition people were not given external memory aids and had to rely on their unaided memory to find the answers to the questions.

Differences in Efficiency and Accuracy between PM Devices and Unaided Memory

All of those experimental prosthetic devices have different properties. PP notes are a schematic and incomplete record of what was said, whereas CC offers a verbatim record. Retrieval Efficiency is also different for these devices. Extracting information from PP or written CC notes is efficient because the eye can rapidly scan text to identify information. CC, however, should support reasonably efficient access to the underlying speech record; using handwritten notes or other visual cues should allow users to quickly identify relevant regions of speech to access and listen to. Of course, both prosthetic devices contrast with unaided memory which is efficient to access but fallible.

5.6 Time: Testing Sections

As in Chapter 4 study, the entire experiment consisted of 3 Testing Sections. The first Testing Section consisted of an introduction, CC training, exposure to the conversational stories and initial memory testing. This took about 50 min. The second Testing Section – a week later - involved remembering certain aspects of the stories presented at the first Testing Section and lasted about 30 min. The last Testing Section took place a month after the first and again involved retrieval of information presented at the first Testing Section and lasted about 30 min. At the end of the experiment users were given a small reward for participating.

Before listening to each story I either gave users a device (CC or PP), or they had to rely on unaided memory (OM). In the CC and PP conditions, they obviously also had OM, which they could choose to use instead of the allocated prosthetic device. Users were instructed to remember the story either with the assistance of the PM or OM, using whatever memory techniques they would normally use. Users heard each story only once - during the first Testing Section. I tested memory at 3 different Testing Sections: same day, 7 days and 30 days later. On each test the users had the same prosthesis as when they heard the original story. For instance if a user had access to CC when first listening to the story, I also gave them CC at subsequent Testing Sections with the same story.

A critical research question was whether and when people made use of devices instead of relying on unaided OM. So, even when users had access to a note-taking device, I made it clear that they were not compelled to use it, and I noted when devices were used in preference to OM.

5.7 Procedure

The experiments were run using a custom built website (see Appendix A). Users were first given a general description of the experiment, the stories and the different types of questions that they would be asked as part of each session.

I then gave them a brief web-based, hands-on tutorial providing detailed descriptions of each memory prosthesis and procedures for the experiment. They carried out 3 practice tasks (one with each prosthesis, and one with no prosthesis). The practice tasks were similar to those used in the experiment. Users were allowed to proceed to the actual experiment only if (a) they felt confident with each device and (b) they had successfully completed all practice tasks.

5.7.1 Experimental Tasks

I read users a story with CC, PP or OM depending on the experimental condition. To control for story/Retrieval Method confounds, the order in which users received stories was counterbalanced, the device they used to carry out each task, and the type of question (verbatim/gist) they were asked. Users answered questions on web based forms (see Appendix A).

A key research question was the relationship between users' confidence in their organic memory and their note-taking and retrieval strategies. Before answering each memory question I asked users to evaluate their *Confidence* in their ability to answer the question without using the device. The Confidence question was asked *after* the user had read the memory question but *before* they answered it: "*How confident are you that you can remember the answer to this question without using your [memory prosthesis name]?*" Responses were generated on 5-point Likert scales.

Users then tried to answer the question. In all conditions I recorded the retrieval time, i.e. how long it took users to answer that question. I also noted whether users relied on OM or the device to answer the question.

Retrieval accuracy was scored as in Chapter 4. I first generated an evaluation metric for each question, by having two coders blind to the experimental hypotheses listen to each story twice. They agreed a set of target answers, specifying keywords and context that needed to be present in that answer. Accuracy scores ranged from 0-5 depending on how much of the target answer the user specified. If an answer included all target keywords (or their synonyms) and context, it received a maximum score of 5. Partially correct answers were defined as either (a) containing all keywords, but inaccurate context, or (b) accurate context and incomplete set of keywords. Scoring was carried out independently by the two judges, and disagreements were referred to a third judge for resolution.

5.7.2 Measures and Variables

The following data was collected and reported:

- *Accuracy* of answers
- *Efficiency* – how *long* it takes users to recall the relevant information
- *Retrieval Method:* when people had notes available with CC and PP, I looked at whether they relied on their notes or their organic memory to answer a question.
- *Note taking behaviour* - *how* people take notes, Note *Quality* and *Quantity* for CC and PP
- User *Confidence* in their ability to remember unaided.

5.8 Hypotheses and Results

Before proceeding with parametric statistical analysis, I applied Kolmogorov-Smirnov (Sheskin 2007) to test for normality in data collected for this chapter. The results of this test have confirmed that data collected is normally distributed.

The results are organised around the following hypotheses:

5.8.1 Overall Memory Benefits: Do Notes help cue recall?

I compared the OM condition with the two note-taking conditions (i.e. comparing CC, PP and OM scores) to assess overall effects of note-taking devices on Retrieval Accuracy.

A two-way ANOVA was conducted with independent variables 1) *Device*– i.e. which prosthesis was used (PP, CC or none in the OM condition); 2) *Testing Section* – i.e. length of time since the user heard the story (same day, 7 days later, 30 days later). The dependent variable was *Accuracy*.

Accuracy scores for digital and analogue note-taking devices and OM are shown in Figure 5.1. As expected there was a significant difference between Devices ($F_{(2,898)}=78.9$, $p<0.0001$ (SD=483.4, mean=241.7)). Planned comparisons showed that there are differences between CC and OM ($p<0.0001$) and between PP and OM ($p<0.0001$) showing the benefits for memory of both types of note-taking prosthesis.

As expected there was a strong interaction between Device and Testing Section ($F_{(4,898)}=11.6$, $p<0.0001$ (SD=142.1, mean=35.5)).

I conducted post-hoc Tukey analyses at each Testing Section. I found no significant differences between CC, PP and OM on the first day, suggesting that there were no immediate benefits for using notes. But 7 days later, I found a significant difference between both CC and OM ($p<0.0001$) and between PP and OM ($p<0.0001$) – indicating strong short-term benefits of both types of note-taking prostheses. During the final session - 30 days later, I found significant differences between CC and OM ($p<0.0001$) and between CC and PP ($p<0.0001$) – indicating the benefits of digital notes over analogue notes at longer Testing Sections, presumably because CC allows access to the underlying verbatim record. PP and OM were equivalent ($p>0.05$) showing that the benefits of analogue notes degrades over time.

I conducted further post-hoc tests looking at Accuracy for each Device at the different Testing Sections. I found no significant difference for CC for all three of the Testing Sections ($p > 0.05$). But for PP there was a significant difference between 1 and 7 days ($p<0.02$), and 7 and 30 days ($p<0.006$). For OM there were differences between 1 and 7 days ($p<0.0001$) sessions but no differences between 7 and 30 days ($p>0.05$) sessions. The fact that CC shows no degradation over time, whereas both PP and OM decay shows the benefit of digital note-taking in protecting recall.

Given that notes do indeed help *overall* retrieval, our next question was: how? What are their effects on PM and OM respectively? I first looked at whether note-taking helps *PM* accuracy by generating rich cues that are useful for later PM retrieval.

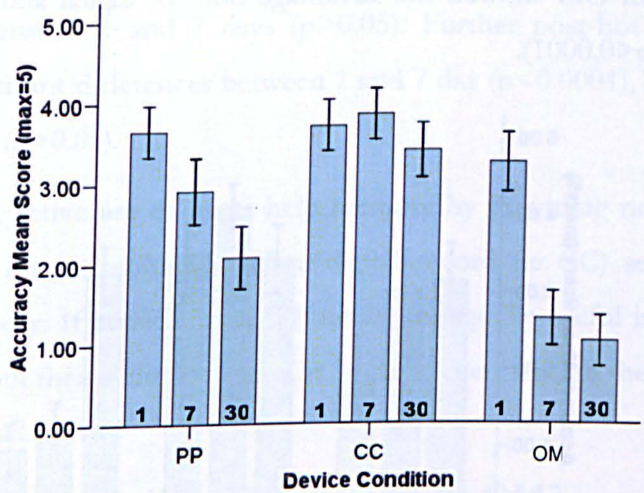


Figure 5.1: Overall Accuracy for different Device Conditions over 3 Testing Sections. For each Device the intervals are 1, 7 and 30 days from left to right.

5.8.2 Do Notes Help Cue PM?

For both CC and PP, the above analysis does not separate cases where notes are available but not used, from cases where notes are taken and used. It may be that users take notes but choose not to use them for retrieval. To examine this, I compared Accuracy when people actively used CC or PP notes at retrieval with the OM condition when no notes were available. Because I wanted to quantify the direct effect of *notes as retrieval aids*, I excluded from the analysis, cases where people had digital notes, but chose to rely on their unaided memory, as in these cases there was no direct prosthetic use of notes. I conducted an ANOVA with *Testing Section* (same day, 7 days, 30 days) and *Device* (CC, PP, and OM) as independent variables. *Accuracy* was the dependent variable. Figure 5.2 illustrates our findings.

As expected, there were significant differences between Devices for Accuracy ($F_{(2, 683)} = 65.7$, $p<0.0001$ (SD=398.4, mean=199.2)). Planned comparisons of CC and PP, with OM showed the advantage both of digital and paper notes over OM (both $p<0.0001$).

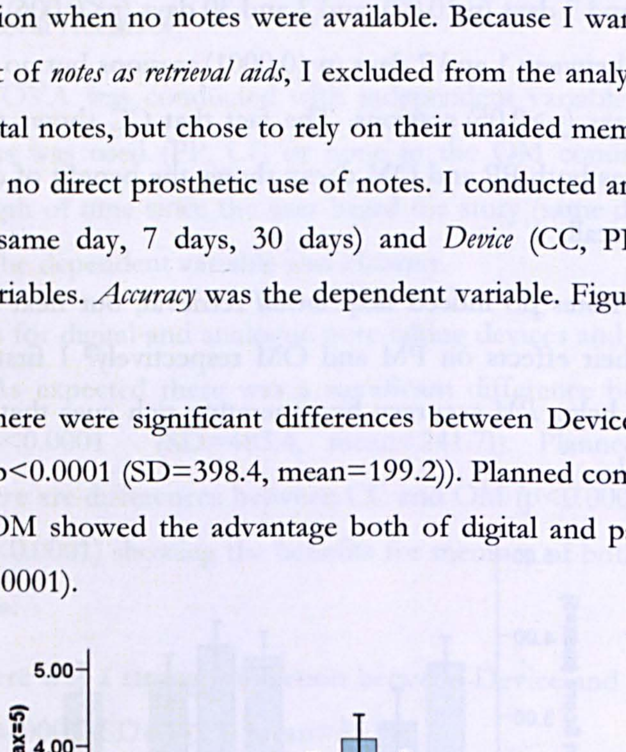


Figure 5.2: Overall Accuracy for the different devices when notes are actively used. For each device the intervals are 1, 7 and 30 days from left to right.

There was also an ANOVA interaction between Device and Testing Section for Accuracy ($F_{(4,683)}=14.0$, $p<0.0001$ (SD=170.3, mean=42.6)). Post-hoc Tukey tests showed that when CC and PP notes were used, performance was better at longer Testing Sections. When I compared CC, PP and OM at day 1, there were no significant differences (all $p>0.05$). At 7 days, there was a significant difference between CC and OM ($p<0.0001$) and between PP and OM ($p<0.0001$). Similarly at 30 days there was a significant difference between both CC and OM ($p<0.0001$) and between PP and OM ($p<0.007$). The data show the benefits of actively using notes as PM cues: PP notes or, with CC, the combination of notes and access to the verbatim record, help by cueing PM at longer Testing Sections, as OM degrades. Further post-hoc tests showed a significant benefit of using CC over PP notes ($p<0.001$) to access a verbatim recording of a conversation.

I conducted further post-hoc tests examining CC at 1, 7 and 30 days and found no significant differences ($p>0.05$). But when I looked at PP, I found a significant difference between 7 and 30 days ($p<0.004$), although there were no differences between 1 and 7 days ($p>0.05$). Further post-hot tests with OM revealed significant differences between 1 and 7 day ($p<0.0001$), but not between 7 and 30 days ($p>0.05$).

In conclusion, active use of notes help memory by providing rich cues for PM. In particular, notes combined with a digital record (ie CC) seemed relatively immune to decay. If actively used, PP notes are equally useful immediately and after a week, but their utility degrades at 30 days, even though they are still better than OM overall.

5.8.3 OM: Do notes distract or help focus OM?

There are competing views about how notes affect organic memory. To investigate whether taking notes *distracts,* or *focuses* OM, I examined the differences in Accuracy when people had taken notes *but chose not to use them* preferring to rely on their unaided memory to remember. I compared unaided

OM in the CC and PP conditions (when notes have been taken but not used), with OM scores where no such notes are available. I conducted an ANOVA with *Testing Section* (same day, 7 days, 30 days) and *Device* (CC, PP, and OM) as independent variables. *Accuracy* was the dependent variable. Figure 5.3 illustrates those findings.

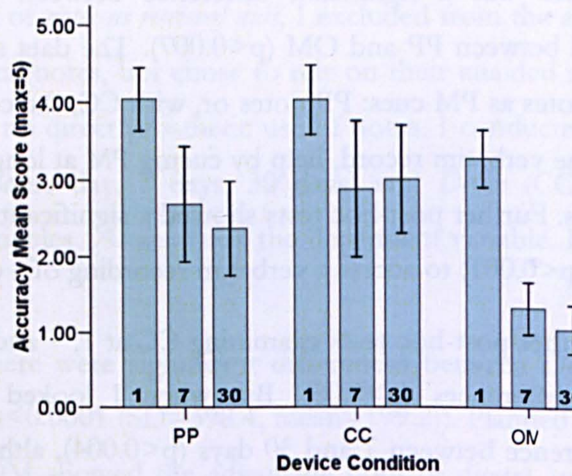


Figure 5.3: OM focus at different Device Conditions over 3 Testing Sections. For each device the intervals are 1, 7 and 30 days from left to right.

As expected, there was a significant difference between Devices for Accuracy ($F_{(2, 514)}$ = 32.1, p<0.0001 (SD=191.1, mean=95.6)) with planned Tukey comparisons showing that in cases where people had taken either CC or PP notes they outperformed OM for unaided memory (p<0.0001).

However, there was no interaction between Device type and Testing Section ($F_{(4, 514)}$ = 1.7, p>0.05 (SD=20.1, mean=5)). As Figure 5.3 indicates the differences between devices are equivalent at each Testing Section. I conducted further post-hoc tests looking at all PM devices at (1, 7 and 30 days). In all cases OM dropped between 1 and 7 days with no subsequent differences between 7 and 30 days.

Overall the results support the focusing hypothesis and contradict the distraction hypothesis. Even when people choose not to use their notes at recall, the initial act of taking notes helps to boost their memory compared with when no notes are taken.

5.8.4 Effects of Note-type: Quality vs. Quantity

Next I looked at the effects of different types of notes on retrieval. Is it better to take a larger numbers of notes (i.e. high *Quantity*) in order to generate exhaustive PM cues? Or should people employ more concise note-taking strategies that try to focus on more critical information (i.e. high *Quality*)?

The *Quantity* of notes was scored by simply counting the number of words that users recorded. *Quality* was more complex. For each story, I devised a marking scheme consisting of the ideal set of notes that would have to be generated to cover all the topics that I asked users about. This included topic keywords plus contextual information about each topic, and both were required to achieve a perfect notes score, see Figure 5.4. Two independent judges applied the marking scheme; they gave 5 marks for complete notes which captured both keywords (or their synonyms), plus context. Marks were reduced for incomplete and partial information. Consistency between judges was 90% and disagreements were resolved by discussion.

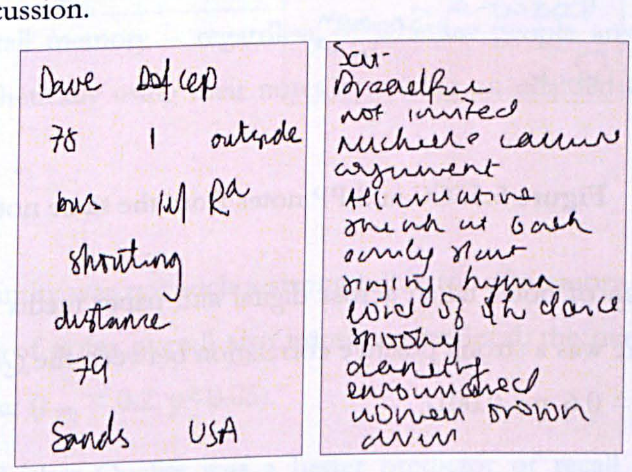


Figure 5.4: CC low and high quantity notes from two different note takers

5.8.5 Similarities between CC and PP notes

Before examining the effects of note type on memory I looked first at how people took notes digitally compared with pen and paper. For instance, when people used CC, did they take more or fewer notes, in comparison to when they used PP?

I found that people had consistent note-taking strategies - taking similar *Quality* notes digitally and on paper. The strength of this relationship is confirmed by a correlation between *Quality* of digital CC notes and *Quality* of PP notes ($r_{(25)} = 0.4$, $p< 0.001$). Figure 5.5 shows CC and PP notes from the same note taker – suggesting similar note-types and strategies in both cases.

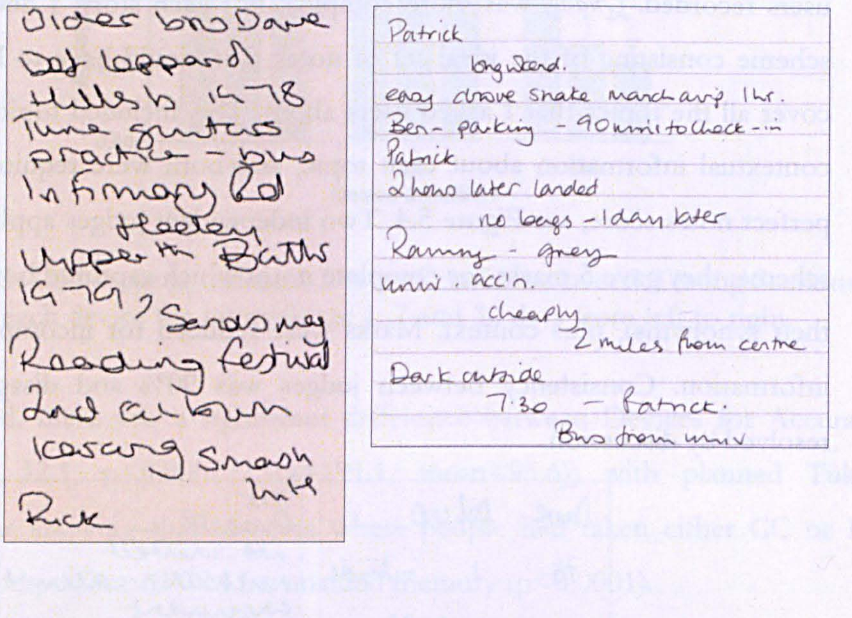


Figure 5.5: CC and PP notes from the same note taker.

The *Quantity* of notes taken across digital and paper media was also consistent. Again, there was a strong positive correlation between the *Quantity* of CC and PP notes ($r_{(25)} = 0.6$, $p< 0.001$).

I was also interested in how people used the space on the Digital PDA screen compared with the paper sheet. I expected people to spread their notes out more on paper but to have more condensed digital notes. I measured space usage in terms of the number of lines used, expressed as a percentage of the number of available lines.

Contrary to my expectations, I found that people used space in a similar way for digital and paper notes. There was a strong correlation between CC and PP space usage ($r_{(25)} = 0.35$, $p<0.01$) suggesting that people tend to apply familiar note-taking strategies to new technology.

I also observed that people used often bullet points to arrange their notes in both CC and PP. Again I found a strong correlation between bullet point use in CC and PP ($r_{(25)} = 0.25$, $p<0.01$).

Having established that digital and pen and paper notes are similar I combined digital and analogue notes in testing the relations between note type and memory.

Quality:

I looked at whether people who take high *Quality* notes in both CC and PP also remember better overall. I found a significant correlation between *Quality* of notes and overall memory – regardless of whether people answered memory questions prosthetically using their notes or relying on unaided OM ($r_{(25)} = 0.5$, $p<0.01$).

Quantity:

In contrast Quantity was not such a strong predictor of memory. While taking a higher Quantity of notes overall also improved retrieval, the overall correlation was much weaker ($r_{(25)} = 0.2$, $p<0.05$).

I also tested whether Quality was a better predictor of recall Accuracy than Quantity and found that this was indeed the case. A comparison of the correlation coefficients using the Hotelling/Williams Test showed that Quality was much the stronger predictor ($t_{(515)} = 4.63$, $p < 0.0001$).

Retrieval Efficiency:

I also looked at the effects of note *Quantity* on *Retrieval Efficiency.* Does taking more notes increase speed of retrieval or are too many notes distracting as retrieval cues?

There was a strong positive correlation between overall note *Quantity* and Efficiency (time to answer each question), ($r_{(25)} = 0.4$, $p<0.01$). People who took more notes tended to take *longer* to generate responses using PM. This suggests that having a large volume of notes decreases speed of retrieval - as there are more notes to scan to find a promising index.

Reasons for Note-Taking:

Finally I looked at what motivates people to take more notes. Do people who are *less confident* take larger volumes of notes, or does confidence result from having good notes?

Contrary to my hypothesis, there was no significant correlation between combined CC and PP note *Quantity* and overall *Confidence* scores ($r_{(25)} = -0.003$, $p>0.9$). People who are not confident about their memory don't act upon this information to take more notes.

However, there was a strong positive correlation between overall note *Quality* and *Confidence* scores ($r_{(25)} = 0.2$, $p<0.01$). This may be because having higher *Quality* notes seems to boost people's confidence that they will be able to remember unaided. Or alternatively people who have better memories tend to take better notes and they are more confident about OM based on their past success of remembering.

5.9 Subjective User Comments

Overall people voiced a liking for memory cueing techniques such as note taking. The majority of users preferred digital over paper notes but this could be due to the audio back up provided alongside the digital notes "*[CC] is similar to writing [PP] notes which I like plus back up of recording*". Nevertheless CC was appreciated for its similarity to PP and its ease of use "*[CC] is very easy to use and it's very accurate*".

But after some time had elapsed people worried that their notes might not be sufficient to guarantee long term retrieval because they were contextually dependent on fallible organic memory to interpret them "*my notes were a mess, and were contextual with my own memory, which itself had faded. Hence the usefulness of the notes was severely undermined*".

A few people acknowledged the importance of taking high Quality notes "*[PP] jogs your own memory. It depends largely on the quality of notes*"...and "*you have notes for prompts if you make the right prompts!*"

Some users discussed the *Quantity* of their notes and the need not to take too many notes "*[after 7 days] the notes I had done were the right amount - not too lengthy*". As I reported earlier, having a large *Quantity* of notes was not perceived as an effective memory cue as it takes too long to retrieve relevant information.

There was also concern about note taking technique "*I've never been very good at making notes*" and "*...the notes are only as good as the user...*" But even with poor note taking skills, people realised that having some notes - whatever their quality - might be better than having no PM backup at all "*I'm dependent on my own ability to make notes but still better than memory alone*".

5.10 Conclusions

This study examined whether users' attempts to prepare for future retrieval using a specific type of prosthetic device, namely notes, led to improved recall.

What can our results tell us about digital memory more generally? They have shown that there are two independent mechanisms by which users' preparations for future retrieval can have effects on memory:

(1) PM cueing – by generating useful cues (notes) that trigger memories when users access them at retrieval. In this study, digital notes in CC were highly robust as retrieval cues showing minimal decay over the month of the study

(2) OM focusing – the very act of generating cues helps memory (even if these cues are never consulted). However, these focusing benefits decay over time.

Also I found some effects for the *types* of cues that people construct. Higher quality cues helped retrieval whereas large volumes of notes only weakly did so. Furthermore there were *costs* to taking too many notes. Generating too many cues leads to more inefficient retrieval with increased retrieval times.

Finally this Chapter clarified the relationship between note-taking behaviour and confidence. People who lacked confidence in their OM were no more likely to take large numbers of notes than those who were very confident. Rather it seemed that taking good notes caused people to be more confident that they would remember unaided.

There were also few observed differences between digital and pen and paper note-taking practices. Digital and analogue note-takers tend to exploit space in similar ways, to use an equivalent number of bullet points and to take a similar volume and quality of notes.

These results have important implications for other studies of prosthetic memory. There has been much recent interest in techniques that allow digital memories to be automatically indexed (Wenyin, Dumais et al. 2001). While such automatic techniques may prove useful, these results show that having users generate *their own cues* is helpful rather than distracting, even when these aren't directly used at retrieval. I did not directly contrast automatic versus manual indexing in this study. At the very least, however, the results in this Chapter suggest that in addition to such automatic techniques we need lightweight ways for people to construct their own retrieval cues, because of the demonstrated benefits they bring.

The confidence results suggest that having high quality personal indices increases confidence and hence the likelihood that the digital memory will be used. Finally the demonstrated relationship between Quantity and Efficiency showed that we need to be careful about how many cues we generate (whether this is done automatically or manually). Too many cues reduce the efficiency of retrieval.

More specifically the study shows the benefits of a new type of digital note-taking device: CC, for helping memory. It is more robust than both PP notes and unaided memory. In the spirit of Web2.0, in social summaries Chapter 7 I describe a further investigation into extending the device to allow the collaborative sharing of notes so that, for example, a class of students could share digital notes that were indexed to a podcast lecture (Davis, Landay et al. 1999). Further I also evaluate a different version of CC that uses pictures rather than annotations and looking at how different *types* of annotations such as pictures support retrieval, which is described in Chapter 6. This work also extends recent studies of how pictures help individuals remember events from their everyday lives (Sellen, Fogg et al. 2007).

In conclusion I have shown the benefits of new types of digital note-taking prostheses in helping memory and clarified some of the different mechanisms by which they achieve their effects. Future work needs to extend these questions to look at how manual cueing contrasts with automatic methods and how well these techniques generalise to other types of indices such as pictures (Sellen, Fogg et al. 2007) or more complex narratives (Frohlich, Kuchinsky et al. 2002).

5.11 Chapter Summary

This Chapter presents the efficacy of memory strategies in the context of digital and paper-based note-taking. It examined pen and paper note-taking as well as a new generation digital note-taking device: ChittyChatty, finding that notes help memory in two ways. First they provide cues that help people retrieve information that they might otherwise forget. Second the act of taking notes helps people to better focus on incoming information *even if they never later consult these notes.* Finally this chapter reports differences between different note-taking strategies. People who take high quality notes remember better than those who focus on exhaustive documentation; taking large volumes of notes decreases the efficiency of retrieval – possibly because it is more time consuming to scan extensive notes to find relevant retrieval cues.

To build on this work, we need to explore the application of PM tools in real world settings. This will provide a clearer understanding of how such memory devices are used when users are presented with complex real-life memory demands. In the next chapter, I describe an extended implementation of the existing PM tool. I examine how the new system can facilitate remembering of important information in a large classroom setting where users have to remember complex new information. I also look at what system features and participant behaviours may influence performance and attention.

Chapter 6

Augmented Digital Records to Support Organic Memory for Learning

6.1 Introduction

The previous two chapters have looked at conversational memory using CC to evaluate various hypotheses about relationship between OM and PM. In this Chapter I extend the analysis of the relations between OM and PM in the context of longer-term memory for complex information in an educational application. I also test a novel variant of CC where pictures instead of text are used to index speech.

Because human memory is fallible, we all rely on various Prosthetic Memory (PM) devices such as diaries, notebooks, sticky notes and calendars to remind us about things that we would otherwise forget. However, recent advances in storage, networking and sensor technology have made it possible to capture huge amounts of digital data relevant to our everyday lives. We can potentially record every experience we have, and every piece of information we touch. One potential benefit of these digital records (DRs) is that they might address the limitations of human memory.

Various 'Lifelogging' visions, have been proposed, starting with Bush's Memex (Bush 1945) and including the influential MyLifeBits (Gemmell, Bell et al. 2006). These visions have led to the development of large numbers of proof-of-concept DR demonstrators that are intended to support our fallible memories (Lamming and Flynn 1994; Dumais, Cutrell et al. 2003; Dickie, Vertegaal et al. 2004; Karger and Quan 2004; Cutrell, Robbins et al. 2006; Gemmell, Bell et al. 2006; Sellen, Fogg et al. 2007). However uptake of DRs has been slow and few working applications have been deployed outside the laboratory. One challenge I address here is to find domains where there are strong memory requirements, where DRs can be practically deployed.

One domain of considerable promise for DRs is education. In many pedagogic situations there is a need to master and reflect on complex information delivered verbally in real-time. Prior research has documented the cognitive problems that students experience in determining what is critical (and hence important to record) while simultaneously processing complex new information (Brown 1987; Bransford, Brown et al. 1999).

DRs might therefore be useful in freeing students from the pressures of 'capturing everything', while trying to simultaneously comprehend novel ideas or contribute to class discussion (Brotherton and Abowd 2004; Munteanu, Baecker et al. 2008). DRs also potentially allow students to be more self-directed, pacing themselves, and allowing more time for personal reflection (Brown 1987). The ability to re-access material may also be of benefit to particular populations, e.g. non-native students who experience additional challenges of trying to master new material delivered in an unfamiliar language (Robertson, Lane et al. 2000). Furthermore, exploration of recording tools is timely, technologies such as MP3 players are now readily available, making it straightforward for students to make repeat listens to podcast recordings at their convenience, and many institutions are now actively experimenting with lecture recording for asynchronous learning

(Hiltz and Goldman 2005; Walker and Moore 2005). Finally multimedia access tools are mature, with well understood techniques developed for controlled access to complex multimedia recordings (Stifelman, Arons et al. 1993; Whittaker, Hyland et al. 1994; Brotherton and Abowd 2004; Munteanu, Baecker et al. 2008).

Despite these arguments, the benefits of DRs may not be clear-cut. Although early studies (Brotherton and Abowd 2004), found that DRs were well liked by students, there was no evidence that they led to measurable pedagogical benefits. There are also potential disadvantages to DR deployment. DRs may change students' learning strategies, making them less likely to attend classes, hence failing to benefit from social learning opportunities (Johnson 1981; Hiltz and Goldman 2005). The increased popularity of these techniques make it crucial to establish whether DRs have pedagogical benefits (Hiltz and Goldman 2005; Walker and Moore 2005) and how they may be used most effectively in educational settings.

I therefore set out to test the benefits of DRs in this setting. I developed two novel UIs that allowed controlled access to DRs, using student-generated handwritten or photo annotations (Figs 1 and 2). I collected naturalistic data about DR use from 98 students over the duration of a course, as well as more controlled data from 35 more students who used DRs to answer class quizzes. I addressed the following research questions:

Overall benefits: Will people make use of DRs in real-world settings? What are the main advantages of DRs? Do DRs help students to perform better on class assignments compared with more traditional instructional techniques such as handouts or personal notes? And do students prefer DRs to traditional tools?

Users: Which students use DRs, and why do they do so? Who benefits most from them? Do non-native speakers exploit DRs to re-access material they might have found initially hard to understand, or are they used as a 'catchup' device by absentee students to listen to material that they initially missed?

Exploitation processes: If DRs do help memory, how do they do so? And when are they used? Do students use them to answer specific course questions or to remind themselves about the gist of an entire lecture? And what types of retrieval indices are most useful in accessing DRs?

6.2 Related Research

Many recent systems support multimedia capture. eClass (Brotherton and Abowd 2004) integrated traditional and multimedia methods to support learning. It supplements the regular learning experience with digital video recordings of lectures, slides, digital whiteboard activity, and personal digital notes. Evaluation surveys showed students felt access to the DR recordings allowed them to participate more effectively in classes. More quantitative learning benefits were not so clear however. eClass users performed no better on assignments than those using regular teaching materials. Other similar systems have also not found huge benefits for annotated lecture recordings (Grudin and Bargeron 2005).

Similar classroom tools that combine DRs with active student sharing of information include Active Class (Ratto, Shapiro et al. 2003), Debbie (Berque, Johnson et al. 2001), Classroom Presenter (Anderson, Anderson et al. 2004; Wilkerson, Griswold et al. 2005), and DyKnow (Berque, Bonebright et al. 2004). A more traditional approach supporting handwritten annotations, NoteNexus (Harvel, Scheibe et al. 2005), showed increased access for materials directly related to assessments. Other research has focused on developing advanced search and browsing to provide remote access to lecture materials (He, Grudin et al. 2000). The Personal Audio Loop (Hayes, Patel et al. 2004) was found to be useful in the social context of recording everyday conversations on a ubiquitous device.

Recent educational theory has emphasized the importance of cognitive monitoring, self-paced learning and self-direction (Brown 1987). A DR of classroom interaction can allow students to engage in more reflection about their understanding, and more active learning. As multimedia records they might also support different learning styles (visual, auditory) (Bransford, Brown et al. 1999), as well as the possibility of seeing material in different contexts - encouraging cognitive flexibility (Spiro, Feltovich et al. 1992).

In addition many institutions are actively experimenting with lecture recording, either using standard recording software such as Camtasia (2008), or dedicated recording tools (Walker and Moore 2005). Tools such as Wimba (2008), iLinc (2008), Elluminate (2008) MRAS (LeeTiernan and Grudin 2001) or SUNY (Shea, Fredericksen et al. 2001) are being actively deployed to support asynchronous learning. Research is now beginning to explore the implications of such deployments (Hiltz and Goldman 2005; Walker and Moore 2005).

6.3 Creating Digital Recordings

Tools

Creating Annotated DRs. Prior research has demonstrated the benefits of time-indexed user annotations in supporting DR access (Stifelman, Arons et al. 1993; Davis, Landay et al. 1999; Kalnikaite and Whittaker 2007). I therefore used 2 access tools to capture end-user annotations: (a) the Sony Recorder (Figure 6.1) which takes photographs to serve as annotations,and (b) ChittyChatty (Figure 6.2) a pen-based UI that creates handwritten annotations. Both tools also record speech, and end-users' annotations are automatically temporally co-indexed to that recorded speech (see Figure 6.3), allowing those annotations to be used for controlled access to the speech.

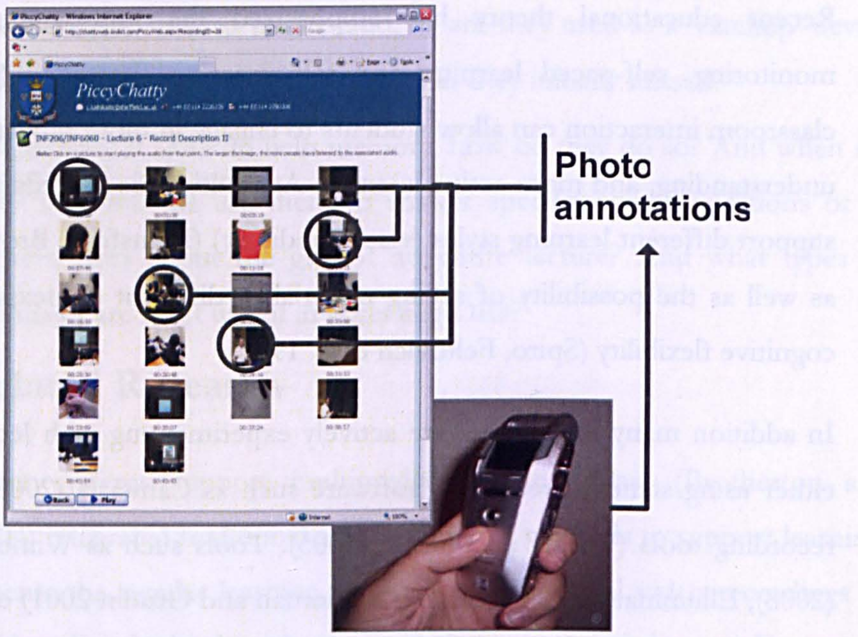


Figure 6.1: Sony Recorder device for capturing speech and end user photo-annotations, along with PiccyWeb UI for retrieving speech using these annotations.

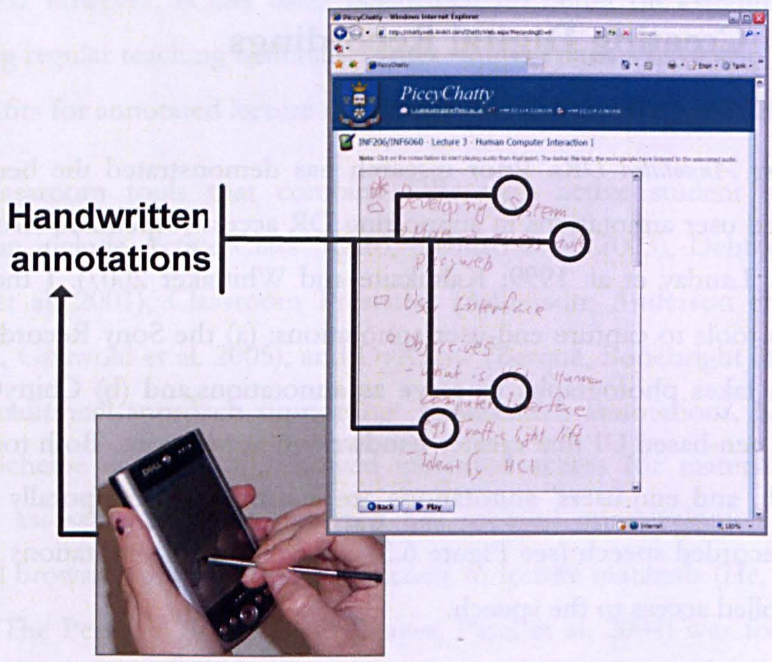


Figure 6.2: ChittyChatty device for capturing speech and handwritten annotations, along with ChattyWeb UI for retrieving speech using annotations.

Accessing the DR using Photo and Handwritten Annotations. I developed two retrieval UIs: one for DR access using photo annotations[2] PiccyWeb (PW, Figure 6.1), and one for DR access using handwritten annotations: ChattyWeb (CW, Figure 6.2). Access is straightforward. Temporal co-indexing of annotations and speech means that clicking on an annotation plays back what was being said at the time that the annotation was made (see Figure 6.3). Notes/photos therefore serve as high-level index to the underlying speech.

For example with PW, if users want to access the part of the DR associated with a specific topic, they scan for a relevant photo annotation (e.g. a photo showing a slide relating to that topic - Figure 6.1). Clicking on that photo means the system begins to play what was being said at the time that the photo was taken. This allows users controlled access to different parts of the verbatim content of the lecture - without having to listen to the entire recording[3]. Of course for the annotator there is a delay between hearing information and constructing the index relating to that information. After extensive piloting, I therefore introduced a short offset, so that playback starts 1.5s before a given index was made.

Context

Our study ran over 13 weeks, as part of an Information Storage and Retrieval course. The course was introductory covering different types of indexing methods, search engines and how they operate. Lectures took place once a week

[2] Note that we use the term '*photo annotation*' to refer to the fact that the photo serves to tag/index the underlying speech. This form of annotation is not to be confused with many digital photography applications that allow users to add textual descriptions that serve to tag a photo.

[3] The practice of using cameras to generate visual reminders is becoming increasingly common, e.g. people photographing important slides in a presentation using phones or digital cameras. Of course unlike our application, these photos do not provide controlled access to what was being said at the time the slide was shown.

and for each two hour lecture, one volunteer used a Sony Recorder (Figure 6.1) to record the lecture - annotating the speech with relevant photos. Another volunteer used ChittyChatty (Figure 6.2) to create a speech record of the same lecture, annotated with their own personal digital handwritten notes. Example annotations are shown in Figs 1 and 2 respectively. While photo annotations are novel, other work has shown that digital handwritten annotations are similar to those taken using pen and paper (Kalnikaite and Whittaker 2008).

In the first lecture, I explained the DR access interfaces to the whole class as well as how to use them. Before each subsequent lecture, all students were reminded about the existence of the DRs and reminded how the PW and CW access systems worked. Volunteer annotators were also given a brief reminder tutorial about how to generate annotations before they began their note/photo-taking. After the lecture, the recording and annotations were uploaded to both CW and PW where they could be accessed by anyone in the class via the internet. Links to the annotated DRs were prominently displayed on the class webpage.

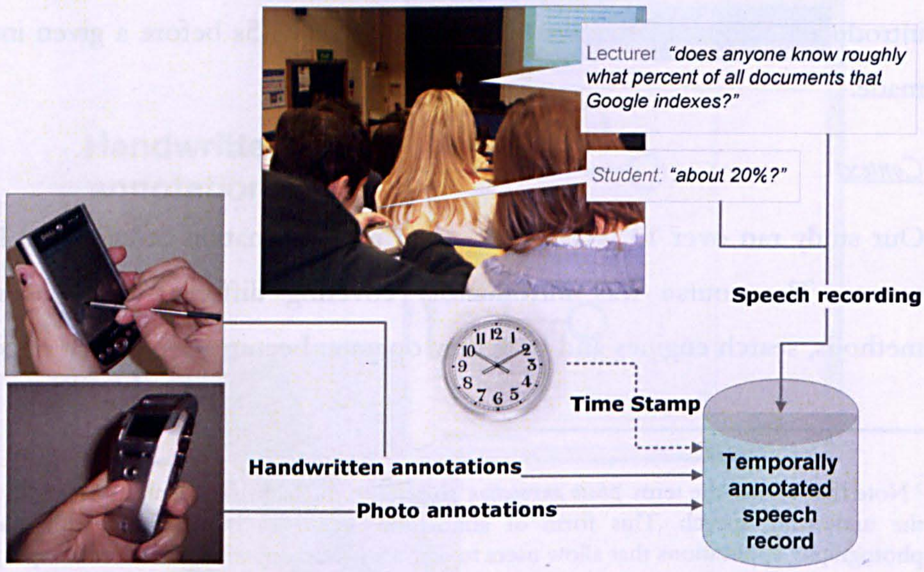


Figure 6.3: Temporal Co-indexing technique – the speech recording is time indexed using handwritten or photo annotations.

6.4 Naturalistic Analysis

6.4.1 Method

I collected *naturalistic* data about DR usage for 98 students for both *handwritten* and *photo* annotations. The students were aged 19-35, 42 were women and 56 men. They attended the lecture once, but they could access DRs of each lecture as many times as they wished, throughout the course using the CW and PW interfaces. I logged details of access sessions for all participants, including their frequency, duration, as well as when they occurred and which types of annotations (photo/handwritten) were used.

I analysed logs to understand the basic characteristics of access sessions and how participants used the DR recordings: did they focus on one part of the recording, e.g. to answer *particular questions* about a part of a lecture, or did they typically have longer sessions to get the *gist* of the entire lecture? I also collected data about students' *native language* to determine whether non-native speakers used the system more, as well as information about class *attendance* to determine whether non-attendees relied more on the digital record to compensate for missing a class. Finally, I wanted to determine the learning benefits of accessing recordings. Did students who accessed recordings more often obtain better *final grades* for the course? I also collected more informal feedback via an email site, and through anonymous surveys circulated to the students to gather feedback about their opinions regarding the DR system (see Appendix B).

6.4.2 Results

Overall users were very positive about the system. People were happy to volunteer for the role of official 'note-taker' whether this was to generate handwritten or photo annotations. Participants found it straightforward to use the system to access the DR, and anonymous comments were largely positive.

Several users pointed out that DRs encouraged active exploration of the lecture, the annotations helped jog their memory as well as encounter information they had originally missed.

"*[DR] gives you the chance to move around … in the lecture, as well as jolt your memory in some cases by seeing the pictures of the lecture at different points.*"

"*after listening to some of the audio there were things that I missed and it has been a useful review.*"

Others felt that DR information was more compelling and engaging than standard handouts or notes, because materials were multimedia.

"*DR is much nicer to listen to rather than staring at loads of text. It is easier to make notes while listening rather than looking back and forth at a computer screen. Using DR is much more interactive than just looking at standard notes so does tend to help keep my attention, as I prefer to be doing something interactive, than just reading through notes.*"

Even when students had missed a lecture they felt that DRs provided a sense of 'being there' that traditional handouts did not afford: "*[with DR] you have a sense of being in the lecture even if you missed it.*"

Some participants also pointed out the *social* learning benefits of DRs: DRs provided insights into what the annotators thought was important. By scrutinizing the notes or photos, students could understand and reflect on what others had thought was critical, which may differ from their own opinions:

"*I also feel that with other people making notes you get an insight in what they think, more than one opinion is very useful.*"

However, there were suggestions that DRs might induce changes in learning behavior: various students felt that DRs led them to attend fewer lectures, or to pay less attention during lectures they were at:

"The availability of [DR] has made me much less worried about missing lectures as, especially [with] textual annotations, you can listen and make your own notes by listening as it's just the same as being in the lecture."

"in parts DR encouraged me to not pay as much attention as I knew that I would be able to look through them at home after the lecture."

6.4.3 Session Characteristics

Our basic unit of analysis was an *access session* - defined as an unbroken interval spent accessing a specific lecture. If a student accessed another lecture or logged off, that session was deemed to have ended. I excluded sessions longer than 3 hours, containing fewer than 5 clicks. Here it was assumed that participants had moved away from their machine and forgotten to close the browser.

There were large differences in access patterns between participants. Some accessed DR multiple times during the course, others more minimally. However, out of the class of 98 people, I had 54 actively accessing the system i.e. for more than one session (making 59% uptake of DR technology). These active users had 7.74 sessions on average, overall spending 44 minutes on average using the system (see Table 6.1).

Before proceeding with parametric statistical analysis, I applied Kolmogorov-Smirnov (Sheskin 2007) to test for normality in data collected. The results of this test have confirmed that data collected is normally distributed.

First I analysed the *types* of sessions for active users, as this offers clues about what functions the system served (see Table 6.1). The predominant overall pattern was to access a given lecture multiple times for relatively short sessions, rather than selecting one part of the lecture and listening to it without interruption. This pattern of multiple operations occurring within a relatively short session, suggests the main use of the system was to access specific material *within a lecture* rather than to get the *gist of an entire lecture.*

	Mean	Std. Deviation
# Sessions	7 74	3 24
Views/Session	1 48	0 78
Clicks/Session	12 10	5 36
Mins/Session	7 80	5 45
Total # Mins	43 96	22 11

Table 6.1: System usage for Active Users (students with more than one access session).

When do students access the system?

I next looked at what *points in the course* students most accessed the system. Figure 6.4 shows the total number of system accesses for all lecture recordings during each week of the course. Thus in week 13, there were 193 accesses of all course lectures. There is a clear relation between active system usage and coursework evaluations. There was an initial period of exploratory system activity in week 2, just after the system was introduced. Coursework evaluations occurred in weeks 3, 8, 10 and 13, and on each occasion, I saw greater system use. These data support the view that the system was being actively used to prepare for coursework evaluations. User comments support this: *"I mainly used [DR] in order to do the tutorial and other coursework."* The importance of the system as a revision aid was further demonstrated by the fact that a system failure during week 13 led to a flurry of emails requesting that the system be reinstated as soon as possible.

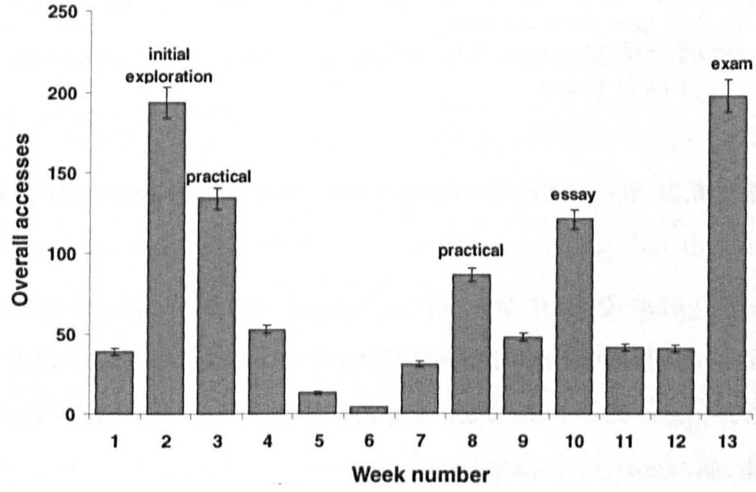


Figure 6.4: Overall frequency of system access over the duration of the course.

Who uses the system and why?

For all students, not just active users, I next examined whether *non-native speakers* made more use of the digital record. The results (see Table 6.2) are striking and confirm our hypothesis: non-native students make far greater use of the system. They both had a larger number of sessions (one way ANOVA: $F(1, 96) = 6.35$, $p < 0.01$), and also a greater number of accesses ($F(1, 96) = 5.41$, $p < 0.02$). Because of these large differences between native and non-native students I analyze the two populations separately in our subsequent analyses. It seems that non-native students made greater use of the system because it allowed them to revisit information that they had missed during the lecture because of their lack of fluency in English. This finding is highly relevant for programs offering international distance education. One overseas student commented:

	Non-native speakers	Native speakers
# Sessions	6 53	1 22
# Clicks/Session	9 79	2 64

Table 6.2: Access behaviour of native and non-native speakers using Digital Records.

"When [Digital Records] are used in lectures, I find I can concentrate more than with traditional methods, because I know that if I miss something important, the lecture can be listened to again. This is also a benefit if I've not fully understand what was being said as I can listen to the lecture at my own pace."

Next I looked at the relationship between *attendance and system usage*. I expected that students who attended fewer lectures would make much greater use of the system. As expected, there was a strong correlation between the number of lectures missed and the number of access sessions. However this was true only for native students ($r(35)=0.36$, $p<0.03$). There was no overall difference in attendance between native and non-native students ($F(1,96)=3.1$, $p>0.05$), so non-native students' greater use of the record does not seem to result from absenteeism.

Effects of System Use on Course Performance

I then explored the relationship between system usage and overall student performance on the course. Given the differences in access behaviour between native and non-native students, I conducted two separate linear regression analyses: one for each population.

I included the following independent variables to characterize system use: # sessions, # access clicks/session, time using the system, session length. I also included a *general measure of student ability* (their general score for their entire degree last year) as another independent variable. This was to control for the possibility that more able students might be more motivated to use the system, making their higher marks on this course the result of their inherent ability rather than system use. These variables were regressed against the dependent variable of overall mark for this particular course.

The overall regression model for native students is shown in Table 6.3. It is highly significant (R squared = 0.744, $p < 0.001$), showing that the independent variables in the model combine to strongly predict course mark. However only two of the independent variables (Degree Score and # Sessions, both **bolded**) are individually significant. As expected, students' overall ability/motivation (as measured by last year's degree score) predicts their final mark for this course ($p<0.001$). More importantly, when I factor out this ability/motivation, the number of DR sessions still predicts their mark for this course ($p< 0.05$). Thus for native students, we can conclude that greater system usage promotes higher coursework scores, when a control for overall ability is introduced.

Independent Variable	Standardized Coefficient	t	Sig.
	Beta		
(Constant)		-3 340	002
Last Years' Degree Score	**.879**	**8.178**	**.000**
# Sessions	**.439**	**2.118**	**.043**
Clicks/Session	- 148	-1 010	321
Minutes	-1 024	-1 024	245
Mins/Session	833	980	336

Table 6.3: Regression model showing user behaviours and general ability as predictors of final course mark for native students. Significant variables in bold.

I conducted a second regression with the same variables for non-native students. There was no relation between system usage and overall course scores, even when I controlled for overall ability.

Which indices are used most, photos or handwritten text?

Finally I compared the use of the two types of indices, namely handwritten text and photos. Overall, people accessed *handwritten annotations* more than photo *annotations*: (means 4.41 vs 2.49, one way ANOVA $F(1, 96)=2.32$, $p<0.03$) and they spent more time listening using *handwritten annotations* than with *photo annotations* after clicking: (means 16.95 vs 7.98, one way ANOVA $F(1, 96)=2.56$, $p<0.01$).

This behaviour is consistent with subjective user feedback. People felt *handwritten annotations* provided more fine grained indices when scanning for relevant information. With *photo annotations* it seemed more difficult to find an appropriate visual index to the underlying speech.

"The [photo annotations] were not very visible and only showed photos of the slides so just the set of slides would have been preferred. However, the [handwritten annotations] were very useful as they were practically an in-depth look into the slides as the lecturer is explaining the slides."

I thought that non-native students may find it hard to interpret others' scribbled annotations and be more reliant on photo indices. However, the opposite was true: non-native speakers were more reliant on *handwritten* annotations than native speakers ($F(1, 96)=5.95$, $p<0.02$).

6.5 Controlled Quiz Study

The naturalistic data offered strong evidence of when, how and who made use of the DR system. In my next study I wanted to *quantify* these effects under more controlled conditions, and to identify the *mechanisms* by which the effects were achieved. I did this by assessing student performance with and without DR on a series of specific class quiz questions that probed knowledge of material delivered in previous lectures. The previous study had also suggested that non-attendees made active use of DR for lectures they had not attended in person. I therefore wanted to determine whether non-attendees used DR in a different way from attendees and the extent to which they were able to compensate for their absence by directly accessing information from the DR. In this study I specifically focused on native students as our results suggested they had experienced greater benefits from system use in the previous study. The specific research questions I addressed were therefore:

- To what extent does DR improve student performance compared with traditional tools such as handouts and personal notes?
- Can DRs compensate for missing a lecture, i.e. can non-attendees use the DR to access information from lectures they missed?
- Which annotations provide better retrieval indices, photos or handwritten text?
- Does having a DR lead students to change the strategies they use for accessing course information, and if so how?

6.5.1 Quiz Questions

I asked participants to individually answer four specific questions from four previous lectures (see Appendix B). The lectures had been presented by the same professor, who also generated the quiz questions. I conducted a pilot to ensure the questions could be easily understood, were of comparable difficulty and did not contain any unfamiliar or unusual terms. Questions were chosen so that they could not be answered directly from lecture notes (to prevent participants from answering questions simply by reading the lecture handouts). An example question was: *'How do search engines match a document to a query?"*

6.5.2 Procedure

The quiz was run using a custom built website. Students were first given a general description of the quiz, the questions and the CW and PW systems. I then gave them a brief hands-on tutorial on how to use each interface.

Participants had access to *handwritten annotated DRs* using CW for half the quiz questions and for the other half access to *photo annotated DRs* using PW (see Appendix B). The annotations had been generated by 2 students who did not participate in this study. Order of presentation of these user interfaces was counterbalanced across students. For all questions, they also had access to *traditional reminders (TRs):* I provided them with the relevant handouts for each lecture and they were asked to bring their own personal lecture notes along to the quiz. Of course students who had missed specific lectures did not have personal notes for those lectures. Participants were also told which lecture the questions were generated from. They answered quiz questions on web-based forms. After the quiz they rated the interfaces and compared them with traditional reminders (such as handouts/personal notes) along various usability and efficiency dimensions.

6.5.3 Measures and Variables

I collected the following data:

Quiz score: Accuracy for each question was scored in the following way. I first generated a target answer for each question, by having 2 graders, blind to the experimental hypotheses, listen to the lecture recordings and read the relevant lecture notes. The graders agreed on the target answers, specifying keywords and context that needed to be present in each answer. Quiz scores ranged from 0-5 depending on how much of the target answer the user specified. For instance, if an answer included all target keywords (or their synonyms) and context, it received a maximum score of 5. Partially correct answers received a proportion of the score depending on the number of keywords and context answered. Scoring was carried out independently by the judges, and disagreements were referred to a third judge for resolution.

Retrieval Strategy: I defined three strategies. Students could answer questions using: digital records (DR), traditional reminders such as lecture handouts or personal notes (TR), or their unaided organic memory (OM). After each question, I asked students to note which of these strategies they had used to answer it.

Attendance: I collected data about whether each student had been present at the lecture related to the specific quiz question – allowing us to examine the effects of attendance on performance and response strategy. This information was available from class attendance lists.

Index Type: For each question, users either had access to *handwritten* or *photo* DR annotations.

Subjective Evaluation: After the quiz, I gave participants a brief survey, asking them to compare retrieval with DR versus TR. They were asked questions about (a) Benefits (b) Preferences (c) Interest (d) Attention and (e) Attendance. Responses were generated as 5-point Likert scales. In addition, I asked open-ended questions about what participants perceived to be the key differences between the tools they had used and why they preferred one to the other.

6.5.4 Participants

Thirty five student volunteers took part in the quiz, (10 women and 25 men, aged 19-35). They had no prior knowledge of the project or our research hypotheses. None had prior experience using handwritten or photo annotations, but obviously all had extensive experience of traditional reminders and their own organic memory. All were native speakers of English.

6.5.5 Hypotheses

I had four hypotheses.

H1: <u>Effects of DR use on performance</u>: I expected more accurate responses when DR was used, compared with TR (traditional reminders) or OM (organic memory). This is because DR provides controlled access to a verbatim record of what was said.

H2: <u>Effects of attendance on performance:</u> I expected students to perform better when they attended the lecture. This should be independent of the retrieval strategy used.

H3: <u>Index Utility</u>: I expected people to perform better using handwritten indices because these offer more fine grained access to the DR than photos.

H4: <u>Effects of attendance on DR use:</u> I expected people who didn't attend a lecture to be more reliant on DR because they would have no recollection (OM) of the information presented, nor would they have access to TRs such as personal notes. In contrast, attendees should be more likely than non-attendees to use both OM and TR.

6.5.6 Results

H1 and H2 both concern performance, so I evaluated both in a single ANOVA with Response Quiz score as dependent variable, and Retrieval Strategy and Attendance as independent variables.

One concern with this analysis is a potential confound of *self-selection*, i.e. stronger students are more likely to attend lectures and to get higher marks on the quiz because of ability, regardless of the tools used. I therefore included in this analysis only those students who had attended *some but not all* of the relevant lectures. I were therefore able to compare the effects of attendance as a *within subject variable.* The analysis thus compares the performance of *the same student* for lectures they had attended versus lectures they had missed, and thus avoids this confound.

H1: Does access to a DR increase performance compared with other retrieval strategies?

For users who had attended some but not all lectures, I compared retrieval quiz scores: (a) when they used DR, (b) when TRs were used (i.e. people used handouts or their personal notes), (c) when they relied on their own memory (OM). The results are shown in Figure 6.5.

The ANOVA showed a main effect for Retrieval Strategy ($F(2,47) = 8.62$, $p<0.001$). Planned comparisons confirm H1 showing benefits for DR over OM ($p<0.0001$), and for DR over TR ($p<0.02$). There were no differences between TR and OM. Providing a DR therefore helps students better than traditional access tools, with students performing 16% better overall using DR. Not only does this show the utility of DR it also shows the potential value of others note- or photo-annotations; people are able to exploit annotations taken by someone else to navigate to an important part of a lecture.

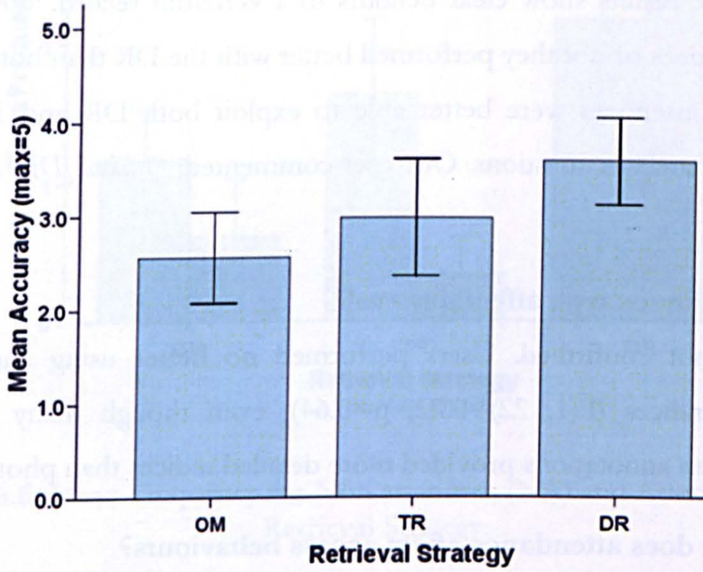


Figure 6.5: Quiz score for different retrieval strategies.

H2: How is performance affected by attendance and retrieval strategies?

The above ANOVA also showed a main effect for Attendance ($F(1,47) = 15.22$, $p<0.0001$). As predicted, students performed better when they attended a lecture. Planned comparisons also confirm that when students attended they remembered more using OM ($p<0.05$). There was also a suggestion that attendees were better able to exploit TR ($p<0.07$). Finally, even though DR provides access to a verbatim record, attendees were still better able to exploit this than those who missed the lecture ($p<0.01$). User comments support this: *"I felt less need to go to lectures because the recordings were available, however, I feel there is still a greater benefit from attending the lecture."*

One question arising from the results is why the non-attendees should use OM at all. Post hoc interviews with participants showed that non-attendees' frustration at being unable to answer questions using either DR or TR led them on occasion to guess likely answers – behaviour they classified as OM.

Overall the results show clear benefits to a verbatim record. Whether students were attendees or not they performed better with the DR than both TR and OM. However, attendees were better able to exploit both DR and TR to respond more accurately to questions. One user commented: *"I liked [DR] to find out what I missed."*

H3: Does index type affect retrieval?

H3 was not confirmed. Users performed no better using handwritten than pictorial indices ($F(1, 22)=0.22$, $p=0.64$), even though many users felt that handwritten annotations provided more detailed indices than photos.

H4: How does attendance affect access behaviours?

I conducted a second ANOVA with Frequency of Access as the dependent variable, Retrieval Strategy, Index Type and Attendance as independent variables. I included all students as I was interested in access behaviour so there was no need to control for confounds of attendance with ability/motivation. The results are shown in Figure 6.6. Confirming H4, there were main effects for Attendance ($F(1,99)=5.53$, $p<0.02$), and Strategy ($F(2,99)=25.78$, $p<0.0001$), and a significant interaction between Strategy and Attendance ($F(2,99)=25.78$, $p<0.002$). Planned comparisons showed, as predicted, that attendees were more likely to use OM ($p< 0.05$), and to use TR ($p< 0.0001$) than non-attendees. However, contrary to our expectations, non-attendees were no more likely to use DR. This may have been because non-attendees were unclear about which parts of the DR to access, leading them to have relatively few, but long, access sessions. Overall the results show that attendance affects strategy use, with non-attendees being forced to rely on DR to access information.

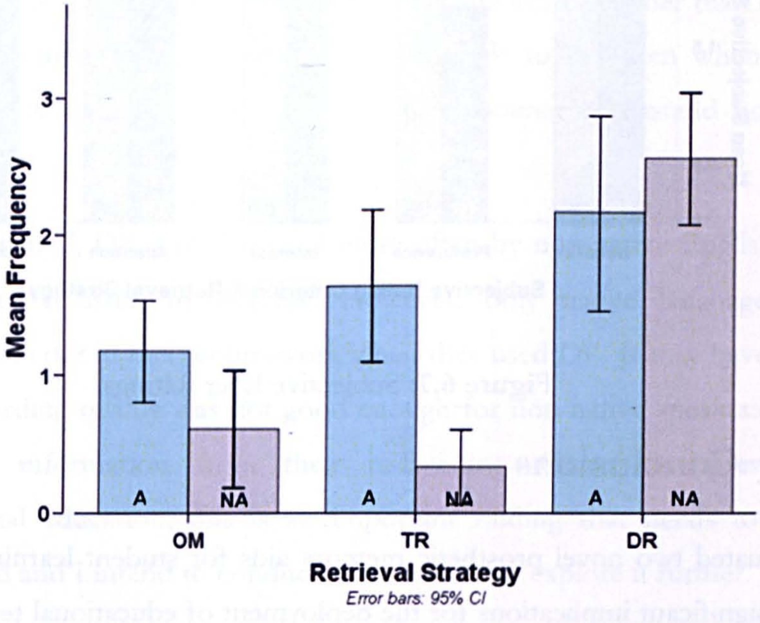


Figure 6.6: Mean Frequency for Non attendees (NA) and Attendees (A) per Retrieval Strategy.

Subjective User Feedback

I also analysed students' subjective feedback on DR tools and how they compare with TR, as illustrated in Figure 6.7. For various Likert judgments, users evaluated that DR was more useful than TR ($t(36)=2.47$, $p<0.02$). They also expressed an overall preference for using DR to access lectures ($t(36)=3.90$, $p<0.0001$). They judged that DR made lectures more interesting: ($t(36)=3.85$, $p<0.0001$), but felt they paid equal attention in lectures regardless of whether they could later use DR or TR ($t(36)=1.45$, $p>0.2$). People also felt less worried about missing a lecture knowing that it was recorded on DR ($t(36)=3.07$, $p<0.004$).

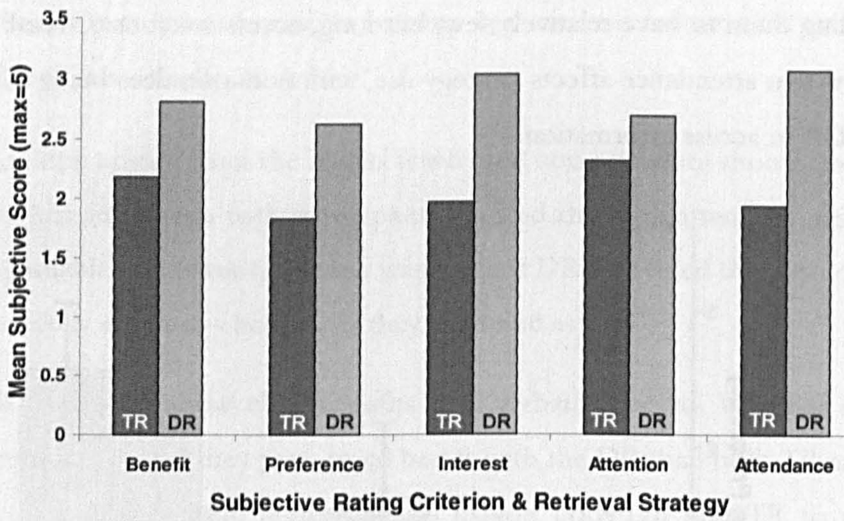


Figure 6.7: Subjective User Ratings.

6.6 Conclusions

I evaluated two novel prosthetic memory aids for student learning. The results have significant implications for the deployment of educational technology. DRs allow students to perform better on quizzes than existing tools such as TR and OM, and DRs were used strategically throughout the course. They were also accessed extensively by non-native students, and using them helped native students improve their grades. DRs were also considered more enjoyable to use than traditional teaching aids. Students were intrigued by the new tools, and once they discovered their full functionality and benefit, they were motivated to use them, particularly to help with their course work and exams.

DR recordings did not serve as a direct *substitute* for attendance. DR was a verbatim verbal record, making it possible for non-attendees to identify anything that was said from the DR of a lecture they had not attended. Although DR allowed non-attendees to answer at least some questions effectively, attendees performed significantly better on quizzes. The naturalistic analysis is consistent with this. Although non-attendees were more likely to access DR outside lectures than attendees, there was little use of DR to listen to an entire lecture. Instead, confirming other work (He, Grudin et al. 2000), naturalistic DR usage tended to be for specific questions or restricted parts of the lecture, rather than listening to an entire lecture. These results show care needs to be taken when deploying asynchronous learning technologies so that students understand how best to exploit them.

As I anticipated, DR was also used more often by non-native English speaking students to re-listen to lectures. However, only native language speakers performed better in their coursework when they used DR. It may have been that audio recording quality was not good enough for non-native speakers to extract important information from their re-listening. Given the prevalence of international education, this is an important finding that needs to be better understood and I intend to conduct more studies to explore it further.

There are important theoretical lessons to be drawn from these results. They are consistent with other prior work (Kalnikaite and Whittaker 2007; Sellen, Fogg et al. 2007) which suggests that memory aids are best used *in synergy* with existing tools, rather than as a replacement for them. I found strategic use of DRs. Most students did not stop attending lectures and rely exclusively on verbatim recordings. Instead the naturalistic data suggested many were trying to optimize what they had picked up in lectures, focusing on specific issues or parts of the lecture they hadn't understood. In particular they made strategic use of the tool when faced with the course assessments.

There are also design implications for our results. In the spirit of Web 2.0 and collaborative tagging, future technologies might combine multiple users' annotations of a lecture (Davis, Landay et al. 1999; Berque, Johnson et al. 2001; Ratto, Shapiro et al. 2003). Common annotations might be used to provide key word summaries of lectures. Similarly our retrieval interfaces might be modified to indicate annotation access popularity, with more frequently accessed annotations being made more salient. I explore this possibility in the next chapter. Finally, content analysis could provide ways to automatically index key lecture events such as main points (Tucker and Whittaker 2006), or even visual summaries (Foote, Boreczhy et al. 1998). These might supplement student annotations, and improve the browsability of the record.

Finally, by providing DRs, we might reduce pressure on students to record all that goes on and allow them to explore new material at their own pace (Bransford, Brown et al. 1999; Brotherton and Abowd 2004). By reducing the need to prepare for retrieval, we can free students to participate more in classroom discussions, allowing them to focus on important new concepts, rather than attempting to record all that they see and hear.

6.7 Chapter Summary

This chapter describes the evaluation of two PM devices, in naturalistic and controlled learning settings. Both devices provide access to annotated digital records (DRs) of lectures, freeing students from taking detailed notes, allowing them to re-access lecture recordings whenever they chose. DRs had benefits over traditional learning aids (e.g. handouts/personal notes): Students were more accurate in answering class quizzes using DR, and spontaneous DR usage outside lectures showed strategic access during important aspects of the course. Native speakers who used DR performed better on coursework, and non-native language speakers used DR extensively. Despite being a verbatim record, DR didn't substitute for attendance. DRs are thus a highly promising teaching tool, but PM devices are best understood as working in synergy with current tools to aid memory.

Chapter 7

Social Summaries to Augment Organic Memory for Learning

7.1 Introduction

The previous chapter demonstrated how PM tools can help individual users improve their long-term memory for complex materials. In this Chapter I explore how adding social feedback mechanisms to a PM device might improve the support the PM device provides.

This chapter explores ways that social interaction could help with everyday memory tasks. There has recently been huge interest in social computing, based on the proposition that the activities of many people can be exploited to help solve individual and small group tasks. For example, search engines rely on user linking behaviour and link labelling to re-rank results generated using textual IR methods (Brin and Page 1998). And social tagging systems such as dogear, dig, deli.cio.us, flickr, citeulike and reddit support the creation of labels and ratings allowing users to access content previously categorised by others.

Such systems tend to operate with large user bases on the web, with user tags and content generally being textual, although not always. Here, in contrast, I built and evaluated a novel system that explores social tagging in a smaller user population, where tags and content are multimedia.

I investigate the utility of tagging to construct *social summaries* of complex multimedia materials. This system allows students to apply time-indexed tags, such as handwritten annotations or photos, to different parts of multimedia lecture recordings. These tags can be used to straightforwardly access different parts of the lecture. The social component of the interface presents information about which tags are most frequently accessed by others: allowing students to infer which parts of the lecture have been of most interest or value to others.

It is well known that students have problems in remembering complex materials presented in lectures, and that their own notes or official handouts are often inadequate for later revision. In this system, a student annotates/tags the lecture in real time, e.g. by taking digital notes/photos when significant points are made. Using the principle of temporal tagging (Stifelman, Arons et al. 1993; Davis, Landay et al. 1999), these tags are time aligned with the original recording (see Figure 7.1). Clicking on a tag (e.g. 'explanation of Nielsen's usability heuristics') will replay what was being said when the note was made or picture taken. This allows the student to easily revisit specific parts of the lecture. The tags are therefore a *visual analogue* to the lecture content - allowing rapid access to key parts of the underlying speech.

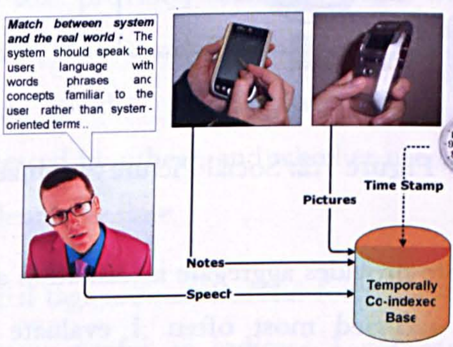


Figure 7.1: Temporal tagging with notes and images technique.

This system also provides social feedback about which tags and hence, which parts of the lecture, *other students* found to be useful. Following the principle of social navigation (Dieberger, Dourish et al. 2000), the new interface provides visual feedback about frequently accessed tags. Social tagging is a promising technique for multimedia summarization where effective content-based techniques have proved hard to develop (Tucker and Whittaker 2006; Whittaker, Tucker et al. 2008).

The interfaces for pictures and notes are shown in Figures 7.2 and 7.3 respectively. Figure 7.2 shows digital photos taken during a lecture. Each time a student clicks on a photo to access a specific part of the lecture, I make that tag more salient by enlarging it relative to other photos. For digital notes, I adjust their salience by highlighting/colour. Figure 7.3 shows how notes that have been used most frequently are highlighted in **bold** and in red.

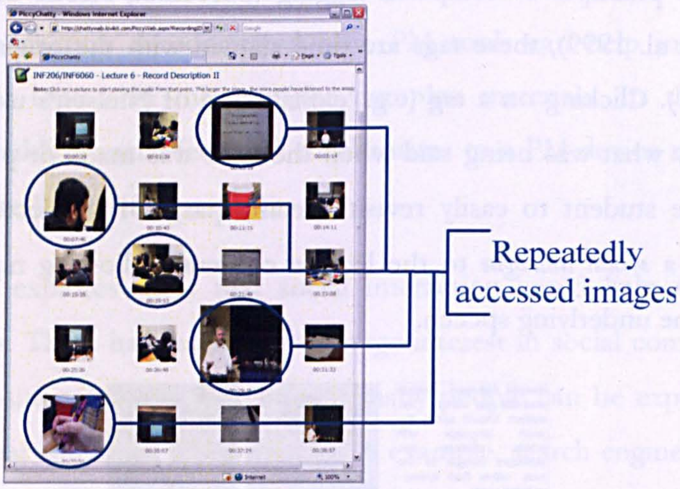


Figure 7.2: Social Picture Summary.

The interface therefore provides aggregate information about which parts of the lecture other users accessed most often. I evaluate social summaries in a naturalistic lecture setting where students used the systems to help with their everyday schoolwork.

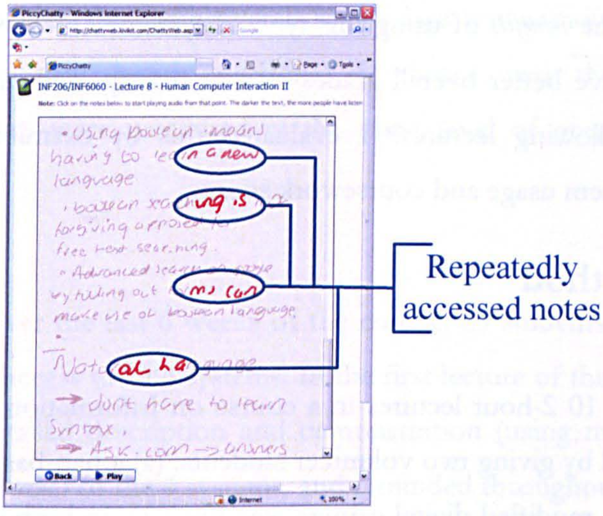


Figure 7.3: Social Notes Summary.

Our main research questions were:

- Do users make *greater use* of systems offering social feedback? While there are potential advantages to providing information about others' behaviours in supporting access to critical parts of lectures, there are also disadvantages: students may find the social interface confusing or want to make their own decisions about what is important. I evaluate these trade-offs by comparing use of system versions that provided social feedback, with versions that did not provide such feedback. I also examined the benefits of social feedback *within each lecture* by looking at whether people preferred to access indices that had previously been accessed by others, and whether use of such social information became more prevalent over time.

- Which *types* of social tags are most useful for retrieving lecture materials? Are photos or notes more useful as indices? I compared the utility of social summaries that presented digital *handwritten* tags with those that used *photos* to tag the underlying speech.

- *What* are the *benefits* of using this type of system? Does using the system help people achieve better overall grades or is it mainly used by students who have difficulty following lectures? I evaluated this by examining the relationship between system usage and coursework scores.

7.2 Method

Systems

For each of 10 2-hour lectures in a course on Information Retrieval, initial tags were created by giving two volunteer students: (a) a pen-based digital note-taking tool or (b) a modified digital camera (see Fig 7.1). In both cases tags were time-indexed to an underlying speech recording. There is normally a time delay between hearing an important part of the lecture and creating the tag. Based on prior work (Kalnikaite and Whittaker 2007) I therefore offset the tag by 1.5s to allow for this delay.

Volunteers were instructed to take notes as they normally would in a lecture. In the case of photos they were told to capture the images they thought would be most useful as retrieval indices. On average students took 174 words as notes, and 60 pictures. Fig 7.3 shows that digital notes were similar to regular handwritten notes. This is consistent with prior research (Kalnikaite and Whittaker 2007). Student comments in post hoc interviews revealed three main types of photo capture (see Fig 7.2): (a) *key slides* to capture verbal discussion associated with these, (b) *people* who asked questions, (c) *contextual images* of surrounding people or the lecture theatre to recreate the context of the lecture. One annotator commented: *"[I] tried to take pictures of people asking questions"* and *"when listening to what the lecturer was saying [I] tried to capture reactions of students."*

I used these tags to create 4 different versions of the system. Two versions included social feedback: Social Picture Summary (Fig 7.2), and Social Notes Summary (Fig 7.3). The other two *Asocial* versions (one with photos and one with notes) provided the same overall set of tags but without social feedback information. Thus, in Asocial versions, all photos were the same size and handwritten notes were presented at the same level of greyscale and colour salience.

Usage Data

The study ran over the last 6 weeks of the course. 25 students took part in the course, and had access to all 4 systems. In the first lecture of the course, students were given a detailed description and demonstration (using materials collected during a prior course) of the 4 systems, and reminded throughout the course that they could use them to access prior lectures. It was explained that the aim of each system was to help retrieve information from lectures, e.g. if they had found the lecture content to be confusing when it was initially presented, or if they simply wanted to revisit the entire lecture.

Access to the systems was provided via a webpage, which was prominently referenced on the course website. The webpage provided tutorials about each of the four systems, as well as links to access each of the systems. Users were allowed to freely choose which version of the system they wanted to access.

Naturalistic data about system usage was collected, logging the number and duration of access sessions for each student; which version of the system they accessed, which tags they used, and how long they listened to speech associated with each tag. I also elicited qualitative feedback via an email survey. Finally, I collected overall coursework scores to explore the relation between system usage and student performance.

7.3 Results

Do people make more overall use of tools that present social feedback?

Fig 7.4 shows the total usage time across all users for each of the 4 systems. There was almost no use of the Asocial systems, but significant use of the Social systems (respective mean usage times/user 10.9 and 462.1 secs, SD = 449.8, Wilcoxon's[4] $W(25) = -213$, $z = -3.2$, $p<0.001$). As one user said: "*…when I missed a lecture, I briefly looked at the [social summary tools] just to get a general idea of what happened in the lecture.*"

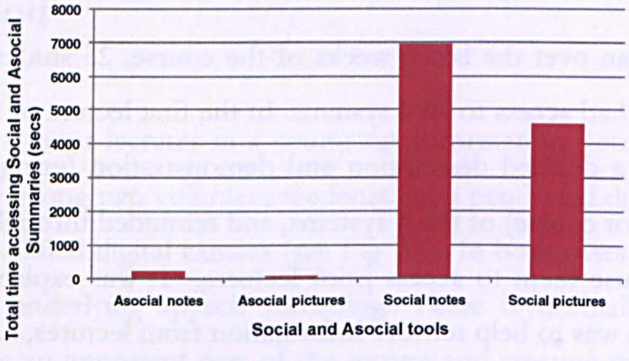


Figure 7.4: Social and Asocial Summary Use.

I also evaluated the use of social information *within* each lecture. Within a given lecture, I can distinguish between accesses that use tags previously accessed by others (i.e. popular tags), from those not previously accessed by others. Overall, the benefits of social feedback were clear: almost half (49%) of all accesses were for *popular* tags. This is well above baseline which would predict 3% repeat accesses for pictures and 1% for words if users had simply chosen tags at random.

Furthermore, this reliance on popular tags became more prevalent over time. If popular tags are indeed useful, then we should expect greater use of these over time. I divided the 6 week test into two halves. I compared the proportion of popular accesses (i.e. number of accesses using tags that had previously been selected by someone else, divided by the total number of accesses) in the first versus the second half of the trial. Table 7.1 shows that for the Social Summaries

[4] Non-normal distributions led us to use non-parametric statistics throughout.

there was a global shift to using popular tags in the second half of the trial. Social Pictures, shifted from 0 selections of popular tags in weeks 1-3 to *all* in weeks 4-6 (respective means = 0, and 0.25, SD = 0.5 Wilcoxon's W(25) = 91, z = 3.2, p<0.001). Similarly, there was also a significant shift in popular tags for Social Notes in the second half (respective means = 0.15, and 0.23, SD = 0.37 Wilcoxon's W(25) = 36, z = 2.2, p<0.03).

Tool	Social Notes Summary	Social Pictures Summary	Overall
Weeks 1-3	44%	0%	35%
Weeks 4-6	56%	25%	65%

Table 7.1: % of accesses of popular tags (those previously selected by someone else), for the first versus the second half of the trial.

Which tags are better: pictures or words?

I also analysed whether people tended to make more use of Social Notes versus Social Pictures. As Figure 7.5 indicates, there was greater overall usage of Social Notes, when I compared mean number of clicks for each type of tag for each participant (respective means = 88.83, and 9.2, SD = 156.4 Wilcoxon's W(25) = 133, z = -3.1, p<0.002) Why was this? It seemed users preferred notes because they provided a finer granularity of access (recall that there were almost three times as many word tags as pictures). Furthermore, some of the picture tags tended to be taken to provide contextual information (e.g. photos of those sitting close to the student in a given lecture), and such contextual information may have been less useful for retrieval.

As one user said: "*...I did not find the pictures themselves valuable...I found it easier to find the relevant part of the lecture using [social notes].*"

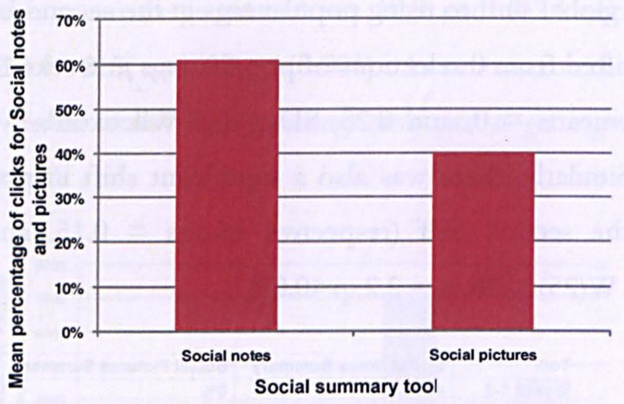


Figure 7.5: Comparing Note and Picture Indices.

Does Social Summary use predict higher coursework scores?

Finally, I looked at how system use relates to academic performance. There was a significant correlation between coursework marks and number of Social sessions (Spearman's rho(25)=0.72, $p<0.001$), but not Asocial sessions (rho(25)=0.42, $p>0.05$). It is possible that more motivated students are more likely to make greater use of our systems. However, this would lead us to expect an increase in coursework scores for *both* Social and Asocial systems. Instead the elevated coursework marks are *only* associated with Social system use, suggesting learning benefits that are specific to Social systems. One student described the benefits: *"Hearing the lecture notes whilst reading the lecture [social summary] notes helps greatly with revision. Today I was scanning through all the [social summary] notes whilst listening to them, and this helps enormously with revision."*

7.4 Conclusions

This chapter gives an overview of both implementation and evaluation of the utility of a novel working system that presents Social Summaries, derived from the analysis of tag popularity. My fieldwork results show, in a real-life practical setting, that this is a promising technique for accessing complex multimedia materials. The results are consistent with other research on social computing

(Brin and Page 1998; Dieberger, Dourish et al. 2000), as well as collaborative note-taking (Davis, Landay et al. 1999; Pimentel, Abowd et al. 2000; Brotherton and Abowd 2004) in demonstrating the utility of social information. They also promise to make an important contribution to multimedia access of conversational data, where effective automatic summary techniques have proven hard to develop (Tucker and Whittaker 2006; Whittaker, Tucker et al. 2008). This approach also finesses a weakness of many social computing applications in motivating users to tag. Here tags are generated as a *side-effect* of existing user practices, such as note-taking and, more recently, taking photos of presentations. At a practical level, the tool seemed to help the education process with greater use predicting higher course marks.

There are a number of possible extensions of the approach to related applications. For example, social tags might be combined for large scale social events such as rock concerts or sporting events where many people attending the same event might upload photos to construct a large scale social summary – allowing group access to favourite moments from the event. And the same techniques might be used to create shared tags for meeting records for the purposes of creating high quality minutes (Davis, Landay et al. 1999).

Another direction might be to develop techniques for presenting *configurable* summaries. Instead of supporting end user browsing through a set of tags, it might pre-compile an audio summary of a lecture based around the audio clips associated with, say, the 10% most popular indices. As with other temporal compression techniques (Tucker and Whittaker 2006; Whittaker, Tucker et al. 2008) users could specify a desired compression level, and the system would compile the appropriate summary.

Finally these interfaces present interesting social challenges that need further empirical investigation. Less dedicated students may increasingly rely on the efforts of more conscientious classmates: e.g., waiting for others to access and 'mark up' the main parts of the lecture before they begin their own revision. And of course there is a potential problem of 'overfitting': students may become so focused on popular tags that they neglect major parts of the lecture. Despite these potential problems, this approach shows demonstrable promise in an area where it has been difficult to develop effective techniques.

7.5 Chapter Summary

This chapter describes an extended notion of social tagging to construct *social summaries* of complex multimedia materials. This system allows students to apply time-indexed multimedia tags such as handwritten annotations or photos to different parts of lecture recordings. These tags can be used to straightforwardly access different parts of the lecture. The social component of the interface presents information about which tags are most frequently accessed by others: allowing students to infer those parts of the lecture of most interest to others. This chapter also demonstrates the utility of the approach in a 6 week fieldwork study. Social summaries are used much more than corresponding systems that do not provide social information. In addition, social tool use was correlated with high course marks.

Chapter 8

MemoryLane: Contextualising Autobiographical Mementos

8.1 Introduction

So far I have looked at memory for conversational data in practical settings. In this final empirical chapter I examine new tools for digitally recording autobiographical events. This Chapter extends the notion of *active* capture with Prosthetic Memory tools, where users actively make a decision and take a deliberate action to capture mnemonic information. The work is motivated by a study done by Petrelli, Whittaker and Brockmeier (Petrelli, Whittaker et al. 2008) which reported on the prevalence of home based physical mementos and how people retrieve associated memories by revisiting physical objects in their homes. Prior research has documented the significance of physical mementos which have strong associations to memories of past events and help with re-experiencing the past (Csikszentmihalyi and Rochberg-Halton 1981; Turkle 2007). These studies focus on understanding memory associations with *physical* artefacts, but with a few exceptions (Frohlich and Murphy 2000; Stevens, Abowd et al. 2003) much less work has been done on developing *digital* tools to 'capture' autobiographical experiences. The aim is to build a prototype exploring reflective and active capture of memorabilia, investigating how people store and contextualise their personal mementos. As established by (Petrelli, Whittaker et al. 2008) all of us are surrounded by physical objects in our homes, but how do I decide what objects to display, keep or add to a *digital* archive? And what *types* of objects do people want to keep as *digital* mementos? In general our attempts to develop digital archives tent to produce large passive storages of photos, videos or documents that are forgotten about, and seldom accessed (Petrelli, Whittaker et al. 2008; Whittaker, Bergman et al. 2009). Prior work on physical mementos (Petrelli, Whittaker et al. 2008) shows a focus on mundane everyday objects which express symbolic relations to significant people and events. However, most of these studies focus on home memorabilia, but do people only want to collect home–based memorabilia? To address this, our system provided support for other contexts where people might have strong reminiscences, such as outdoor places, communication items and other people. Previous research has shown that contextualisation is crucial for recall and reminiscence about past events (Smith 1979). By contextualising memories in this way, I hoped to create rich cues that might provide relevant information for finding an appropriate reference to an individual memory (Smith and Vela 2001). There have been a number of psychology studies investigating the relation between encoding context and subsequent memory showing that adding context enhances peoples' ability to remember (Bodner and Lindsay 2003). In a broader way, there are more complex contextual structures in place, such as social contexts (Zerubavel 1996), that we all apply periodically to help retrieve memories and create other associations to the past. Such social contexts are particularly active when memory concerns other people and relationships. Social context also applies to the study reported in this chapter, explaining why people re-cluster mementos of other people as their relationships with others change.

One of the best known Prosthetic Memory tools, MyLifeBits (Gemmell, Bell et al. 2006) has the primary aim of providing an appropriate context to help re-collect memories. Another tool for providing contextual information to help with remembering is Infocockpit (Tan, Stefanucci et al. 2001), which uses place based information to provide additional context. Other context aware systems such as Cyberguide use context of an outdoor location to help remember personal travel routes (Abowd, Atkeson et al. 1997).

Context has also been a primary focus for other Prosthetic Memory agents such as the Context-base Video Retrieval System (Hori and Aizawa 2003), Ubiquitous memories (Kawamura, Fukuhara et al. 2007) and SharedLife (Mori, Basselin et al. 2007) whereby real mementos are given associative attributes that help people recall of the past. However there is little research into how and when people actually utilise these Prosthetic Memory tools, and how do these tools map between physical and digital environments?

I also looked at narrative in the context of digital mementos. People like telling stories about their past and the objects that they have around them remind them of events, places and people they know (Frohlich, Kuchinsky et al. 2002; Crabtree, Rodden et al. 2004). Previous research shows that active creation of narrative is an essential part of remembering our past (Brockmeier and Carbaugh 2001) which helps us to reminisce about the past (Hoskins 1998). However there is little understanding of how and when people choose to narrate memory-related physical objects, places, people or communications. People like to tell stories about their lives: according to Tversky (Tversky 2004); they tell stories mainly to their friends, but less so to family and other groups of people. Tversky found that 58% of life narratives are told to inform and 38% to entertain. In this chapter, I also investigate how people tell stories for their own re-listening.

There also is a view that people tend to remember better when there is an emotionally strong association to mementos. According to Conway (Conway 1990), emotions play a role in the formation and maintenance of vivid memories. Tversky analysed what people narrate about their lives (Tversky 2004), and she noticed that most events were rated as emotional, intense or very intense, with half positive and half negative emotions. This chapter further explores mementos and their relationship to emotional states.

Furthermore, I examine what additional benefit could new generation tools provide that goes beyond the prevalent ways of organising digital mementos, namely digital picture collections. Previous work has shown that people are resistant to using novel software to organise such digital mementos, instead relying on existing file systems (Kirk, Sellen et al. 2006; Whittaker, Bergman et al. 2009). I wanted to explore whether a rich contextualised system would be preferred overall to existing context-free tools.

To explore the above questions, I designed and implemented a novel exploratory Prosthetic Memory tool called *MemoryLane*. Our tool allows people to capture, re-organise, archive and tell stories about events in their past through representations of their physical artefacts. I evaluate *MemoryLane* to understand how, when and why people decide which physical mementos to digitise and how do they go about doing it.

This chapter presents an exploratory software tool for contextualising collections of digital mementos, which is then evaluated with 31 users at *two* different time intervals to explore the creation of digital mementos. In particular, this tool supports contextualised archiving of:

1) artefacts: physical items that people keep in their homes;
2) locations: capturing and contextualising outdoor places on a map;
3) people & relationships: logging contacts and important relationships with other people; and
4) communications: preserving physical and digital communications.

It also allows people to annotate their mementos by creating associated narratives using either text or speech as well as to register an affective evaluation of each memento. I therefore explore multiple issues around capture and annotation of multimedia data.

Specific research questions that this exploratory study addresses are:

1. What ***type*** of mementos do people digitise and how ***frequently*** do they digitise these types of mementos? What types of mementos do people prefer to capture? Are they most focussed on people, places, or communications?
2. What ***attributes*** of mementos affect user's about whether to include in the digital archive, for example memento convenience.
3. How ***important*** are these mementos and do people digitise their most important mementos first? What types of mementos are more important?
4. What stories do people ***tell/narrate*** about their mementos? Do people find it easier to tell speech-based stories about their mementos or do they prefer textual annotations?
5. What types of ***emotions*** dominate among digital mementos? How do these change over time? Do people choose more emotionally positive or negative mementos?
6. Do people contextualise their mementos based on a ***timeframe*** to represent an overview of a certain period in their lives, or are they picked at random, or the basis of importance?
7. Do people ***reminisce*** at capture or recollection? Does active capture or passive recollection evoke a greater reminiscence effect?
8. Finally, what are the ***overall benefits*** of digitising physical mementos? Do people want to keep their physical mementos for later retrieval or is it enough to have them in their current physical form without a digital archive? Do current tools support memento capture and archiving well?

8.2 Experimental Method

This study investigates what type of memorabilia people digitise. There were *three* stages to the study: (a) **Capture** (b) **Reconstruction** and (c) **Long-term exploration.**

8.2.1 Capture

In the Capture stage, participants were given *two* devices for memento capture: 1) a Digital camera; and 2) a Dictaphone. They were instructed to capture at least 5 instances of each of *four* different types of memorabilia namely: physical mementos, outdoor locations, other people and physical/electronic communications.

At the beginning of the study, I gave participants a brief hands-on tutorial providing a detailed description of the memento capture collection technology and procedures for the experiment. They carried out practice tasks (one with each type of memento (home, places, people, communications) as well as for each of the different system functions e.g. narration and picture taking). Users were allowed to proceed to the capture phase only when they felt confident they knew how to use the technology they were given.

When people received the camera and the Dictaphone, they were given a clear set of instructions on the 4 theme areas: Home, Locations, People and Communications and told to capture at least 5 mementos in each of these areas. They were instructed to decide on a *timeframe* from which their mementos originated. This *timeframe* could be any previous period of their life that had a connection with their current life e.g. the time since starting a PhD, or since moving to the city where they now lived.

At the end of the 3 day capture session, each participant was invited for a brief interview and to hand over the digital mementos they had collected. During the interview, participants were asked a series of question about their capture experiences (see Appendix C).

I also asked them questions about how easy it was to digitise physical mementos. In particular about their experiences in capturing different types of mementos. I asked them: *Did you find it equally easy to capture different types of mementos?* I was also interested in *differences* between types of mementos so I asked participants to nominate their favourite capture theme. Here I asked them: *Do you have any preference as to what type of things you would like to capture more?* (see Appendix C).

8.2.2 Reconstruction

After **Capture**, the experimenter took each users' mementos and uploaded them into *MemoryLane*. They then discussed and manipulated these mementos in the **Reconstruction** phase.

Reconstruction (see Figure 8.1), followed the **Capture** stage and took place in two sub-stages: 1) Rearrangement of digitised mementos and association of relevant *speech* narration in *MemoryLane* (described in detail in the subsequent section); and 2) Evaluation of uploaded mementos: participants were asked to add an importance rating and emotional association to each memento. This involved users looking at their personalised *MemoryLane* populated with the memento data they collected in the previous **Capture** stage. *MemoryLane* ran on a tablet PC, Figure 8.1, allowing participants to use a hand held pen sized stylus for "grabbing" and "re-locating" thumbnails of their mementos on the screen.

Figure 8.1: Participants in their homes rearranging & evaluating their mementos using *MemoryLane*.

8.2.3 Long-term Exploration

After the **Reconstruction** stage they entered the **Long-term Exploration** stage: users were given *MemoryLane* containing all their mementos to take away on a personalised 2GB memory stick Figure 8.2, which when inserted into a personal computer ran independently (there was no proprietary software required to be installed on participants' machines). Participants were asked to use this tool as though it was a "multimedia photo album" on their own computers and to add as many mementos as they wished from any period of their life. They were instructed to add mementos from their existing digital archive or capture new mementos using their mobile phone (I ensured that all participants had mobile phones that could record *speech* and capture digital images). With their own everyday technology, participants created *speech* based stories, or later added textual annotation for each memento, rating its importance and associating emotions with their captured mementos. In addition, participants were asked to look again at their mementos in *MemoryLane* and were asked to modify and supplement what they had already recorded in the previous session (see Appendix C). There was a one month time limit for the **Long-term Exploration** stage.

Figure 8.2: 2GB memory stick with *MemoryLane* given to participants at the **Long-term Exploration** stages.

8.2.4 Reconstruction and Long-term Exploration Procedures

More specifically, during both **Reconstruction** and **Long-term Exploration** stages participants were asked to carry out the following tasks:

Evaluate

Look, listen, write or read about each memento in their personal *MemoryLane* and assess whether the information still applies or *update* it if anything has changed.

Modify

Re-locate mementos with *MemoryLane*, e.g. if people in winter took a picture of a plant in their home and then in spring moved this plant to the garden they should move this between different *MemoryLane* locations. In this case, the person should move the plant out of their House space in *MemoryLane* into the Garden space. I observed a lot of adjustments and re-shuffling in the People space, primarily due to the dynamics in the relationships that people have with other people. For instance, if there was an argument, that person would be moved out of the privileged people space to a space outside.

Rate Importance

During the **Reconstruction** and **Long-term exploration** stages, people were asked to rate the *importance* of each of their mementos. *MemoryLane* allow users to assign each memento an importance level on 5-point Likert scales using the common notion of popularity stars. All stars contained textual tags, so users could check what each number of stars represented. For instance, if a user selected one star, they would see a text tag – *"not important"*, for two stars, the text tag contained *"slightly important"*, for three stars – *"indifferent"*, for four stars *"very important"* and for five stars *"extremely important"*. Participants were able to see the textual tags before deciding on how many stars to select. All participants were familiar with using stars to rate importance from a similar concept used on film rating sites, e.g. YouTube, iMDb, or Guardian film reviews. This importance rating allows us to explore different reasons for choosing mementos.

Tell/Write Narrative

I was interested in exploring how people capture and retrieve narratives about their past. Thus I asked our participants to record spoken stories on a voice recorder/phone or textual stories using the digital stylus or by typing. We know that people generally like telling stories, particularly about their personal experiences. I anticipated that speech based story telling would therefore be more popular than textual annotations. I also anticipated people would generate more annotations (regardless of whether these were speech or text) in strategic cases where participants knew that they might forget the whole or fragments of a story at a later date.

Emotional Ratings

I wanted to understand if mementos uploaded onto *MemoryLane* had particular emotional attributes. Thus I asked participants to indicate their emotional state for each memento they uploaded. How did I decide which emotional states to include in the list? From emotion research, there is little consensus about what emotions to present and how to illustrate them. *MemoryLane* uses the popular notation of Emoticon expressions, which most people are familiar with. These emoticons are also used in text messaging chat clients, e.g. MSN or Skype. The emotion list in *MemoryLane* consisted of Ekman's *six* basic facial expressions (Ekman 1999): *anger*, *disgust*, *fear*, *happiness*, *sadness* and *surprise*. However from previous explorations into emotional states and memory (Kalnikaitė and Whittaker 2008), I discovered that people wanted more variation in their emotional expression. Therefore from that study I added the following emotional states to Ekman's list: *ashamed*, *confused*, *depressed*, *disgusted*, *love*, *relaxed*, *sarcastic* and *scared*.

At the end of each stage, I asked participants to compare *MemoryLane* and the Windows File Navigator (in thumbnail view mode), because it is known that people are very resistant to using novel software applications for organising mementos (Whittaker, Bergman et al. 2009). I logged all the data collected by participants their ratings, and narratives as well as any changes and additions made between the sessions.

8.3 MemoryLane User Interface

This section describes *MemoryLane* and its functions for contextualising memorable physical mementos at Home, of People and relations, for outside Locations and Communications. As mentioned before, *MemoryLane* runs on a tablet PC. It is initially accessed through a personalised area which depicts the person whose *MemoryLane* it is, as illustrated in Figure 8.3. The central picture shows an image of the participant to give a personalised impression. The menu bar on the left hand side only becomes active when the user enters any of the *four* main areas: Home, Locations, People and Communications.

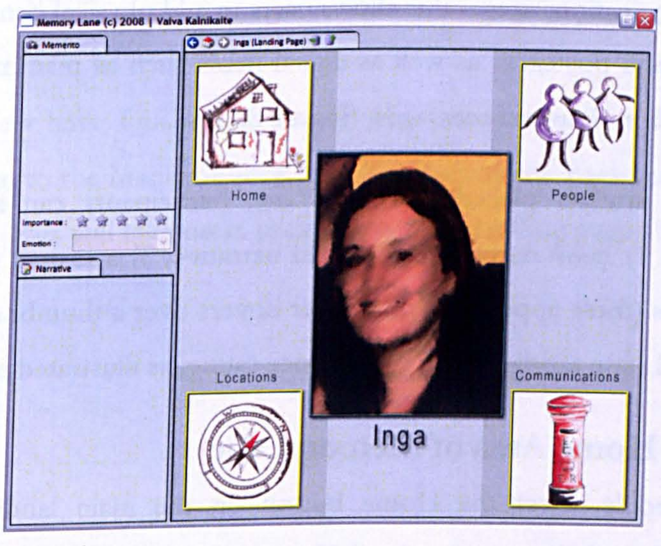


Figure 8.3: Main landing page of *MemoryLane*
(Image displayed on this illustration has been made anonymous).

From the main navigation area, users can access any of their *four* personalised areas:

1) Home area – contains a conceptual sketch of a home which users can populate with digital thumbnails of physical artefacts from their home, (Figure 8.4);

2) Locations area – world map which allows users to place their memorable outdoor places on the map as well as access localised, more granular maps (Figure 8.5);

3) People area – contains a conceptual 'photo frame' where users can place images of other people they want to remember and cluster them based on their relationships, (Figure 8.6); and

4) Communications area – this allows users to add physical communications such as letters and postcards as well as digital items such as msn messages, facebook postings or mobile text messages, (Figure 8.7).

For each artefact placed in *MemoryLane*, participants can add the following attributes: 1) *speech* narrative; 2) textual narrative; 3) a textual tag consisting of a few words (these appear when the user hovers over a thumbnail of an object); 4) emotional state rating; and 5) importance rating, as illustrated in Figure 8.4.

8.3.1 Home Area of MemoryLane

When people select the Home button on the main landing page, they are redirected to their conceptualised Home area. After the Capture stage, this area was populated with individual home mementos visually presented as thumbnails (*hotspots*) dotted around the house, see Figure 8.4. Participants were able to "select" these thumbnails with the stylus pen and "re-locate" them to other parts of their digital home by dragging them around the screen. When participants select a thumbnail, an enlarged picture of the selected memento is displayed on the left hand side navigation bar, as illustrated in Figure 8.4. If there is *speech* associated with the selected memento, it begins to play automatically. Users can rewind, pause or re-play the *speech* with the player functions illustrated in Figure 8.4. Underneath the enlarged memento picture, users have the option to rate the importance of their memorabilia. Importance ratings are carried out on 5 point ratings star system, where 5 stars mean that this memento is "*extremely important*", and 1 star means that this memento is "*not important*". Underneath the importance ratings, users are able to add emotional status from a drop down menu. Finally, participants can add textual narrative about the selected memento, which is done in handwriting with automatic character recognition.

It is important to note that when participants upload their mementos, they have to provide a description tag (which can consist of one or a few words) for each memento. This tag becomes visible when hovering over the memento's thumbnail.

When the users have finished uploading their mementos into the Home area, they can return to the Main entrance page by selecting the Back button at the top of the screen. This will take users back to the main landing page shown in Figure 8.3.

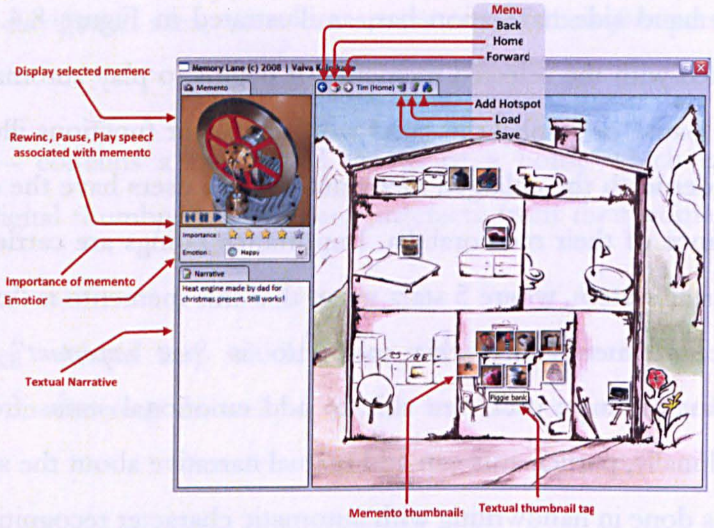


Figure 8.4: Home area of *MemoryLane* populated and annotated with images of memorabilia. The thumbnails show where the mementos are located in the Home.

8.3.2 Locations Area of MemoryLane

In the Locations area, users have the same functionality as described in the previous section Home area (i.e. to add, manipulate, annotate and rate mementos). However the Locations area not only provides a map context, but also a hierarchy functionality to allow participants to zoom in and out of maps. Since people like travelling, which can be widespread, this demands access to a complete map of the world. However participants also want to be able to view local places, e.g. their work place. Figure 8.5 illustrates a hierarchically interactive thumbnail structure that supports global to local navigation. Thumbnails that are surrounded by a yellow border, are the top level mementos, that link to local maps, e.g. the bottom thumbnail in Figure 8.5 would redirect the user from the global/national map to a local map with further thumbnails at their local level of interest.

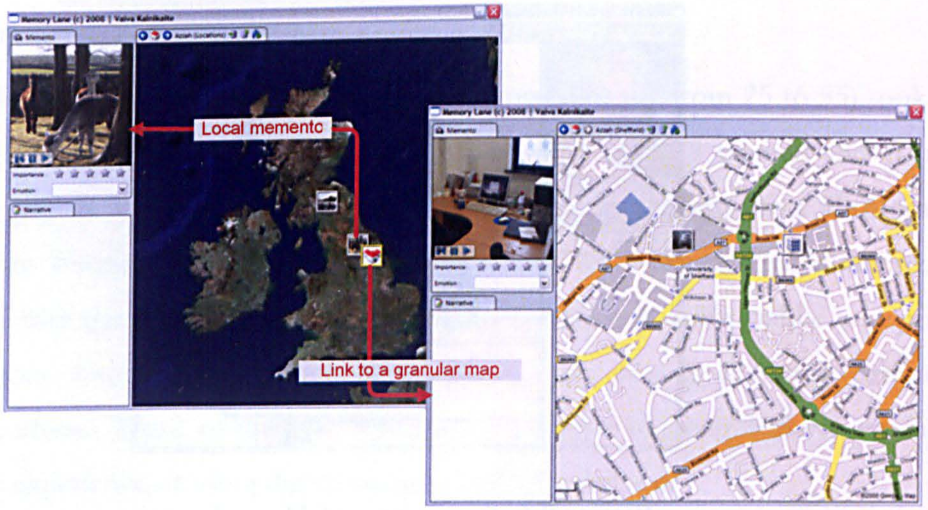


Figure 8.5: Locations area of *MemoryLane* – hierarchical thumbnail structure – zooming from global/national to local map level.

8.3.3 People Area of MemoryLane

The People area, as illustrated in Figure 8.6 is accessible from the main page shown in Fig. 8.3. This space contains a sketch of a picture frame displaying thumbnails of people that participants might want to remember in the future. This area has the same functionality as the Home area described in the previous section (i.e. people can add, manipulate, annotate and rate mementos). People can also organise objects within the frame to represent different social organisations such as friends/family etc.

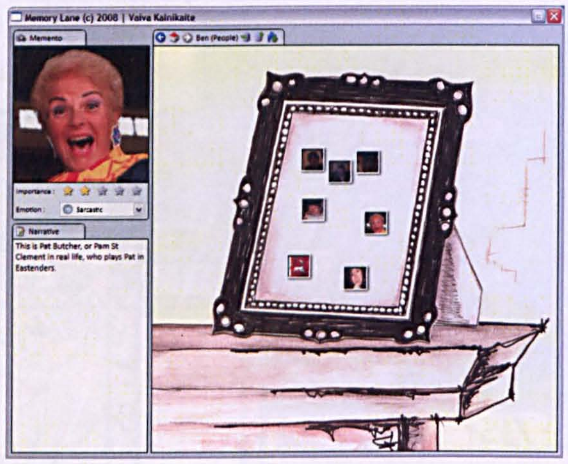


Figure 8.6: People area of *MemoryLane.*

8.3.4 Communications Area of MemoryLane

This area contains a sketch illustrating different forms of communication, suggesting to participants that they can upload digital as well as physical items, e.g. e-mail, mobile text messages, social site postings or digital representations of physical letters and postcards, as illustrated in Figure 8.7. For physical communication artefacts, people were encouraged to digitise in one of the following ways: a) take a digital picture of an artefact; b) scan the artefact; or c) take a screen shot or a photo.

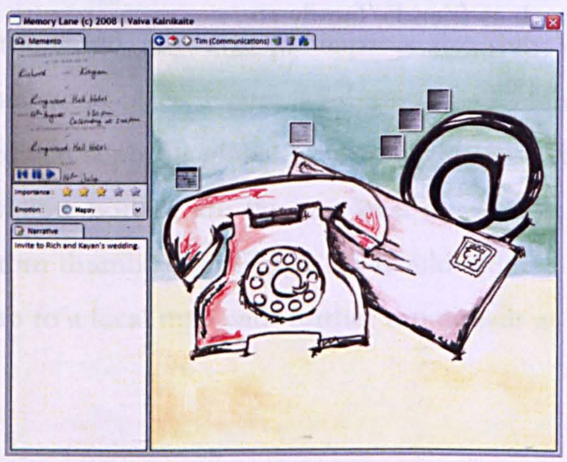


Figure 8.7: Communications area of *MemoryLane.*

8.4 Participants

Thirty one users (16 women and 15 men, ranging in age from 25 to 55) took part in the **Capture** phase of this study. 19 of these users (9 women and 10 men, ranging in age from 25 to 55) also completed the **Long-term Exploration** phase. Users were volunteers consisting of university researchers, administrative and management staff, as well as other professionals from public and private sectors. Users had no prior knowledge of the project or our experimental hypotheses. None of the users had prior experience of using *MemoryLane*, but all had experience of using digital cameras and a Dictaphone.

8.5 Hypotheses

These hypotheses address motivations and affects for capturing and re-accessing digitised physical memorabilia.

- H1 – *Importance*: It was anticipated that more important mementos would be added first. I also expected that mementos from Home and of People would be rated more important than outside Locations or Communications. I also expected that in their narratives people would have more to say about mementos that are more important to them, so their *speech* or textual narrative would be more detailed for more important mementos.
- H2 – *Memento type and upload frequency*: Although I expected people would try to capture important mementos, capturing different types of mementos requires different levels of effort. For example, capturing images of significant remote locations might require people to visit those locations. In contrast, capturing images from the home or of people seen everyday is considerably easier. It was anticipated that people would upload more of the mementos that are easier to capture.

- H3 – *Memento distribution in the home context:* I expected that people would capture more mementos from public spaces in their homes, e.g. living room or the kitchen. Less capture would be done from private places e.g. bedroom. Even less capture would be done from spaces such as the attic, where memorabilia are often hidden in secret boxes that are only occasionally opened (Petrelli, Whittaker et al. 2008).

- H4 – *Narration:* Given the centrality of narrative, it was anticipated that more speech, than textual narratives would be created, since speech tends to be easier to generate (Tucker and Whittaker 2006). In terms of the relation between narrative type and memento types, I assumed participants would be more inclined to *tell* longer stories for mementos in People area, because there might be more to tell about people than about other types of mementos, because people are more important in general.

- H5 – *Emotions:* From the photos literature, it was expected that people would upload more emotionally positive mementos (King 1986; Chalfen 1987). I also anticipated that different areas, in particular Locations, or Communications might have more neutral emotion.

- H6 *Timeframe:* It was expected that people would choose mementos that represented more recent events, mainly because I explicitly asked them to focus on their "current life period" and also because more mementos related to the recent past might be easier to create. However during the long-term exploration stage, it was expected that people would upload mementos representing events from the more distant past, because they were given more time to reflect on their past.

- H7 *Reminiscence:* I was interested whether people would experience richer feelings of reminisce during the Capture stage or Reconstruction stage. I anticipated that people would reminisce more during the Capture stage, because this was the time they were reflecting on what they wanted to capture, as well as actively recording the narrative. In contrast, the later stages seemed to involve passive re-listening to those narratives.

- H8 *Overall Benefit:* Prior work has shown a preference for using familiar tools to organise mementos (Whittaker, Bergman et al. 2009). However, I anticipated that context based tools such as *MemoryLane* would be judged as more fun to use, compared with everyday tools like Windows File Explorer. I also anticipated that tools allowing contextualisation would also be judged as more appropriate for archiving personal digital memorabilia.

8.6 Results

8.6.1 Memento Collections

There was a general pattern in the kind of mementos participants collected. The following outlines some of the main themes in the different *MemoryLane* areas:

Home area: as illustrates in Figure 8.8, from left to right. Participants collected artefacts that: (a) reminded them of certain regular activities, e.g. cycling, playing piano or ironing, keeping a light on during the night while children are asleep, or teaching a toddler how to use a potty; (b) were received as presents, e.g. a small sculpture of a lighthouse, a mug; (c) things participants habitually collected over time, e.g. pebbles or rugby balls; (d) wall displays and artworks, e.g. posters or paintings; (e) objects associated with intellectual activity, e.g. music CDs/DVDs, lecture notes or books; and (f) favourite foods, e.g. dinner ingredients for a Friday night.

Figure 8.8: Examples of artefacts in the Home area of *MemoryLane.*

<u>Locations area:</u> as illustrated in Figure 8.9, from left to right. Participants collected artefacts that: (a) reminded them of daily tasks, e.g. going to work or the gym; (b) local places they liked to visit at weekends, e.g. the Botanical Gardens, Llama sanctuary or a local park; (c) places associated with certain

interests, e.g. stadium of a favourite football team; (d) places they liked to eat out, e.g. a pub, "Kebabish" take-away or a Chinese restaurant; (e) holiday places, e.g. Inca trail; (f) outdoor objects of interest, e.g. trams; (g) visiting friend's places, e.g. a view from a friends' balcony.

Figure 8.9: Examples of artefacts in the Locations area of *MemoryLane*.

People area: as illustrated in Figure 8.10, from left to right (personal images were intentionally blurred). This area generally contained collections of images of friends and family. However participants collected other images that represented: (a) family members at gatherings, e.g. family dinners; (b) friends or housemates in action, e.g. a housemate cleaning a fridge; (c) work colleagues, e.g. a supervisor; (d) children and partners, e.g. new born son; (e) famous people that participants met or saw, e.g. Ken Doherty – famous snooker player; (e) pets, e.g. my cat.

Figure 8.10: Examples of artefacts in the People area of *MemoryLane*.

Communications area: as illustrated in Figure 8.11, from left to right. In this area, people collected the following: (a) facebook entries, e.g. arranging to go out with friends or funny citations; (b) mobile text messages, e.g. messages exchanged with a friend on a specific topic; (c) typed or handwritten letters, e.g. winning a photographic competition, rediscovered letter from dad written a long time ago or the last letter from grandmother; (d) correspondence that evoked further planning, e.g. wedding invitations; (e) tickets, e.g. concert tickets; (f) guides and other manuals, e.g. a tourist guide to Venice.

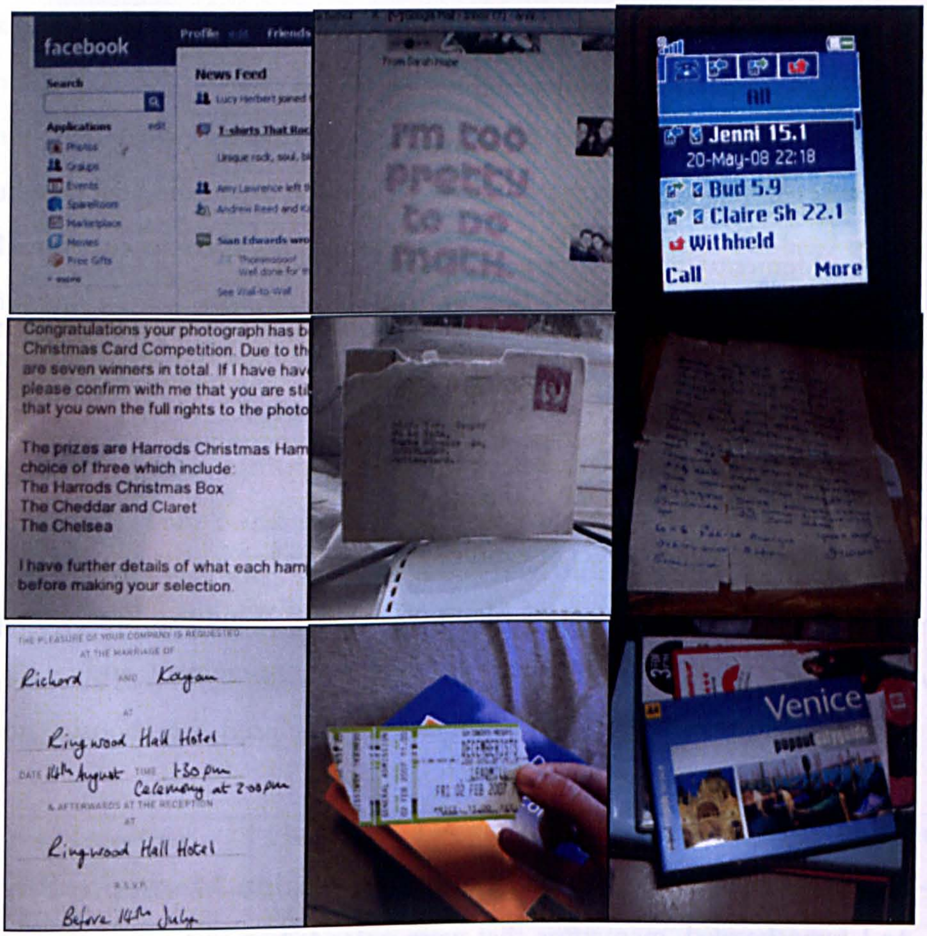


Figure 8.11: Examples of artefacts in the Communications area of *MemoryLane.*

8.6.2 Re-construction of Mementos

Overall in this study I collected 460 mementos, including 160 pictures in the Home area (examples in Figure 8.8), 128 pictures in outdoor Locations (examples in Figure 8.9), 108 pictures of People (examples in Figure 8.10), and 64 pictures of Communications (examples in Figure 8.11) items.

There were *two* recall phases: **Reconstruction** – initial exploration of mementos that people collected in the 3 days that they were given the capture technology;

and **Long-term Exploration** – further upload of new and re-accessing of existing mementos in *MemoryLane.*

8.6.3 Measures

The following data were collected:

- Importance rating for each memento;
- Memento type uploaded;
- Memento upload frequency in each session;
- Memento location within a given context, e.g. in the Home whether the memento was located in the living room or the bedroom;
- Memento narrative type and length of the narrative;
- Memento emotional association;
- Reminiscence affect at capture and retrieval;
- Comparative ratings of MemoryLane and Windows File Navigation for preference, performance, ease of use and engagement; and
- Subjective understanding of time-frames used for choosing mementos from.

8.6.4 Identifying Factors that Determine Memento Properties

As I hypothesised, mementos that were uploaded more frequently, were judged as more important, they were more likely to have associated narrative, and to engender a strong emotional association. This section presents the results from all the sessions and shows what, how and why users captured, uploaded and organised their mementos in each context.

Importance of Mementos

I hypothesised in H1, that different types of mementos will have different levels of importance. In the **Reconstruction** stage, mementos in the People area had significantly higher importance ratings than Home ($t=6.3$, $df=28$, $p<0.0001$) or Locations ($t=3.7$, $df=28$, $p<0.002$). Home mementos were also related more highly than Communications ($t=2.2$, $df=19$, $p<0.04$), see Table 8.3. This is interesting as Communications seem to be more people-oriented. One possible explanation as to why Communications received low ratings was hard to capture Communications, e.g. it involved replicating digital communications and generating digital representations of physical.

In the **Long-term Exploration** stage there were no significant differences between Importance ratings in any of the MemoryLane areas, as also illustrated in Table 8.1. This could be due to a 'ceiling effect' where the number of uploaded items was low, so there were few differences, thus there were no effects. As Table 8.2 illustrates, people seem to have got the most important memorabilia into their *MemoryLane* during the *first* session which may explain why they did not add many more items in the *second* session.

Memento Importance (max=5)/Session	**Home**	**Locations**	**People**	**Communications**
Reconstruction stage	3.0	3.2	4.1	2.3
Long-term exploration stage	0.2	0.1	0.1	0.2
Total	3.2	3.3	4.2	2.5

Table 8.1: Mean importance ratings in *MemoryLane*

One participant said that their choice of which memento to include was purely based on the importance value of that memento: *"I chose my mementos based on their importance to me and what memory they bring to me when seeing them."* Another participant concurred: *"They were things that were fairly easily accessible (so I did not have to hunt around to dig things out). But I don't keep many mementos, so I only had a limited* *choice anyway. They were also things that had some meaning to me, not just random items with no attachment."*

In the Long-term **Exploration** session, it seems that, some participants changed strategy and decided to capture more of the essence of the changes that might be currently taking place in their life, for instance moving and adjusting to a new living situation, as one participant said: *"I decided to use cell phone pictures since I have virtually no space left on my hard drive to upload 'real' pictures. I have taken several pictures on my new phone (partly because it was a novel thing to do.) I chose a few pictures that capture the essence of a new experience moving to a new place. Also I haven't had much time to take pictures or go around because I am still focused on moving in and getting adjusted, so there weren't too many to choose from."*

Which memento types are uploaded most frequently?

To understand which area was more popular across all sessions, I counted the number of mementos that people uploaded. I then averaged the number of objects in each area per participant and used paired t tests to investigate significant differences between each area of *MemoryLane.* Even though people were told to capture 5 instances of each item, there was variability in the numbers chosen.

At the Reconstruction stage, the number of mementos uploaded in the Home area was significantly greater than in the Locations area ($t=2.7$, $df=30$, $p<0.01$), People area ($t=4.6$, $df=30$, $p<0.0001$) or Communications area ($t=4.2$, $df=30$, $p<0.0001$), see table 8.2. One participant gave their reasons for collecting more Home, rather than other mementos: "*The Home area, definitely. Locations and People get used less, although if I used MemoryLane for recording holidays etc then I would use locations much more. The People area is useful for recording chance encounters with interesting people who you might not necessarily see again but unfortunately I haven't had many of those!*" Similarly another participant said: *"My favourite area was the Locations part. I liked how* *you could geographically locate places on a map of the UK. I also liked the objects part [Home area], because you could place items in their relevant areas within a house."*

Furthermore, during **Reconstruction**, only 10% of Home mementos resembled what can be described as daily habits i.e. remembering to iron a shirt for work. These types of mementos increased to 23% during **Long-term Exploration,** which suggests that participants started recording more mundane everyday tasks they had to carry out more, resembling Lifelogging behaviour.

Memento Count/Session	**Home**	**Locations**	**People**	**Communications**
Reconstruction stage	5.5 *(83%)*	3.6 *(68%)*	2.9 *(52%)*	2.6 *(93%)*
Long-term exploration stage	1.1 *(17%)*	1.7 *(32%)*	2.7 *(48%)*	0.2 *(7%)*
Total	6.6 *(100%)*	5.3 *(100%)*	5.6 *(100%)*	2.8 *(100%)*

Table 8.2: Average Upload Frequency in *MemoryLane*

For the **Long-term Exploration** stage there was a shift in how people used *MemoryLane.* There were significantly fewer uploads than observed in the **Reconstruction** stage even though people had more time to prepare for the **Long-term Exploration** stage (1 month). In the **Long-term Exploration** stage, the Home area again contained significantly more mementos than the Communications area ($t=2.8$, $df=30$, $p<0.01$). Similarly, Locations also contained significantly more uploads than Communications ($t=3$, $df=30$, $p<0.006$). But overall, Long-term exploration stage proved to be most popular for the People area with significantly more mementos being uploaded than in the Home ($t=2.6$, $df=30$, $p<0.01$) or Communications ($t=4.3$, $df=30$, $p<0.0001$) areas. There were no significant differences between Home and Locations areas ($t=-1.4$, $df=30$, $p>0.1$), see Table 8.1.

Also, as table 8.1 shows, the People area contains the key mementos in *MemoryLane* in the long-term, while the novelty of digitising Home, Locations and Communications items wears off over time. The People area remains strong because links with other people seem to be stronger than memories of events associated with objects or places. As our design suggests, it might be that instead of digitising objects, locations or communications that remind the user of a person, it might be better to digitise that person, and build conceptual or physical links to the objects and places relating to people. One user said: *"I liked uploading people onto the MemoryLane and I have a lot of the pictures of people. But I would have like to link them to the locations that these pictures were taken"*.

As hypothesised in H2, and Table 8.1 shows, the Communications area was least popular. We know that communications come in many different forms: facebook, e-mail, text, paper letters, postcards and other. Perhaps people did not feel they really wanted to transfer their existing communications to *MemoryLane*, particularly those which were already in a digital form, which meant duplication of the item. It was clear that the mapping between the communication artefacts and their representations in the memory and *MemoryLane* was smooth. One user said: *"I did not see the point of a Communications area"*. In addition, a lot of the communication items already represented alternative links to that memory, for instance, they came from people. Future designs of such tools should focus on links between communications and people, which tend to be mutually inclusive.

The People area was very popular for grouping thumbnails of pictures to show complex relationships between people, just as in other studies (Whittaker, Jones et al. 2004). There was a tendency to group friends in one corner and family in the other. Sometimes people chose to put some of their thumbnails outside the frame (particularly when there were arguments and disagreements). This user behaviour was primarily evident during the **Long-term Exploration** stage.

During the **Long-term Exploration** phase there were a lot of people who took pictures of their daily events or of artefacts used in their daily routines. Several participants said that they would like to use *MemoryLane* as a way of keeping a digital record of their everyday snapshots: *"I will be adding all the photos and videos I've collected in the last month because I tend to take about 1-3 photos per week plus about 1 video per week."* This is also evident from the artefacts representing everyday routines, such as ironing or using a potty, as illustrated in Fig. 8.8. This type of collection suggests a shift from more static "Who I am" starting memorabilia collected during **Capture** stage, to more of a "Life-logging" (i.e. a role for tools that automatically capture everyday events or habits) behaviour where participants capture more mundane everyday event based experiences during the **Long-term Exploration** stage.

Next I analysed which areas of the house people uploaded their mementos to more frequently. Consistent with (Petrelli, Whittaker et al. 2008)'s findings and H3, people tend to include more mementos in the public display space in their home, primarily their living room. However more personal memorabilia were allocated to the bedroom - which was the second in popularity for memento placement. As anticipated, attics, halls and bathrooms were the least popular. Memorabilia kept in the attic are not accessed daily, and attics tended to contain mementos from the more distant past and often hidden in boxes. As I expected, hall and bathrooms were not the prevalent places for keeping mementos in general. As a frequently used public space, I anticipated the kitchen to be popular, which was not the case. This perhaps could be due to the kitchen holding objects that are far too mundane. Table 8.3 illustrates these results.

#LivingRoom	#Bedroom	#Kitchen	#Garden	#Attic	#Hall	#Bathroom	#Total In House
32%	29%	14%	12%	5%	4%	4%	100%

Table 8.3: Overall memento distribution in Home Area in *MemoryLane.*

Memento Narratives

It was expected that people would find it easier to generate speech based narratives for their mementos. As hypothesised in H4, participants generated *speech* narratives during the **Capture** stage. However, with the exception of 2 participants there were no *speech* based narratives included in the Long-term **Exploration** stage, as shown in Table 8.4. There was a switch from using speech to using text for narrating mementos. During the collection, 75% of users captured speech narrative and during the **Reconstruction** stage, 25% of users added textual narrative. However, during the **Long-term exploration** stage only 3% of users added speech based narrative, but 97% added text based narrative, as illustrated in Table 8.4 below.

Narration Count/Session	**Home**		**Locations**		**People**		**Communications**	
	Speech	*Text*	*Speech*	*Text*	*Speech*	*Text*	*Speech*	*Text*
Reconstruction stage	4.8	1.5	3.2	1.2	2.7	0.9	2.9	1.0
Long-term exploration stage	0.0	0.7	0.1	1.0	0.0	1.2	0.0	0.0
Cumulative Total	**4.8**	**2.2**	**3.3**	**2.2**	**2.7**	**2.1**	**2.9**	**1.0**

Table 8.4: Average number of Narratives by type collected in *MemoryLane*

Table 8.4 shows that participants often included both types of narration across all areas.

Narrative & Memento/Area	**Home**	**Locations**	**People**	**Communications**
#Combined Narrative	7.0	5.5	4.8	3.9
#Mementos	6.6	5.3	5.6	2.8

Table 8.5: Average number of combined narratives in *MemoryLane*

Table 8.5 illustrates that overall, 60% of all mementos in the Home area had both types of narrative: *speech* and text. Similarly, 37% of Locations and 39% of Communications were annotated with both types of narration. However in the

People area, there were fewer annotations than actual mementos. Not all items had an associated narrative. Synchronous annotation containing both text and *speech* narrations were not popular in People area.

Text stories tended to be shorter and more brief. Speech based stories were longer and included ambient sounds and often captured speaker's mood and variation in voice. An example of a memento provided by one participant is shown in Figure 8.12 and this is what they said about it: *"OK…ah…this is a T-shirt that I got when in Moscow…and it is part of my collection of T-shirts in random languages…I am recording this, because I am not sure how long this lovely T-shirt is going to last…it's getting a bit shrunken and faded…so will probably have to donate it pretty soon…it's just a nice memory of the trip that I went on with my mother last year"*. It's clear that participant was trying to remember where they got this T-shirt and the reasons why they wanted to capture it. You can hear utterances and the tone of voice, which sounded happy and somewhat nostalgic.

Figure 8.12: Memento narrated with speech in *MemoryLane.*

Another participant wrote this textual narrative for memento in Figure 8.13: *"This is a chocolate man in a chocolate shop in Devon. I took this picture on holiday, I love chocolate so this shop was a real treat, but I went at the end of the holiday so didn't have enough money to get more chocolate!"*. This form of narration is briefer and often excludes any contextual information.

Figure 8.13: Memento narrated with text in *MemoryLane.*

So why it was that people did not use much speech to narrate during **Long-term Exploration** stage? When I asked people about the lack of long-term popularity of speech annotation, they said the following: 1) They did not like to narrate stories in certain contexts, e.g. for Locations as it meant standing in the street and talking into the Dictaphone, which felt odd. On the other hand, they saw definite advantages to speech, e.g. they felt they might forget what they wanted to say if they postponed their narrative until later when they were back at their homes; 2) Participants found it cumbersome to capture *speech* on their mobile phones and then transfer it onto their *MemoryLane*; and 3) Participants found it awkward and embarrassing to listen to their own voice.

One participant said they did not like to narrate their own stories: *"No speech – I don't like listening to my voice. Typed explanations – yes, but mostly done for the benefit of explaining the pictures to other people who might view the [MemoryLane]".* Another participant said: *"I liked some speech on the photos to remind me of the context to the photo which helped me remember better."* Thus, people generally liked having the narratives, but they found it hard to generate them. Participant comments show the value of speech based narrative helps remember better.

In contrast, the majority 97% of participants made use of textual annotation feature in *MemoryLane* in the **Long-term exploration** stage, as one participant explained: *"I would have liked to added speech, as it would have been able to capture my feelings/ emotions at the time of the memorable event. I did add [textual] narrative to all the* *mementos - this was really helpful as it enabled me to describe the memento and why it was important or memorable at the time. It's also good when looking back at mementos as you can re-live the moment when reading the narrative."* There were also editing problem in capturing speech based narrative: *"... I feel more comfortable with notes anyway – if you make a mistake typing something it's much easier to fix than when recording speech."* Overall, participants interpreted speech as a more permanent and final type of record than text. *Speech* seemed to be more complete, and it is perceived as being harder to erase or change when compared to text.

People were strategic in the ways they captured narrative for *MemoryLane,* focusing on what they thought they might forget: *"I tried to record things which I wouldn't necessarily remember without MemoryLane in the long term – so generally people who I see every day I didn't add, but little objects which I might not have in the future I did. I also added things which have a bit of a story attached to them and detailed that with them (e.g. the Festival of Britain coin) so if my own memory fails me I will always have the MemoryLane and stories to remind me."*

Emotional Memento Associations

Consistent with H5, I expected that people would be uploading mementos that had positive rather than negative associations. This turned out to be true with *MemoryLane* participants. Table 8.6 shows that *happy* mementos dominated in all areas of Memory across both sessions. However in addition to happiness, there was an emotion of *love* associated to the People area in the **Capture** and **Reconstruction** stages that was not present in the **Long-term** exploration stage. What could have caused this shift in emotional states? Perhaps people uploaded mementos of their loved ones for the **Reconstruction** stage, but because this applies to relatively few people they 'ran out' of them for the Long-term **Exploration** stage.

It was also interesting to find that the Locations area in the **Long-term Exploration** stage had an even distribution of emotion ratings for *happy* as well as *relaxed*. This could be because when people have time to go to other locations, they do so because these places make them feel *relaxed*, as the results in Table 8.6 show. However, there were some negative emotions for the Locations and Communications areas. This was primarily for mementos that had nostalgic associations, e.g. a last letter from grandmother or the last place visited with dad.

	Area	Love %	Relaxed %	Happy %	Other Neutral %	Other Negative %
Reconstruction stage	Home	15.40	21.30	**45.60**	9.60	8.10
	Locations	9.10	24.70	**42.90**	5.10	**18.20**
	People	**41.80**	2.90	**41.80**	6.20	7.30
	Communications	12.00	20.00	**48.00**	2.00	**18.00**
Long-term Exploration stage	Home	4.10	29.20	**54.20**	4.20	8.30
	Locations	0.00	**45.10**	**45.10**	3.80	6.00
	People	17.10	26.80	**46.40**	4.80	4.90
	Communications	14.40	7.10	**57.10**	21.40	0.00

Table 8.6: Emotional Associations in *MemoryLane.*

Participants found it interesting to think about the emotion that might be associated with their mementos, but as the results show most mementos were of a happy nature. One participant explains their memento decisions and emotional states: *"I decided on which mementos to upload based on events that I wanted to remember in the future. These were all positive, happy events which it would be nice to look back on in the future, especially the narrative that was added to the memento, which described the event and feelings/emotions attached to it. An example was my birthday, when we went to a medieval* *banquet in a castle and went in fancy dress - this is something we are unlikely to do again, therefore it's good to be able to reflect on the memory by looking at the photos and reading the narrative."* These findings are consistent with results for home photos which are overwhelmingly positive (King 1986; Chalfen 1997)

8.6.5 Benefits of MemoryLane

Time-frames for digital mementos

When participants were instructed to collect mementos related to their current life, I asked people to choose their own time-frame. They were allowed to choose any time-frame that related to their current life, e.g. since they started their new job, or since they moved to the new country. According to (Cohen 1996) people tend to organise their memories in lifetime periods, sometimes called "extendures" (such as schooldays, college days, working in Sheffield, etc.) Although I asked people to use a time-frame, I found that in the **Reconstruction** stage only 28% of all participants applied a structured time-frame organisation to their memento collection. The remaining participants picked their mementos based on importance. The evidence contradicts (Tversky 2004) in showing that importance, rather than timeframe, was crucial at deciding which mementos should be included and which excluded.

However, consistent with H6, during the **Long-term Exploration** stage there was a shift consistent with (Tversky 2004). 50% of participants used a personal time-frame to pick their mementos during their **Long-term Exploration** stage. During this stage people collected mementos that represented everyday activities, e.g. capturing the usual path to work. This resembled Lifelogging behaviours of recording the mundane, unlike the **Capture** stage where participants captured mementos that represented their identity and static relationships.

<u>*Reminiscing at capture or retrieval?*</u>

I looked at the reminiscence effect at both stages: (a) Capture; and (b) Reconstruction. I anticipated that people would reminisce more at the Capture stage, as they were re-telling the story about their past events, and while doing this, they would experience a reminiscence effect. Contrary to our expectations in H6, 84% of participants reminisced during the Capture stage, but 87% reminisced at the Reconstruction stage, suggesting that people equally reminisce during both: capture and retrieval.

8.6.6 Overall Ratings

<u>*Subjective Comparisons: File System vs MemoryLane*</u>

At the end of the **Capture** and **Reconstruction** stages all participants were asked to subjectively evaluate their experiences of using *MemoryLane* and compare them with Windows File Navigation system, illustrated in Figure 8.14, which they were currently using for organising their digital picture and audio mementos.

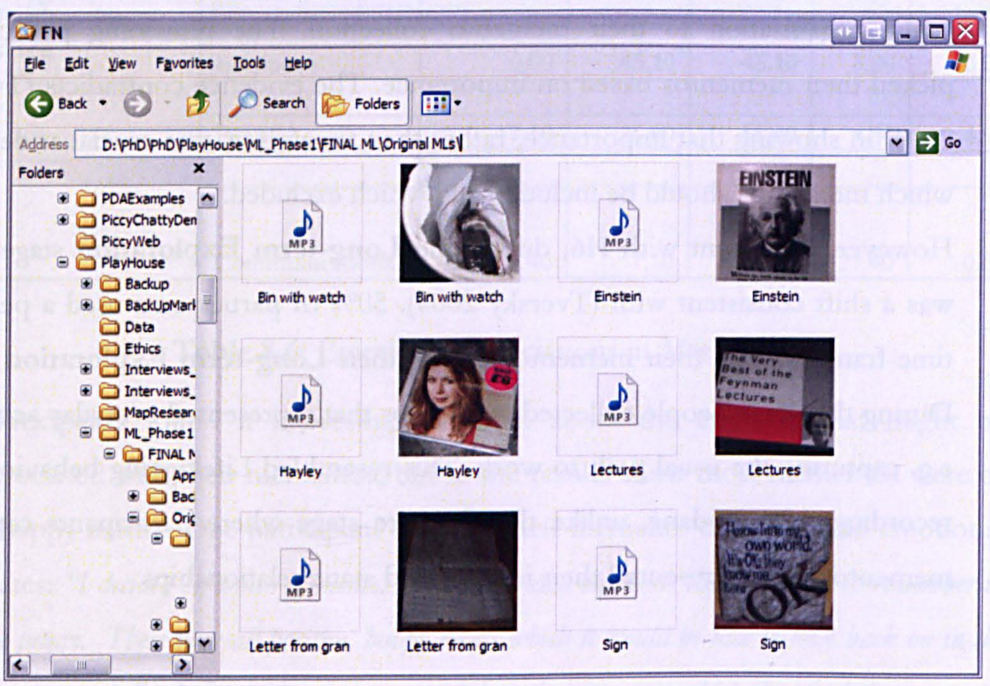


Figure 8.14: Windows File Navigation System.

Users were asked to judge which tool: (a) *MemoryLane* or (b) File Navigation tool was the most suitable with re-collection of memories. 94% ($t=9.4$, $df=30$, $p<0.0001$) of users said that they preferred *MemoryLane* as a way for viewing their mementos. Participants were then asked to subjectively rate the following on the 5 point Likert scale: 1) *Ease of Use*: how easy was it to use each tool for uploading and managing digital mementos? 2) *Engagement*: how engaging were each of those tools for arranging personal mementos?

The subjective scoring results revealed a significant difference in scores between *MemoryLane* and File Navigation tools. In particular, a significant number of participants said that *MemoryLane* was *Easier to Use* than File Navigation ($t=6.8$, $df=30$, $p<0.0001$). In terms of which system was more *Engaging*, significantly more people said that MemoryLane helped them to engage more in managing their digital mementos ($t=12.6$, $d=30$, $p<0.0001$). These differences are illustrated in Figure 8.15.

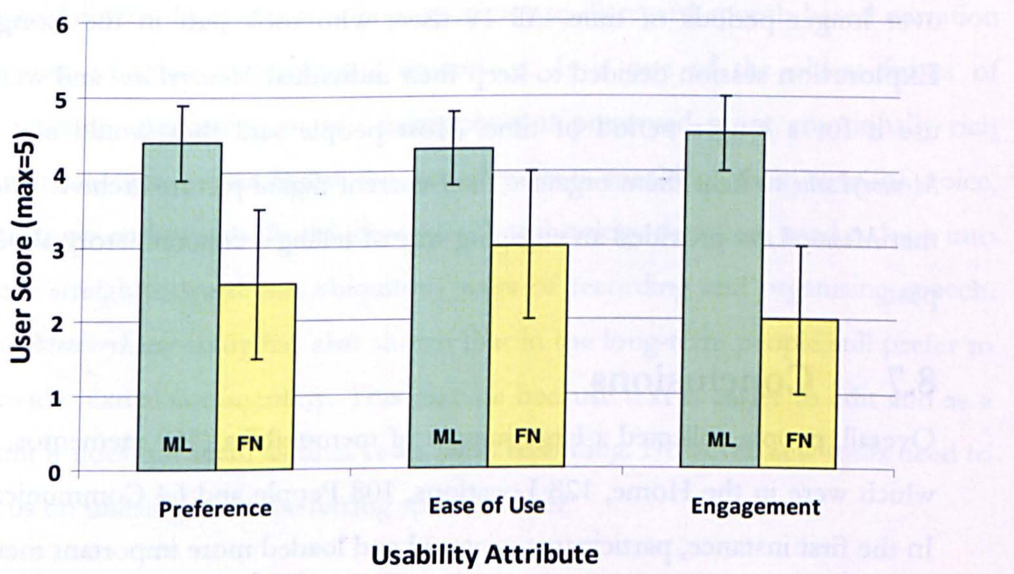


Figure 8.15: Subjective ratings for *MemoryLane* (*ML*) and Windows File Navigation system (FN).

Most participants enjoyed using MemoryLane and they said that they could see using it for their longer term memento archiving. As one participant said: "*I would definitely like to use [MemoryLane] for a longer period of time. If we ever sell our house and move, it would be great to add mementos of our life in the current house (well, the good times anyway!). It would be nice to look back on these in time, after we'd moved house. It would also be good to add speech to mementos to capture my thoughts, feelings and emotions attached to them. I think this would be useful when looking back at a memento that maybe happened some time ago and had been forgotten. Speech would bring back the memory straight away.*"

Another participant voiced their thought: "*Overall, I think MemoryLane is a great idea and really easy to use*".

8.6.7 Long-term MemoryLane Trials

To understand how *MemoryLane* could help people contextualise and preserve their physical mementos, it was essential to give the tool to participants to use over longer periods of time. All 19 users who took part in the **Long-term Exploration** session decided to keep their individual *MemoryLane* and wanted to use it for a longer period of time. Most people said they would like to use *MemoryLane* to help them organise their current digital picture archive. They said that *MemoryLane* provided an engaging way of telling a coherent story about their past.

8.7 Conclusions

Overall, people collected a large sample of memorabilia (460 mementos, 160 of which were in the Home, 128 Locations, 108 People and 64 Communications). In the first instance, participants captured and loaded more important mementos to them which may have led to fewer, less important mementos being loaded at later sessions. Initially there was tendency to capture mementos that were of static nature that represented more permanent relationships, e.g. the identity of a person in their home. However, subsequent sessions showed Lifelogging behaviours, where participants were capturing mundane everyday events.

With the proliferation of devices that allow people to capture and digitise their physical environment, it is assumed that because the technology available, people will use it. The capture stage of the study shows that although people have the technology, it is not always convenient to use. People seem to feel that some technologies, particularly concerning the capture of speech narrative are cumbersome. Walking in a busy street and talking into a Dictaphone, you not only making a fool of yourself, but you might not feel comfortable enough to tell an accurate and detailed narrative, jeopardising the coherence of a story (Tversky 2004).

Another reason for the unpopularity of *speech*-based narrative could be that it is hard to manage when it has been captured. These recordings are often large and require conversion to a conventional format or physically transferring it off a mobile device. However it is important not to disregard speech based narration because of these technological limitations. It is one of the oldest forms of storytelling. According to our participants it preserved more emotionally rich data than text – you heard the actual person speaking, their tone of voice, utterances and moods. In this case speech technology designers need to look into more straightforward and ubiquitous ways of recording and organising speech. The *MemoryLane* study has also shown that in the long-term people still prefer to provide textual commentary. This may be because text is easier to edit and as a result it does not seem as final as an *speech* recording. However new tools need to focus on utilising and transferring speech better.

Context has a crucial place in telling a fuller and richer story of the past. Tools such as MyLifeBits (Gemmell, Bell et al. 2006) have explored capture tools for mementos and other mnemonic information. However understanding how people explore this context rich data is still lacking. Our study of *MemoryLane* has shown that because of time constraints and convenience, people will first capture their immediate environment – artefacts in their home. However, over longer periods of time people prefer to capture other people and from this build further associations with objects, locations and communications.

Consistent with previous research (Chalfen 1997; Whittaker, Bergman et al. 2009), people's captured events were mainly positive, in particular stronger emotions, such as "love" were associated with people. There were some negative emotions in the Communications and Locations areas, but these were more of a nostalgic nature.

I observed that during the first session, participants tended to capture mementos from multiple periods of their lives (contradicting Tversky). However in subsequent sessions, people used more structured approach which led to capturing more recent events and mundane everyday experiences.

Our explorations about the reminiscence effect show that participants reminisce equally at capture and retrieval.

Finally, participants expressed a preference for the new system as a better way of organising, managing and contextualising personal digital mementos.

8.8 Chapter Summary

This chapter illustrates that there is a long-term potential benefit in digitising physical mementos. I built a viable system that was popular with users as a way to capture and organise digital memorabilia. The study has demonstrated that people choose mementos based on their importance and convenience. They also tend to primarily choose mementos that evoke a *"happy"* memory. People still prefer to generate textual rather than speech based narratives to describe their mementos. Overall, they want to digitise and preserve their narrated physical mementos in a context for future perusal and recall.

Chapter 9

Contributions, Conclusions and Future Work

Building user-centric Prosthetic Memory and Lifelogging tools, by its very nature, requires integration of research in several disciplines, in particular Human Computer Interaction (HCI), the Psychology of Organic Memory (OM) and Prosthetic Memory (PM) and tool implementation from Computer Science. In this thesis, I have built and evaluated novel PM tools to support memory in *three* domains: for conversations, for learning and for remembering personal memorabilia. Through a series of systematic lab and field-based user evaluations, I have shown: (1) a synergistic relationship between PM and OM and the principles that govern PM use; (2) how and why PM tools are used in the long-term in settings where people have a genuine need to remember; and (3) how to design tools for capturing personal autobiographical memorabilia.

As observed in the Literature Review in Chapter 2 of this thesis, there has previously been: 1) a prevailing focus on technology driven research and rather less empirical or theoretical exploration of how this technology relates to OM; 2) few long term usage studies exploring long-term effects of these technologies on OM; and 3) few tools designed for capturing autobiographical memorabilia. The main contributions of this thesis are to address these gaps, by specifically looking at: 1) what factors determine the usage of PM tools through controlled laboratory studies and in field trials with real users; 2) what long-term benefit PM tools can provide; and 3) implementing and evaluating better tools for supporting capture and retrieval of autobiographical memorabilia.

9.1 *What* determines the usage of PM tools?

People naturally exchange a lot of important information through speech and most of this requires some form of capture and retrieval. In Chapters 4 and 5, I conducted a controlled laboratory based user study, where I investigated *how* and *when* people use PM, as opposed to their OM, to help them remember everyday conversations. For this I recruited 25 volunteers for a within-subject study. Participants were given everyday conversational stories to listen to in a set of capture and recall conditions at three different Testing Sections (same day, week later and month later). A combination of qualitative and quantitative set of results showed that:

1) Users rely on PM for retrieval when they feel least confident in their own OM;

2) Veridical PM tools provide more accuracy than OM at longer Testing Sections;

3) There is a preference for *efficient* rather than *entirely accurate* PM devices.

4) Use of PM tools helps focus *attention* at capture, even when PM tools are not used later;

5) User performance increases when they use PM tools strategically, e.g. by taking appropriate notes;

6) Cues such as notes help memory in two distinct ways. They serve as retrieval cues, but the act of constructing cues (e.g. notes) help users encode and hence remember information.

The underlying theoretical principles learnt from this short-term study were firstly that PM and OM work in *synergy* with each other and future PM tools should focus on designs that support this principle. Rather than aiming to replace OM, PM tools should focus on supporting OM. We should develop tools that better support situations where users' OM is poor. This could be by focusing on areas where normal people are liable to forget (e.g. prospective memory), or by targeting specific populations who suffer from particular memory deficit (e.g. Alzheimer's – see (Berry, Kapur et al. 2006)) Second, in this type of application we should focus on the *efficiency* of PM retrieval over exhaustive, accurate capture. At the moment researchers tend to assume that people are purely oriented to accurate recall, but my work shows there are clearly situations where efficient approximate recall is enough. Future work needs to more precisely identify these situations, as well to develop more efficient tools for accessing memory cues.

9.2 *What* are the Long-term Benefits of Using PM tools and *Why*?

Learning is a situation that places large demands on memory. Students have to remember a lot of new material in situations where complex information is being constantly presented to them. To explore whether students' OM could be supported with new multimedia based PM tools, I extended the tool I had developed for capturing everyday conversations to lecture settings. I did this to investigate long-term usage of PM tools - to examine *how and why* students would use them in real-life settings. I implemented *two* new novel interfaces that allow users to access and re-listen to lecture presentations. I conducted a long-term field trial in a real lecture theatre environment, creating a tool that was used actively by 54 out of 98 student participants. I also conducted an additional controlled laboratory study. Both studies used within subject experimental design. The results showed the following:

1) Students looked at important segments of a lecture rather than listening to the whole lecture;

2) Participants were able to utilise other people's notes and pictures to access relevant parts of a lecture;

3) Non-native students used these tools more, but native students who used these PM tools more obtained better scores for their coursework;

4) Students accessed notes more than pictures. I believe this is because notes offered more fine-grained retrieval indices;

5) Attendees who used PM tools did better than non-attendees, although non-attendees accessed PM tools more. Again this result indicates that PM tools are unlikely to totally replace OM;

I then extended these classroom based PM tools to develop a web 2.0 inspired social summarisation tool that incorporates social feedback - where notes or picture indices become more salient as they are accessed more frequently by users. This is a new form of index tagging: one benefit over most web 2.0 applications it does not require any effort from the user, except to access the items of interest. Instead of tagging with text; users provide feedback by clicking on handwritten strokes and pictures. Following the introduction of these new PM tools, there was a demonstrable shift from using the old (non-social) PM tools described in the previous paragraph, to the new social web 2.0 based PM tool, even when both were available. To investigate the benefits for students of social summarisation, I conducted a field trial which involved logging participant access to social summaries over a period of weeks giving them a questionnaire at the end. 25 students actively used these new social PM tools, even though they still had access to the previous (non-social) PM tools. The results show that access to social tools over long-term improves student coursework scores, more than individual PM tools. In addition, I again observed greater use of notes vs. photo indices presumably because the finer granularity of notes makes them more effective retrieval indices.

9.3 Better PM tools for Supporting Autobiographical Memory

Unlike memories in other domains, autobiographical memories are more personal, emotional and often hold more personal significance. It is well known that people expend a great deal of effort in creating collections of personal memorabilia. The literature review showed a relative lack of PM tools to facilitate the capture and organisation of autobiographical memorabilia. In order to investigate the types of mementos people want to capture, I designed and implemented a novel PM tool for uploading and organising digitised representations of physical autobiographical mementos. I then conducted a field trial with 32 participants over 2 time intervals (2 weeks and 1 month after being given the tool) to investigate the benefits of this new PM tool. The main focus of this study was to investigate *what* types of mementos participants choose to digitise and *why*. This field trial was design focused. I encouraged users to critique the new PM tool they were evaluating and suggest new directions. The results revealed unusual patterns for annotating personal autobiographical mementos with complex views on capturing audio narratives. The overall results were positive, with users uploading large collections of mementos during the first session and fewer during the subsequent sessions, indicating preference for important mementos. Participants also uploaded more mementos expressing associations with other people and relationships. In general, the chosen mementos evoked positive emotion with stronger emotional ties in the People area.

Overall, the user feedback to the benefits of being able to organise autobiographical memories in contextualised and more engaging ways.

9.4 System – Data - Theory Conclusion

In this thesis I have shown that there is a strong relationship between OM and PM while highlighting a theoretical and design principle that it is essential to have OM and PM work in *synergy* with each other. This has important implications for approaches such as Lifelogging which tend to be focussed on exhaustive "capture" of all experience, rather than areas where users may have systematically poor OM. The thesis also points to the importance of metamemory in the use and deployment of PM tools: people tend to use PM tools only when they believe that they will otherwise forget.

This thesis has contributed a long-term naturalistic study investigating how and when people use PM as opposed to their OM. The results show overall long-term effects of using PM, as well as *when* and *why* people use this type of tool. Again, we see that actually recording will not replace OM, as absentees never performed as well as students who attended lectures. Furthermore, the typical use of the system was for short queries as opposed to accessing an entire lecture – again suggesting that people want to incrementally improve what they already remember instead of consulting an entire complete recording.

Finally, this thesis has contributed a novel PM tool for organising autobiographical memento data, in a domain where there is a shortage of such tools. Evaluation results show the advantage of having a digital application to store captured personal memorabilia. Further the study indicated that people upload important memorabilia with strong emotional associations, with the focus being on People and the Home areas. The study also revealed some of the limits of speech for annotation.

9.5 Future Work

With the proliferation of ubiquitous, mobile and wearable devices, we now have potential access to huge amounts of personal data; therefore we need to design effective user-centric ways of capturing, organising and retrieving memories from a variety of everyday events. Future work should continue investigating domains where such data will be potentially beneficial, and propose new PM and Lifelogging tool designs. That work should focus in the following areas.

9.5.1 Lifelogging and Wearable Tools

Chapter 4 looked at capturing digital records using manually operated cameras and Dictaphones. Chapter 8 looked at long-term capture of personal memorabilia with mobile phones. However, the proliferation of wearable Lifelogging tools such as SenseCam (Hodges, Williams et al. 2006), raises some interesting research questions. With exception of (Sellen, Fogg et al. 2007) there has been little work in evaluating the effect of these tools on long-term memory recall. Such evaluations show there are potential advantages to using Lifelogging technologies, as everything is being recorded without user intervention. There is no longer a need to decide what to capture and prepare for it. This unintentional capture also allows people to participate more in events. However, an obvious disadvantage is that we will generate vast amounts of data, when there are few effective current methods for accessing these large archives. We need to further investigate how people capture, organise and re-access their long-term Lifelog data as well as new access techniques. One way of doing this might be to augment passively captured visual data with other types of meta-data such as GPS or even bio-data such as heart-rate. This would allow people to cluster and organise data based on location or other metadata. Another potential downside of lifelogging is that people may remember less: in Chapter 6 we showed that people remembered better (with or without cues) when they actively constructed such cues. There is a danger that we may not pay attention to our environment in the same way if we believe that everything is being recorded for us.

In addition, many people consider this type of technology to be an invasion of their private life. Over longer periods of exposure to this technology, people habituate to it as it becomes more integrated into their everyday daily routines, or if they can see clear benefits for it. Future work needs to explore trade-offs between privacy costs and the benefits accessing such data.

Such wearable and ubiquitous technology would also allow explorations of people's habits. It would not only show where people went and what they did each day, but it would allow them to view and analyse a set of days - highlighting habitual patterns and divergences from these patterns. This might allow people to reflect on behaviours to better understand themselves, or to have greater insight into how they might reduce undesirable behaviours (e.g. understanding what events precipitate smoking or overeating)

These types of investigations could lead people to new ways of using wearable and Lifelogging technologies that support current remembering practices and also help improve OM rather than seek to replace it or overwhelm users with a mass of low-level personal information.

9.5.2 Towards Tangible Embedded Memorabilia

One reason why photo albums will never go out of fashion may be that these are tangible objects that are embedded in our homes that are associated with common social practices. They are easy to pick up, they never run out of batteries, and they are easy to share and show to others.

Future research in PM and Lifelogging should explore these familiar practices and focus on building new tools that have these tangible properties. For example, we could further explore new, tangible ways of capturing, organising and browsing audio snippets. Although speech is crucial in conveying everyday information, it is hard to capture, organise, access or browse (Whittaker and Hirschberg 2003; Tucker and Whittaker 2006; Whittaker, Tucker et al. 2008). Speech is very expressive, conveying more emotion than any other form of communication (Dib and Kalnikaitė 2008). Audio based data not only contains speech based stories, but it can also include other forms of sounds such as ambient sounds (Frohlich, Kuchinsky et al. 2002; Oleksik, Frohlich et al. 2008). Although sound and story capture is important to memories of the past and remembering in general, there are only a few tools such as the Memory Box (Frohlich and Murphy 2000) and Living Memory Box (Stevens, Abowd et al. 2003) that focus on augmenting memories with ambient sounds and stories for retrieval. Furthermore, there are no PM tools that allow users to currently capture audio on their phones and seamlessly transfer to an archive, to help organise and re-access it later.

One new design might be for a PM tool that interacts with a mobile phone. This PM tool could take the form of a tangible artefact in one's home (such as a globe) and use Bluetooth technology to allow users to upload and physically interact with their audio snippets.

We expect that these types of tangible technologies would be less effort to use, fun, appropriate and engaging for managing audio based memorabilia.

9.5.3 Augmenting Memory through Social Summarisation and Web 2.0

From Chapter 7 it is evident that Social Summarisation was very useful for users who need to re-access large amounts of memory information, e.g. in an education setting. It is also evident that people can successfully exploit feedback generated by other people - confirming the quality and validity of such information. This raises interesting research questions about how and when people generate and exploit new web 2.0 technologies (Kalnikaite and Whittaker 2008).

Combining multimedia and existing tagging practices, we propose to further explore how social summarisation might be applied to new domains with richer media such as video using widely available YouTube services. With the proliferation of mobile telephones, more people tend to use their camera phones to capture moments of their lives on video, e.g. scenes and songs from a rock concert or a jazz festival. Although these videos snippets are initially perceived as a personal event record, they might form a different point of view be seen as a mass capture event. People have also begun to upload such short videos snippets onto shared video sites such as YouTube. I propose, instead of having separate video snippets of an event shared publically on YouTube (as is currently the case), we could derive techniques of analysing the scenes that people upload more frequently, rate higher or access more often and derive new types of social summaries that are based on any, or all of these criteria. This approach would allow people to infer which parts of the video were valued most by other users and provide social highlights of these shared events. For example, we already have built a successful application that exploits users uploading behaviours to YouTube identifying video highlights from the most uploaded parts of a film (San Pedro, Kalnikaite et al. 2009).

9.5.4 Empirical Evaluations investigating *What* Type of Records People Keep of Their Past

Before building even more PM and Lifelogging tools, we need to better understand what type of records people tend to keep of their past. This would allow us to build tools that better fit people's archiving practices. There are a few studies (Petrelli, Whittaker et al. 2008) investigating what physical mementos people keep in their homes and this suggests that there is a need for some form of tangible interaction with physical mementos. Another study (Petrelli, van den Hoven et al. 2009) investigated what people would like to keep for the future in a form of a time capsule - showing that people tend to focus on symbolic objects, but make few attempts to annotate these in preparation for future retrieval. Other studies into how people re-find their digital memorabilia (Whittaker, Bergman et al. 2009) have also shown that people tend to keep a lot of digital photos but forget where they stored when asked to find them. Part of the problem is that people seldom access these archives, so they are unaware of their existence, and their lack of organisation. This also happens with physical mementos too, since organising analogue pictures is time consuming; we tend to simply place them in a shoe box (Frohlich, Kuchinsky et al. 2002). However analogue mementos stored in a shoe box accumulate at a far slower rate than digital photos and so end up being easier to manage and access. Again we need to explore new ways of making mementos tangible and exploiting users' current practices to better understand how such interactions are incorporated into users' everyday lives.

These studies have focused away from technology in order to understand current practices. Technology nevertheless dominates our everyday lives and there is greater shift from physical to digital.

9.5.5 Mapping Between the Physical and Digital

Studies investigating personal memorabilia collections e.g. (Petrelli, Whittaker et al. 2008; Petrelli, van den Hoven et al. 2009) have provided further understanding into what people capture in terms of their personal memorabilia and why. This knowledge of personal archiving practices can inspire novel PM tool designs that work in synergy with OM. However, there is a shortage of PM tools that effectively support this mapping between physical and digital, which seems to be one of the most important properties for collecting, organising and re-accessing autobiographical mementos.

In future studies, I propose to investigate new PM tool designs that facilitate this mapping between physical and digital. In particular, PM and Lifelogging tools that help integrate current digital practices, such as mobile communications with current archiving practices of autobiographical mementos - taking images or capturing audio and integrating it with GPS data, calendar data or e-mail archives need to be automatically populated with easily transferable links between data associate with contacts, places, communications, multimedia or temporal data (e.g. (Ringel, Cutrell et al. 2003)).

This recommendation differs from 9.5.2 in the sense that it replaces or replicates physical with digital.

9.5.6 Applying Memory Theories to the study of PM tools

The need for synergy between OM and PM is clearly motivated by our empirical results. However, this is only one possible design requirement. Psychology research for decades has investigated how people remember and forget and this knowledge has not been exploited in the design of effective PM tools. I have shown that people not only remember using their OM, but they rely on metamemory to help them find information stored in an external PM tool. Our memories undergo immense challenges on a daily basis in terms of the amounts of information that is being presented to people which is expected to be recalled one way or another. Such examples are not only in the most memory demanding areas such as learning, but access other areas of our everyday lives.

Thus I propose a further exploitation of psychological memory theories and data for the design of such new everyday PM tools. Current tools tend to address recall of predominantly factual information. We need new tools that support different memory phenomena such as social remembering, reminiscence, and reflection, and psychological work can help us better design tools to support these.

Appendix A

Resources relating to the work in Chapters 4 and 5

- Stories used in ChittyChatty experiment and follow-up
- Task questions used in ChittyChatty experiment – same day, 7 days later and 30 days later.
- ChittyChatty evaluation PM review questionnaire
- ChittyChatty evaluation score sheet

Pilot Story – "Loud Neighbour"

Hi! Did you hear that music last night? It was blasting till about 2 o'clock in the morning. By the time I got to sleep, I could hear birds singing outside. That neighbour downstairs, Kelly makes me crazy! I guess you don't hear much of it while living opposite.

The other night I wrote her a note explaining my situation. I didn't get the reply until 3 days later. She thanked me for giving some guidelines on the time that I go to bed and the time I have to be up in the morning. Kelly's reply was not much of an apology that I was expecting. It had a promising tone thought. She assured me that she will stop playing her music loud after 11.30 at night. She claims that this is the only time she gets to practice her piano.

To be honest, I think she is good on the piano. I could listen to her piano all night. Most of the time, it is not the piano that keeps me awake. It's the cheesy 80's tunes that she is so obsessed with. Yup, it is almost at the obsession stage with some songs being repeated 5 or even 10 times. Last night, she discovered Red Hot Chilli Peppers, and that was the only reason why I didn't go to interrupt her.

Think she might be a music teacher specialising in 80's music.

I think I'll leave it for now. Hope that she will stick to her promise and forget her 80's cheese for a bit longer. But if it doesn't calm down, I will have to pay a visit or invite her for a cup of tea and to listen to some music from downstairs.

Story 1 - "Mad Dave"

Oh, do you remember my older brother, Dave? Let me tell you how he got here. He has loved Def Leppard ever since he was 15 years old and saw them play at the Sheffield Show, Hillsborough Park in 1978. The hair, the tight trousers, the heavy guitars, the thunder of the drums and the screaming vocals. He was particularly entranced with their Yorkshire lyrics.

To be honest, he was obsessed. They used to rehearse in some old warehouses and he would hang around outside listening to them tune their guitars. He found it entertaining. When their practice sessions were over, they'd catch their bus home and Dave would pretend he was getting the same bus. He would sit in front of them all the way to Infirmary Road and then get off and catch his own bus back to Heeley. The band gradually got annoyed with him being under their feet all the time and one of them gave Dave a right shouting at when he followed him into Record Collector. I remember Dave looking really embarrassed when he got back home. He told himself to keep a distance from now on.

In 1979 Def Leppard were one of the biggest rock bands in the country, but then a strange thing happened. A journalist for Sounds magazine wrote that the band had "sold out" to America. Dave wasn't sure what that meant. Like just about every other band, they wanted to be successful in America, but so what? It's not like they had cut their hair, but suits on and started singing mushy ballads. Most of their original fans believed this story and when they played a Reading festival in England, they showered them with bottles. It was another 7 years before their home country would ever really accept them again.

In that time they became incredibly popular. Their second album sold well in America and their third sold six million. I remember Dave cheering and hopping up and down in WH Smiths when he read the news in Kerrang magazine. Two girls reading Smash Hits looked at Dave's spotty face and greasy hair in disgust. But he didn't care. His band was bigger and better than Duran Duran or whatever rubbish was that they liked.

But heartache soon followed when Rick the drummer lost his arm. David visited the place on the Snake Pass where he crashed. Then he spent a whole week with his one arm tucked inside his T-shirt until Dad clipped him around the head and told not to be so stupid. He was 22 at the time so that left him feeling foolish for a bit.

Anyway success followed again at the end of eighties, followed by the inevitable decline. The albums began to lose their edge and when Steve, the guitarist died, Dave thought they would pack it all in. But they kept going, keeping the tour bus rolling, last night they came home to Sheffield to play the Arena, and as usual Dave was right at the front going crazy. Suddenly, Joe, the singer spotted Dave in the crowd. He'd recognised Dave after all those years, thought obviously he was a bit fatter and his long hair was thinning a lot. To Dave's complete surprise Joe pulled him out of the crowd and introduced him to the whole arena as Def Leppard biggest fan. They nicknamed him "Mad Dave". Dave raised his arms into the air to bash in his glory and then dived forward back into the crowd. Obviously they didn't fancy a fat, balding and middle-aged rocker landing on their heads. So that's how he got here, Northern General hospital with crushed ribs, a fractured arm and a broken nose.

Story 2 – "Trip to Dublin"

We went to Dublin last summer. A friend of mine at work, Patrick, is from a small town just outside Dublin and he invited us over to stay for a long weekend. We'd never been to Ireland before and didn't know what to expect. I was surprised about how easy it was to get there. We drove over the Snake Pass and arrived at Manchester Airport in just over an hour.

Another friend at work, Ben, who we travelled with, had sorted out the parking and we were checking in only 90 minutes after leaving home. Patrick was a seasoned Manchester to Dublin traveller and he showed us how to avoid the long queue for checking your hand luggage and everything. I hate airports so I was glad to get through so quickly.

We landed in Dublin two hours later just to realised that our bags have not been flown with us in the same place. Instead they will be arriving a day later. But we weren't too worried about that, it was only a long weekend, and they have shops in Dublin.

It was raining and my first impression was that Dublin looked pretty much like any other city in Britain, grey, full of the same sort of shops and the same vacant looks on the shopper's faces. We stayed at University Accommodation. It was the end of term so the students had gone home and the University was letting rooms out fairly cheaply, though it was about 2 miles away from the centre.

By the time we'd collected our keys and dropped our bags off, it was dark outside. Patrick had left us not long after landing in order to catch up with some of his friends and had arranged to meet us in the centre at half past seven. We had planned on catching a bus from the University but it was getting late and we weren't sure which was the right bus to catch, so we opted for a taxi. Those Dublin taxi drivers certainly like to talk. The man who drove us in was really friendly though, not like in Sheffield. By the end of the journey we knew everything you could possibly want to know about what Ireland was like 10 years ago and how much it has changed with all the EU money. They were all very excited about the new state of the art swimming pool that was being built in the centre of Dublin.

When we arrived at the pub where Patrick had arranged to meet us, he was already on his third pint of Guinness. The place was buzzing because England were about to play Argentina in the World Cup and a big screen at the back of the pub was showing the game. I went to the bar and ordered three pints of Guinness. You've got to drink Guinness when you're in Ireland. Patrick told me it was nothing to do with the water from the river, as some people claim, just care and attention in how it is kept and transported to the taps.

England scored and about four people in the whole place cheered. I kept my celebrations to myself, not wanting to return home with a black eye and watched the rest of the game unfold nervously. We almost held on right to the end and then Argentina scored. The whole pub erupted as if Robbie Keane had just put the Republic of Ireland into the finals. I looked at my half pint of Guinness in despair. Well at least we had a draw I thought. But no, Sol Campbell did that mistake in almost the last minute and we lost. Now the whole pub was dancing a jig. I hoped the rest of the holiday wasn't going to be like this.

Story 3 – "Barrington's Life"

Oh, do you remember Barrington Vera-Smith from the sixth form? I ran into him the other day and he was telling me he lived in Naseby now, which is just a few miles south of Harrogate, in a beautiful two bedroomed cottage with his wife, Jenny. They've just had their tenth wedding anniversary, which they spent in Cornwall.

He had never been to Cornwall before. The beautiful scenery and clear waters left an impression on Barrington. I have never thought of him as a traveller, but he was saying that he would like to go travelling the world next year, health permitting. He hasn't decided where to go and what to see yet, but he and Jenny felt that they could do with some more time off together. Their current responsibilities would make it hard to go away for longer than 2 weeks, but Barrington was positive about that. I think he had a plan, or at least it seemed so.

However, the day that I met Barrington, he was organising an evening dinner party for his friends. He was looking for Jenny's favourite wine, which unfortunately was all sold out on that day. "It's a complete disgrace" he told me. He also asked to speak to the manager of the wine merchants and since he wasn't in, he said he would be writing a strong letter to the chief executive later that day. I sincerely hope it didn't ruin the whole night for him.

Barrington and Jenny are nationally recognised dog breeders. Their dogs, Richard and Susan, are their treasures, Barrington told me. I'm sure you've heard of cocker spaniels and I'm certain you know what a poodle is. Well, Richard and

Susan are a cross between the two. They are cocker-poos. Some people seem to think this is a funny name, but not Barrington and I had to try really hard not to laugh. Barrington was telling me for almost an hour about his dog breeding business and how Richard and Susan are widely respected as professional show dogs. So far they have won three 1st prizes in the toy category and a couple of best in shows. Barrington admitted it was a struggle to get to this standard.

Apparently, the dog grooming always falls on Jenny. Barrington takes responsibility in booking the shows and organising the transport. He also does all the driving which is almost every day. Barrington boasted about the money his business was making. He has been approached by the local college to run a few workshops in starting-up a successful business. He was keen on accepting the opportunity.

Anyhow, Barrington is five years older than Jenny and he told me he was working for the RSPCA when they met. Jenny then had just finished her A levels and was considering a biology degree at a university. It sounded like it was definitely love at first sight. With a little help from Jenny's family they decided to make a career in breeding.

I mentioned Barrington about my sister's new baby. He looked a bit uncomfortable and hinted that Jenny's family drop the odd hints now and then about when they plan to have their own children. When I asked, Barrington gave me a funny look. He told me that Richard and Susan are their children. They require round the clock love and attention. They just wouldn't have the time for anything else.

Our conversation had to end there unfortunately as Barrington still had a lot of preparation to do for his evening party, including buying a Spiderman's outfit for himself. To be honest, I was glad to get away.

Story 4 – "Dreadful Wedding"

I've got to tell you about the wedding I went to last Saturday. It was dreadful! Well, when I say I went to the wedding, I should explain first that I wasn't invited. Actually what happened was that Michelle was going to go with Callum but they had a big argument on Tuesday, which I'll tell you about later, and so Callum asked if I wanted to go instead of Michelle. Well, you know about me and Michelle, so I said yes, I'd love to.

The service was one of those, you know, alternative ceremonies. We arrived a few minutes late so had to sneak in at the back. The service had started earlier than expected, even before the arrival of the bride. Everyone was singing when we came in. What is that hymn called? Lord of the Dance, that's it. So the people at the front, they were really lost in the whole experience, waving their hands in the air, eyes closed, it was a bit spooky I thought. Then some of them started dancing. I couldn't believe it, I felt embarrassed just watching them. Then, after the song finished, the bridal march music started. But do you know how it's usually played on the church organ? Well, instead of the organ, they had this woman in a light brown flowery dress playing it on an accordion. Me and Callum were really trying hard not to laugh, but everyone else looked really serious. And then I saw the bride, well...what can I say?

As she walked in, everyone turned to look at her, as people always do when the bride enters. However, instead of a white gown as everyone expected to see, she was covered, I mean the whole of her body, in green body make-up and she was just basically wearing a green bikini. She also had this sort of head-dress made of leaves and twigs. So she comes down the aisles, following by these dancing children, dressed up just like her.

When she reached the other end of the aisle where the groom was waiting, he was actually dressed in his tuxedo, so they made quite a strange looking couple. They joined hands and started spinning around together. At the same time everyone started clapping. That's when we both realised that we were the only people not clapping, so we felt we had to join in. Next came the vows and to me that was the most surreal part of the whole ceremony.

In between the vows there were intervals where people stood up from their seats and started reading poetry in different languages. I remember one guy stood up and started reading Pushkin's poem from Eugene Onegin in Russian, which I found most bizarre, because we had to learn that at school.

Reception was like walking to a health food store. I like that kind of thing. Carrot patty and rice cakes were my favourite. I wasn't keen on trying out the globe of pickled gherkins or drinking sour milk as a starter. It was all walk around and pick your own bits of food that you fancy. The desert was the classic wedding cake, but instead of raisins there were bits of carrots.

The best man speech was the most strange. It was a play rather than a speech. With the help from a number of friends he acted out the bride and groom's university years. It was hilariously funny thought. Callum's character was also played by an old bald guy, which caused him to leave the room for that part.

I think this wedding has changed my opinion about the alternative weddings. Although it was unusual and strange, it was fun. In the end, I didn't know what to expect and I enjoyed that.

Task questions used in *ChittyChatty* experiment

Recall Questions – Same Day

Pilot Story – "Loud Neighbour"

1. What time did Kelly promise to stop playing her music loud?
2. What does the speaker think is Kelly's profession?
3. Did the speaker mind Kelly playing piano?
4. Did the speaker have any intentions of complaining to Kelly again?

Story 1 - "Mad Dave"

1. Which year did Def Leppard became one of the biggest rock bands in the country?
2. What was Dave's nickname?
3. how did the local fans feel about Def Leppard's success in America?
4. How did the crowd respond to Dave diving onto them?

Story 2 – "Trip to Dublin"

1. What time was the evening meeting with Patrick arranged for in the centre?
2. Who made the last minute mistake in the England v Argentina game?
3. Do taxi drivers in Dublin think that EU is a good thing?
4. How did the speaker feel about watching the football game?

Story 3 – "Barrington's Life"

1. How many first prizes have Richard and Susan won in the toy category?
2. What outfit did Barrington have to buy for this evening's party?
3. Doe Barrington feel confident about his knowledge on running a successful business?
4. How does Barrington feel about having his own children?

Story 4 – "Dreadful Wedding"

1. What instrument was bridal music played on?
2. What colour was the bride wearing?
3. What mood was everyone in at the beginning of the ceremony?

4. Were there any traditional wedding foods served at the reception?

Recall Questions 7 Days Later

Story 1 - "Mad Dave"

1. Which festival were Def Leppard showered with bottles?
2. What was the name of the guitarist who died?
3. How Dave feel about other bands ?
4. How Was Dave popular with the band during their revival tour?

Story 2 – "Trip to Dublin"

1. What kind of accommodation speaker did stay in?
2. Which countries played the football game that the speaker watched whilst in Dublin?
3. What did the speaker think of Dublin architecture?
4. Did the speaker think that Guinness in Ireland is better?

Story 3 – "Barrington's Life"

1. What breed were Richard and Susan?
2. What degree was Jenny thinking doing after she finished her A levels?
3. Does Barrington give the impression that dog breading is hard work?
4. How did the speaker feel about talking to Barrington?

Story 4 – "Dreadful Wedding"

1. Who's poem written in Russian could speaker remember?
2. Was the guy who played Callum's character instead of the best man's speech, bold?
3. Did the groom give an impression of being smartly dressed?
4. Did Callum enjoy the best man's play?

Recall Questions - 30 Days Later

Story 1 - "Mad Dave"

1. Which magazine were the two girls in WH Smiths reading?
2. Which hospital Dave was taken to?
3. What did Dave think about Def Leppard "Selling in" to America?
4. How did Dave feel about Rick, the drummer losing his arm?

Story 2 – "Trip to Dublin"

1. How far away from Dublin city centre was the student accommodation that the speaker was staying in?
2. What was the tournament that the speaker watched whilst in Dublin?
3. Does the speaker prefer Sheffield taxi drivers?
4. How did the spectators behave like in the pub where the speaker watched the football game?

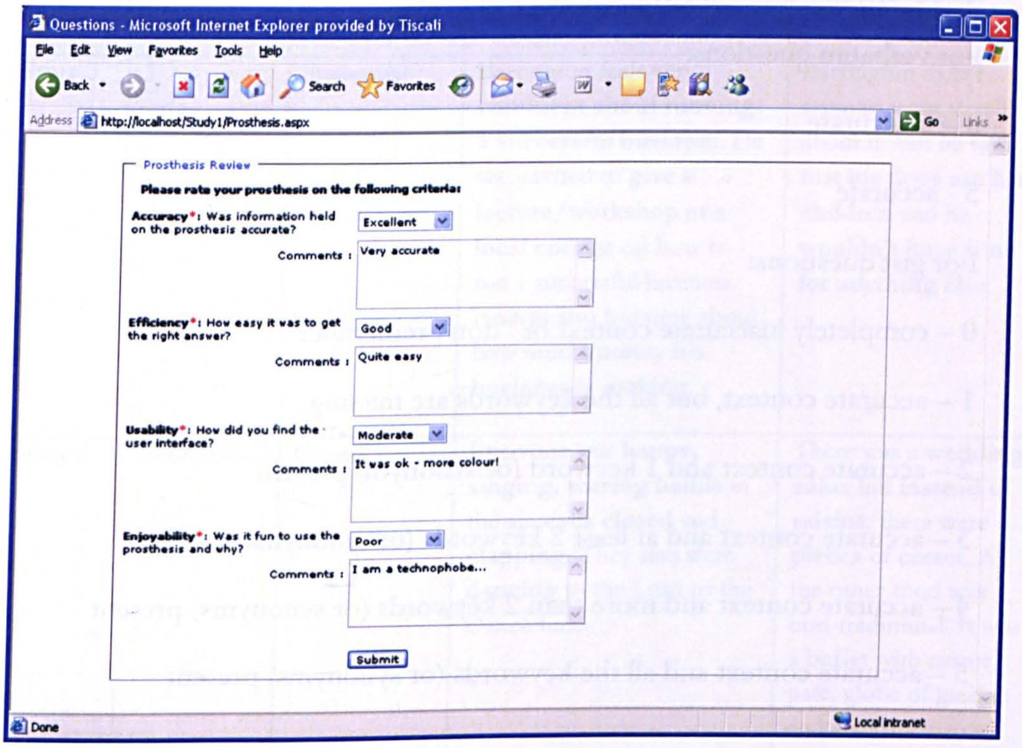
Story 3 – "Barrington's Life"

1. How many prizes of "best in show" does Richard and Susan have won?
2. Who was Barrington working for when he met Jenny?
3. Was Jenny's family keen on Barrington and Jenny becoming dog breeders?
4. Was Barrington ready for his party?

Story 4 – "Dreadful Wedding"

1. What was speaker's favourite dish at the reception?
2. What was in the wedding cake instead of raisins?
3. Did the speaker recognise any of the poetry read in between the vows?
4. Did speaker find the wedding fun?

ChittyChatty evaluation PM review questionnaire



ChittyChatty evaluation score sheet

Marking scale

For verbatim questions:

0 – inaccurate

5 - accurate

For gist questions:

0 – completely inaccurate context or "don't remember"

1 – accurate context, but all the keywords are missing

2 – accurate context and 1 keyword (or synonym) present

3 – accurate context and at least 2 keywords (or synonyms) present

4 – accurate context and more than 2 keywords (or synonyms) present

5 – accurate context and all the keywords (or synonyms) present

Answers – Same Day

	Answer 1	Answer 2	Answer 3	Answer 4
Story 1	1979	Mad Dave	They felt that the group **sold out** to America and most fans **didn't like them any more.** They also **showered them with bottles** at the Reading festival. But Dave didn't care, he still thought they were the best.	The crowed parted as they didn't fancy **middle aged balding** rocker **diving** on to them.
Story 2	7.30	Sol Cambell	The taxi drivers were chattier than Sheffield ones. They told how Ireland changed over the past 10 years. They thought that EU money was **good** as	The speaker felt a bit nervous as majority of the pub were supporting Argentina. **He was England's supporter**. The

			they were funding for the new state of the art swimming **pool** to be built in the **centre of Dublin.**	speaker **didn't was to show true feelings** as he didn't want to **end up with a black eye.**
Story 3	3	Spiderman	Barrington feels very **confident about running a successful business.** He was invited **to give a lecture/workshop at a local college** on how to run a successful business. He was also boasting about how much **money his business is making.**	Barrington didn't seem to want to talk about it. But he said that **his dogs are his children** and he **wouldn't have time for anything else.**
Story 4	accordion	Green	Everyone was **happy, singing, waving hands** in the air, **eyes closed** and **clapping.** They also were **dancing** to the Lord of the Dance tune.	There was a **wedding cake,** but **instead of raisins,** there were **pieces of carrot.** All the other food was non-traditional. It was a buffet with carrot pate, globe of guirkin and sour milk.

Answers - 7 Days Later

	Answer 1	Answer 2	Answer 3	Answer 4
Story 1	Reading	Steve	Dave **doesn't like other bands**. He thinks other bands such as **Duran Duran are rubbish**. He always liked and **followed** Def Lepard.	**Yes.** He was invited onto **the stage** and presented as Def Lepards **biggest fan.** The band remembered him and that made Dave feel good.
Story 2	Student Accommodation	England v Argentina	The speaker thought that the **Dublin looked like any other city**. It was **grey** with the **same**	**Yes** - The speaker thought that the Guinness in Ireland **tasted better**. It was **nothing to do with**

			shops.	**the water out of the river.** It was to do **with the way it was kept and transported.**
Story 3	Cocker-poos	Biology	**Yes** – Barrington was saying that it is **hard to get to the standard they currently are at.** He also said that it's **time consuming** and there is **no time for anything else.**	The speaker was **glad to get away.** He got **tired** of listening to **Barrington talking about his dog business.**
Story 4	Pushkin	Yes	**Yes,** he was wearing a **tuxedo**	The speaker though that the play was funny, but Callum **didn't enjoy** it as he **walked out** during the play. He **didn't find it funny** that his role was played by an old bald guy.

Answers - 30 Days Later

	Answer 1	Answer 2	Answer 3	Answer 4
Story 1	Smash hits	Northern General	**He didin't mind them being popular in America.** He was **happy** for them when he read that they were becoming famous in America and selling a lot of albums.	**Dave felt sad** about Dave loosing his arm. He **tucked his own arm in the t-shirt** until his dad clipped him round his head and told him not so be so **stupid.**
Story 2	2 miles	World Cup	**No. He liked the Dublin taxi drivers** because they were more **chatty** and **friendly**.	The spectators were **loud** and **erupted** when their supported Argentina team scored the goal. They didn't like English team winning.
Story 3	2	RSPCA	**Yes**. With **a bit of help** from Jenny's family, they were **able to set themselves a dog breeding business.**	**No**. He **still a lot of preparation to do** including **buying a spiderman's outfit** and buying jenny's favourite **wine**.
Story 4	Rice cakes and carrot pate	Carrots	**Yes**. It was **Pushkin's** Evgeny Onegin. **Russian poetry.**	At first he thought it was **strange** and unusual. But in the end he thought it **was fun** and **he would go to another alternative wedding in the future**.

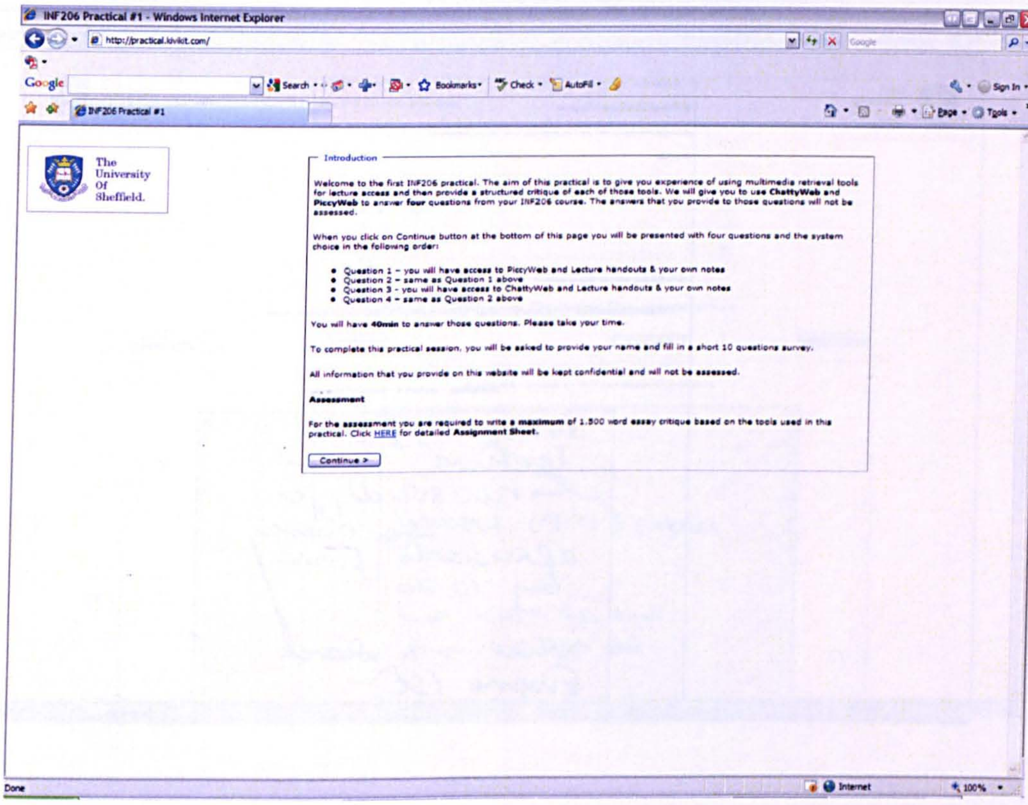
Appendix B

Resources relating to the work in Chapter 6

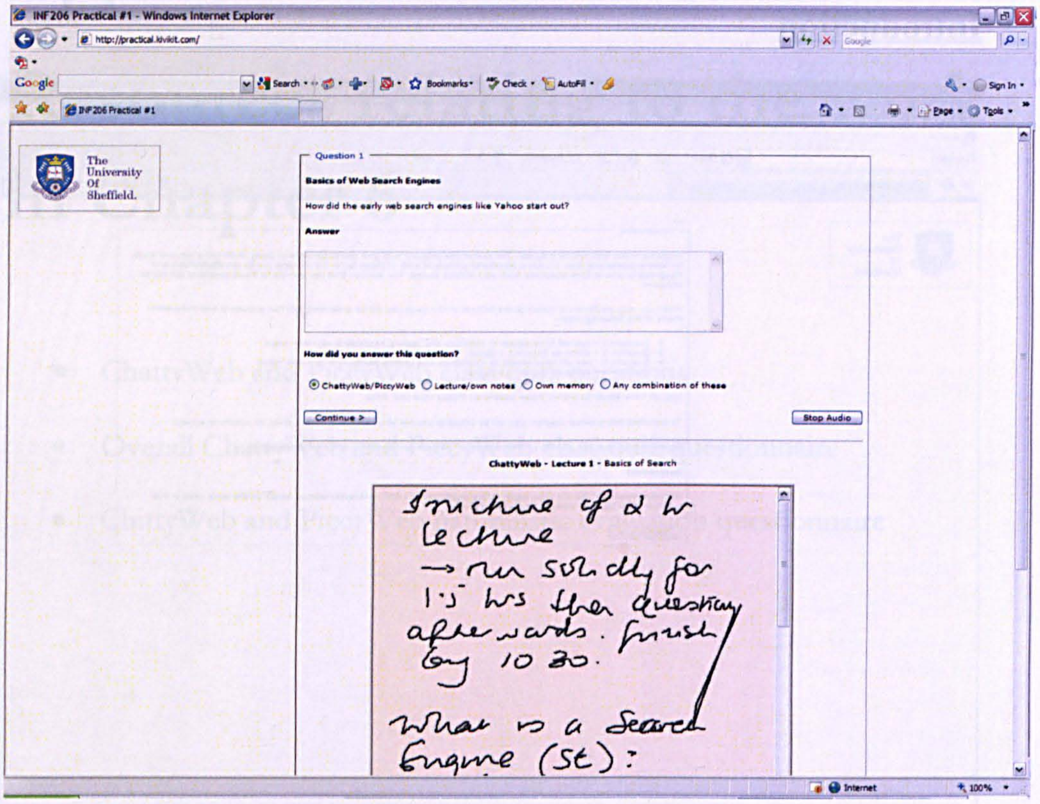
- ChattyWeb and PiccyWeb class quiz questions
- Overall ChattyWeb and PiccyWeb class quiz questionnaire
- ChittyWeb and PiccyWeb naturalistic evaluation questionnaire

ChattyWeb and *PiccyWeb* class quiz questions

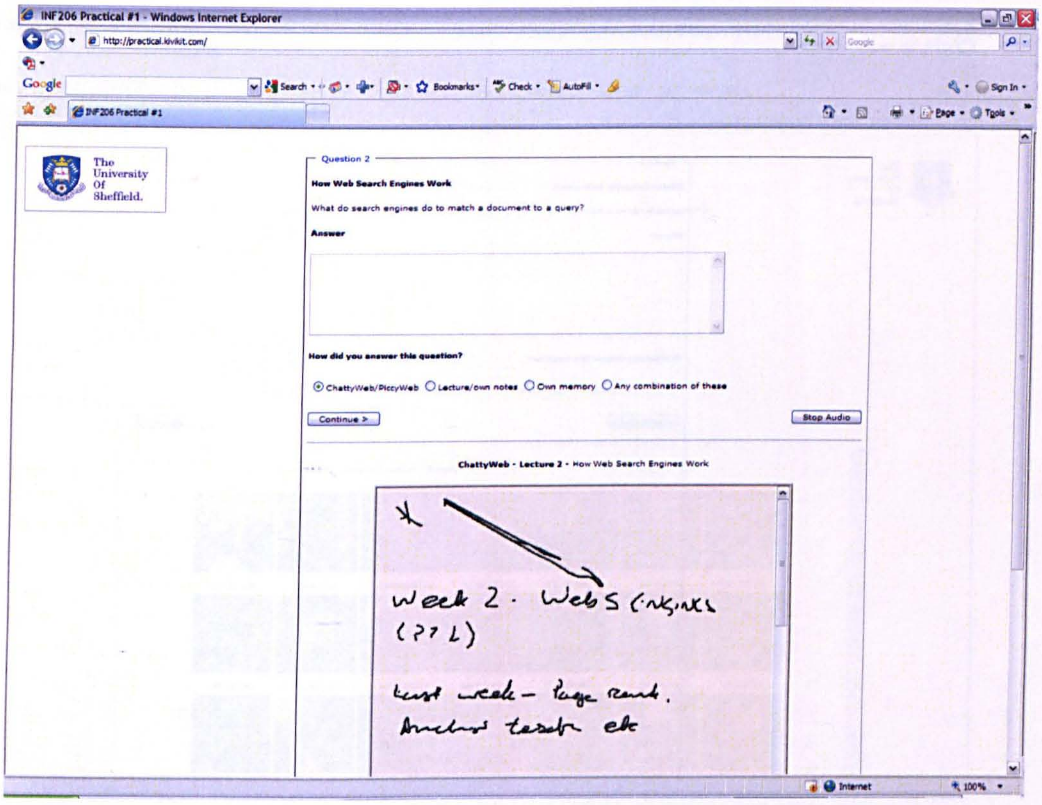
Introduction



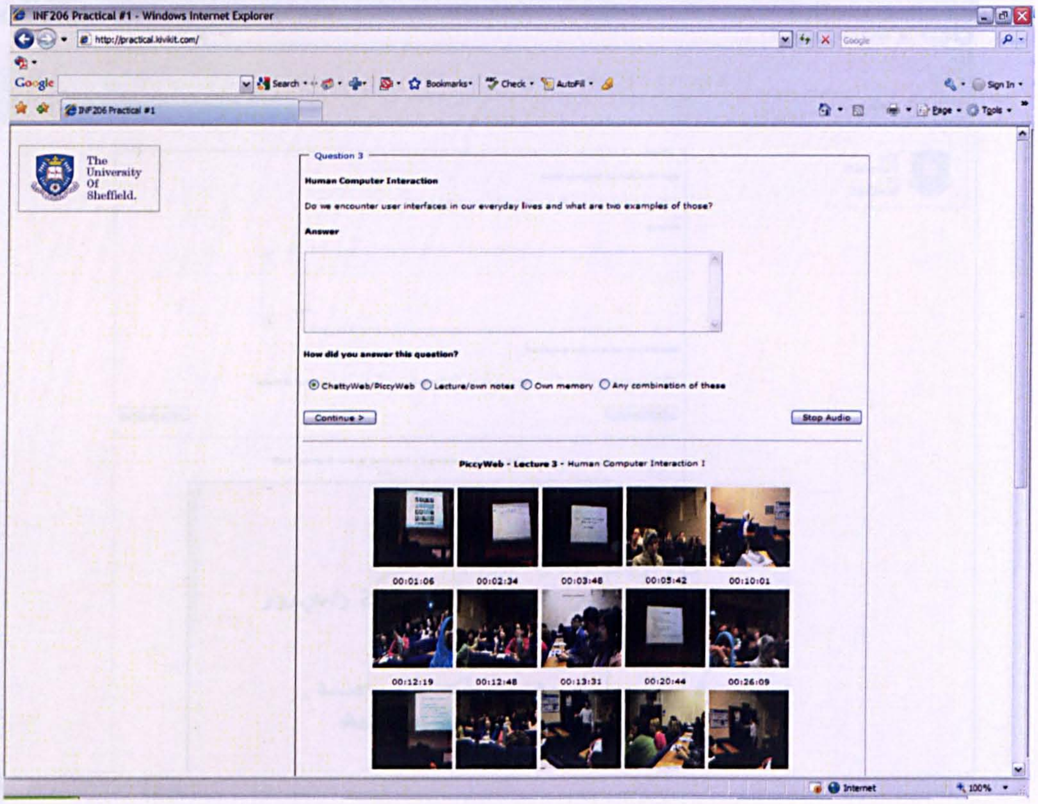
<u>**Quiz Questions 1 – using ChattyWeb**</u>



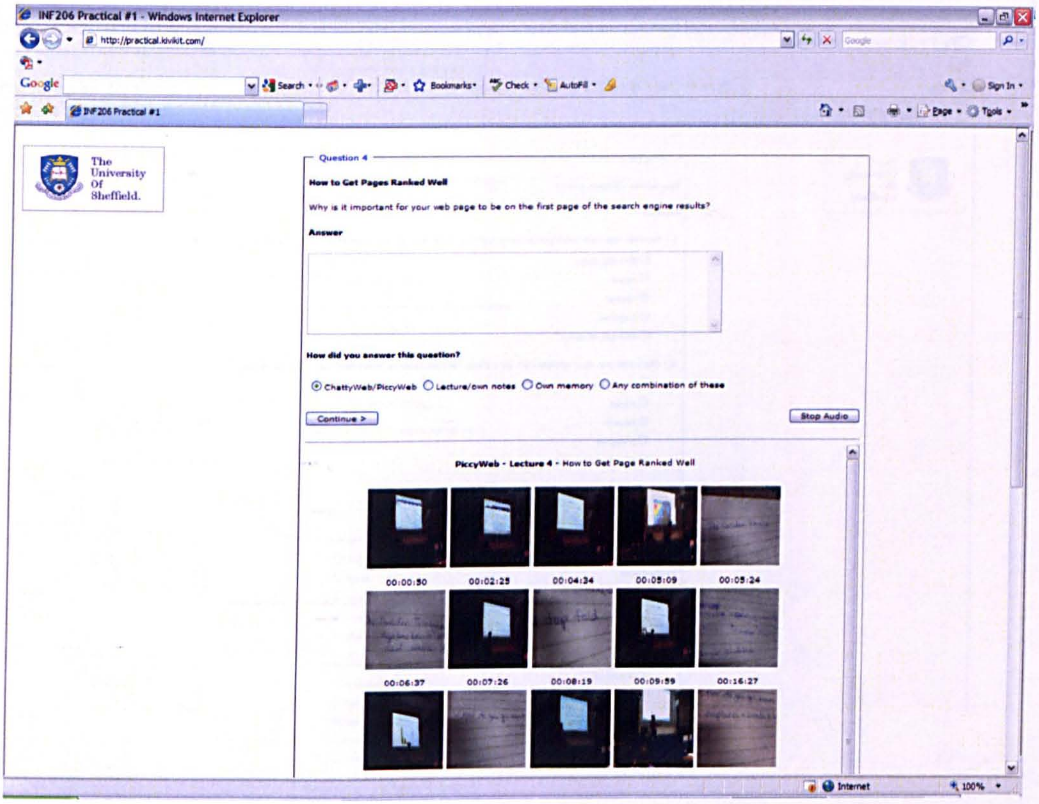
Quiz Questions 2 – using ChattyWeb



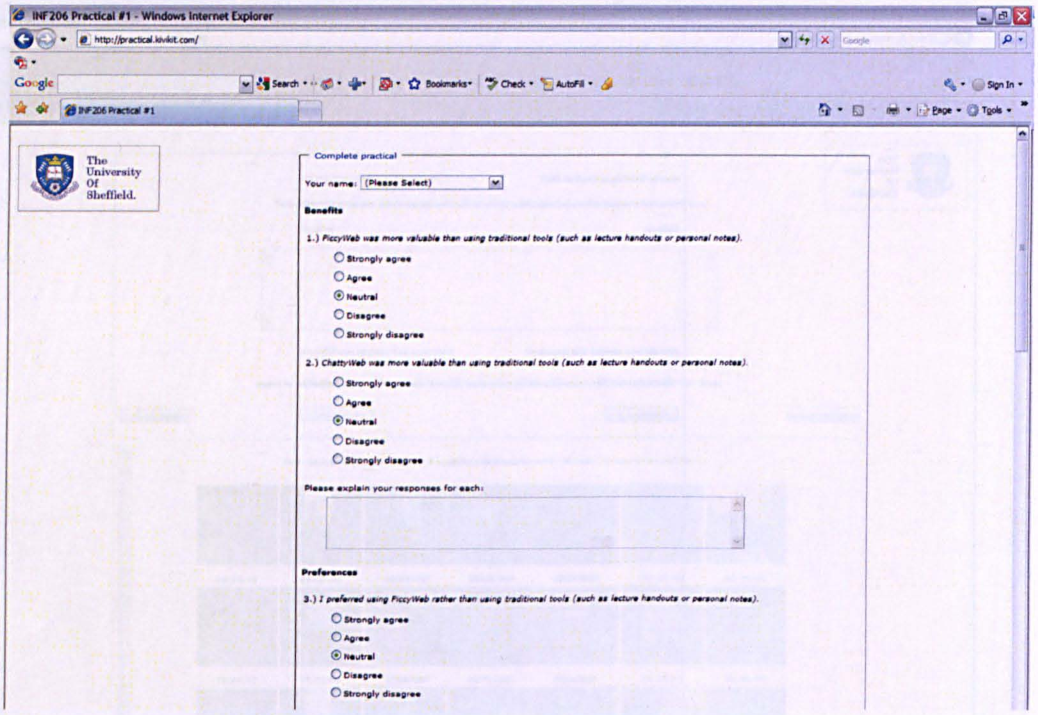
Quiz Questions 3 – using PiccyWeb

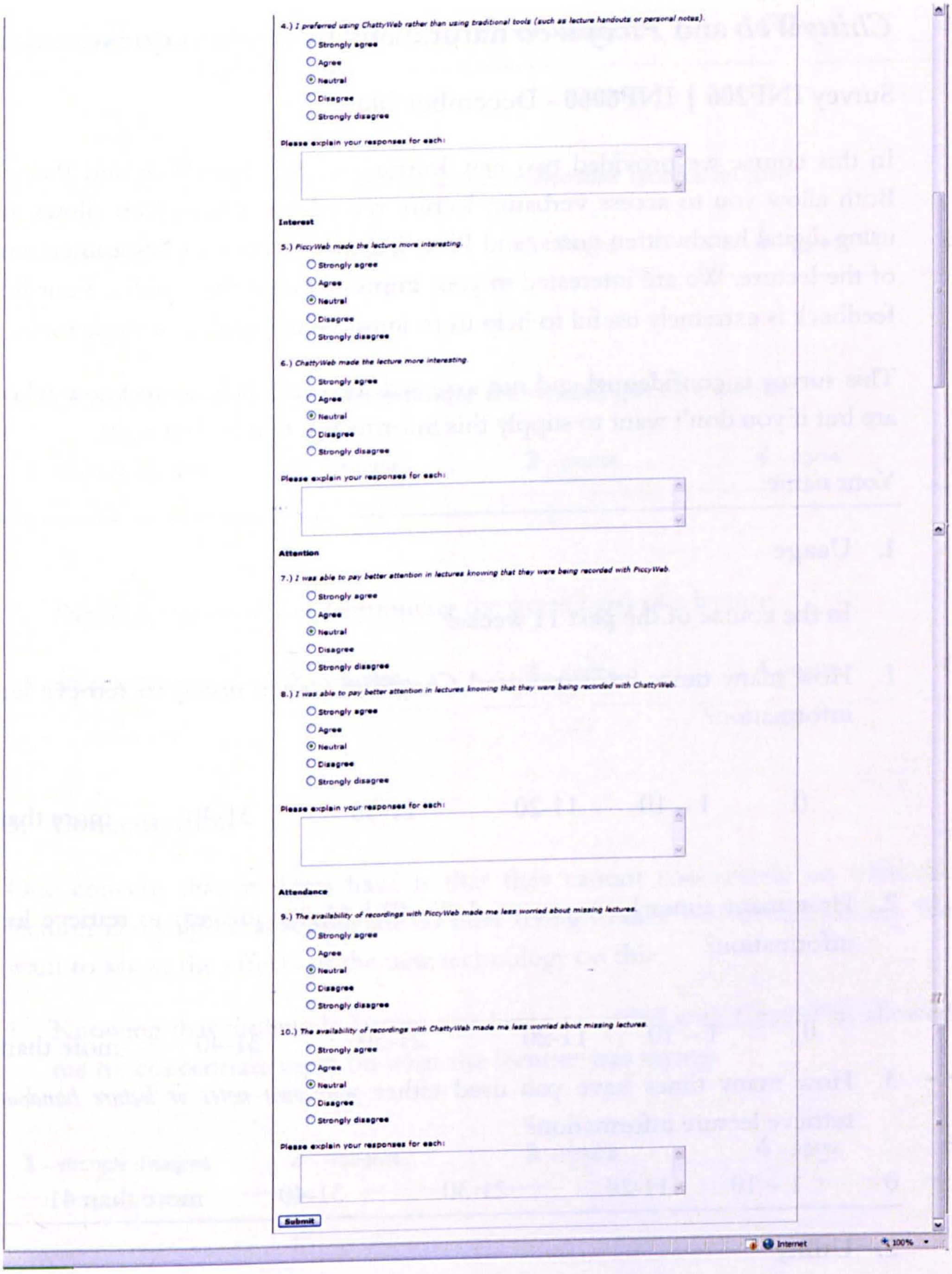


Quiz Questions 4 – using PiccyWeb



Overall *ChattyWeb* and *PiccyWeb* class quiz questionnaire




4.) I preferred using ChattyWeb rather than using traditional tools (such as lecture handouts or personal notes).

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Please explain your responses for each:

Interest

5.) PiccyWeb made the lecture more interesting.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

6.) ChattyWeb made the lecture more interesting.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Please explain your responses for each:

Attention

7.) I was able to pay better attention in lectures knowing that they were being recorded with PiccyWeb.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

8.) I was able to pay better attention in lectures knowing that they were being recorded with ChattyWeb.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Please explain your responses for each:

Attendance

9.) The availability of recordings with PiccyWeb made me less worried about missing lectures.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

10.) The availability of recordings with ChattyWeb made me less worried about missing lectures.

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Please explain your responses for each:

Submit

Internet 100%

ChittyWeb and *PiccyWeb* naturalistic evaluation questionnaire

Survey INF206 | INF6060 - December 2007

In this course we provided two new learning tools, ChattyWeb and PiccyWeb. Both allow you to access verbatim lecture recordings. ChattyWeb allows access using digital handwritten notes, and PiccyWeb allows access using camera images of the lecture. We are interested in your impressions of these tools. Your honest feedback is extremely useful to help us to improve the quality of these tools.

This survey is confidential and not assessed. It would help us to know who you are but if you don't want to supply this information that is your right.

Your name:

---

1. Usage

In the course of the past 11 weeks:

1. How many times have you used *ChattyWeb* (digital notes) to retrieve lecture information?

   0   1 – 10   11-20   21-30   31-40   more than 41

2. How many times have you used *PiccyWeb* (digital images) to retrieve lecture information?

   0   1 – 10   11-20   21-30   31-40   more than 41

3. How many times have you used either *your own notes or lecture handouts* to retrieve lecture information?

0   1 – 10   11-20   21-30   31-40   more than 41

---

2. Utility

1. *ChattyWeb* was good for accessing *specific information* from a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. *PiccyWeb* was good for accessing *specific information* from a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

3. *ChattyWeb* was good for determining the overall gist of a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

4. *PiccyWeb* was good for determining the overall gist of a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

3. Concentration

One concern that students have is that they cannot concentrate on what the lecturer is saying because they are so busy trying to take very detailed notes. We want to assess the effects of the new technology on this.

1. Knowing that the whole lecture was being recorded with *ChattyWeb*, allowed me to concentrate more on what the lecturer was saying:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. Knowing that the whole lecture was being recorded with *PiccyWeb*, allowed me to concentrate more on what the lecturer was saying:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. The *lecture handouts* and my own notes allowed me to concentrate more on what the lecturer was saying

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

4. Note-taking

1. I took fewer lecture notes knowing that *ChattyWeb* provided an *exact recording* of what the lecturer said:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

2. I took fewer lecture notes knowing that *PiccyWeb* provided an *exact recording* of what the lecturer said:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. I took a *different type* of lecture notes knowing that *ChattyWeb* and *PiccyWeb* provided an exact recording of what the lecturer said:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

5. Attendance

1. Knowing that the whole lecture was being recorded with *ChattyWeb*, made me less worried if I happened to miss a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. Knowing that the whole lecture was being recorded with *PiccyWeb*, made me less worried if I happened to miss a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

3. Having *lecture handouts* made me less worried if I happened to miss a lecture:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

Please tell us how many lectures you missed over the last 11 weeks:

____

6. Benefits

1. Having access to *ChattyWeb* helped me with my coursework:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. Having access to *PiccyWeb* helped me with my coursework:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. Having access to *my own notes and lecture handouts* helped me with my coursework:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

7. Fun

1. *ChattyWeb* makes the course more fun:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

2. *PiccyWeb* makes the course more fun:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. *Lecture handouts and taking own notes* makes the course more fun:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

8. Efficiency

1. Retrieving information from the lectures was very straightforward with *ChattyWeb:*

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. Retrieving information from the lectures was very straightforward with *PiccyWeb:*

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

3. Retrieving information from the lectures was very straightforward with *my own notes and lecture handouts:*

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

9. Information Provided

1. *ChattyWeb* allowed me to access information that that I would otherwise have missed:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

2. *PiccyWeb* allowed me to access information that that I would otherwise have missed:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. *Lecture slides/own notes* allowed me to access information that I would otherwise have missed:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

10. Accessibility

1. With *ChattyWeb* I could always easily access the recordings using the internet:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

2. With *PiccyWeb* I could always easily access the recordings using the internet:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. With *lecture handouts/own notes* I could always easily access information from the lectures:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

11. Information quantity

1. *ChattyWeb* provides me with exactly the *level* of detail I need for learning purposes:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

2. *PiccyWeb* provides me with exactly the *level* of detail I need for learning purposes:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

3. *Handouts and my own notes* provide me with exactly the *level* of detail I need for learning purposes:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

12. Information quality

1. *ChattyWeb* provides me with exactly the *type* of information that I need for learning purposes:

1 – strongly disagree	2 – disagree	3 – neutra	4 – agree	5 – strongly agree
☐	☐	☐	☐	☐

2. *PiccyWeb* provides me with exactly the *type* of information that I need for learning purposes:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

3. *Own notes/handouts* provide me with exactly the *type* of information that I need for learning purposes:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

13. Quantity vs. Attention

1. When I take more notes in a lecture, the less I am able to pay attention to what the lecturer is saying:

1 - strongly disagree	2 - disagree	3 - neutra	4 - agree	5 - strongly agree
☐	☐	☐	☐	☐

14. Your Choice

ChattyWeb, *PiccyWeb* and *own notes/handouts* provide different ways to access lectures. Which did you like the best?

ChattyWeb	PiccyWeb	Own Notes/Handouts
☐	☐	☐

Please say why:

If you did not use *ChattyWeb* or *PiccyWeb* at all please say why you didn't use them:

If you used *ChattyWeb* or *PiccyWeb* only *one or twice* please say why you didn't make more use of them:

Please state the *main* ways in which we might improve *ChattyWeb*:

Please state the *main* ways in which we might improve *PiccyWeb*:

Appendix C

Resources relating to the work in Chapter 8

- MemoryLane Reconstruction stage questionnaire
- MemoryLane Long-term Exploration stage questionnaire

MemoryLane Reconstruction stage questionnaire

CAPTURE

1. When did you choose your memories from? Please say why you chose this period?
2. How easy it was to decide what to capture and what to leave out? Please rate (1=very difficult to decide....5 very easy to decide)
3. Did you spend longer or shorter than 3 days to capture your mementos. Please say why.
4. Are there any mementos that you thought about, but decided not to include? Please say what it was and why you left it out?
5. Did you delete any mementos which you already recorded in the process? Please explain.
6. If you could capture more data more effortlessly, would you like to do that? E.g. if you had a device recording everything that you do and captures everything that surrounds you. Please explain.
7. What did you think of capturing different types of objects: house objects/outdoors/people/communications? Did you find some more difficult than others? Please explain.
8. Any preference as to which type of objects you would like to remember more/less? Please explain.
9. As you were capturing your memories and recording stories did you re-experience those memorable events, otherwise known as reminisced. Please say which memories caused reminiscence and why?

RETRIEVAL

1. Which representation: FN or ML did you prefer for viewing your autobiographical mementos? Please explain.
2. Please rate your preference for FN and ML (1 = lease preferred and 5 = most preferred)
3. Please rate the ease of use in both (1=very difficult and 5 = very easy)
4. Which system was more fun to use FN or ML. Please rate (1=least fun and 5=most fun)
5. Would you like to have more important mementos highlighted or emphasised? Please explain.
6. What other information you would like to associate to your mementos? E.g dates/people involved. Please explain.

7. Was it easy to rate the importance of your mementos? Please explain.
8. Was it easy to assign emotion to your mementos? Please explain.
9. I have noticed that you sometimes recorded textual narrative for your mementos. Why did you do this for certain things and why not for the others? Please explain. Was it useful? Please explain.
10. As you were looking at the things that you captured, did help you reminisce through those events? Please say which system evoked this feeling and explain the differences.
11. Can you see yourself arranging your mementos in the way presented in ML? Any design aspects of ML you would like to change?
12. When you perhaps move away and you will be surrounded by new things, would you come back to this and revisit your memories? Which system, FN or ML, would you choose for this? Please explain.

SHARING

1. Now that you have your mementos in ML, do you think you would like to share this with other people e.g. friends, parents? Please explain.
2. Would you want to *pass on* your ML onto your kids (if you have them) so they could see what your mementos were in the past.
3. Please rate the relevance of each system for sharing (5=very relevant and 1=least relevant).

MemoryLane Long-term Exploration stage questionnaire

USAGE

1. How many times did you use ML in the past month to upload your mementos?
2. How did you decide what mementos to upload?
3. How old are your mementos that you uploaded onto ML?
4. Which areas of ML were your most favourite (e.g. where you uploaded most pictures).
5. Did you include any audio narrative in your ML?

FUTURE

1. Would you like to use it for longer and add more mementos?
2. What other media you would like to add to your ML?
3. How else would you like to organise your mementos?

PROBLEMS

1. Did you encounter any problems while using ML?

DESIGN SUGGESTIONS

2. Do you have any design suggestions?

Bibliography

(2008). "Camtasia." from http://www.techsmith.com/camtasia.asp.

(2008). "Elluminate." from http://www.elluminate.com/.

(2008). "iLinc." from http://www.ilinc.com/.

(2008). "Wimba." from http://www.wimba.com/.

(2009). "CMU's speech recognition engines." from http://cmusphinx.sourceforge.net/html/cmusphinx.php.

Abowd, G., C. Atkeson, et al. (1997). "Cyberguide: A mobile context-aware tour guide." Wireless Networks **3**(5): 421-433.

Abowd, G. D., L. D. Harvel, et al. (2000). Building A Digital Library of Captured Educational Experiences. ICDL, Kyoto, Japan.

Abrams, D., R. Baecker, et al. (1998). Information archiving with bookmarks: personal Web space construction and organization. Proceedings of the SIGCHI conference on Human factors in computing systems. Los Angeles, California, United States, ACM Press/Addison-Wesley Publishing Co.

Aizawa, K., D. Tancharoen, et al. (2004). Efficient retrieval of life log based on context and content. Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences. New York, New York, USA, ACM.

Alea, N. and S. Bluck (2003). "Why are you telling me that? A Conceptual model of the social function of autobiographical memory." Memory **11**(2): 165-178.

Anderson, J. R. (2004). Congnitive Psychology and its Implications, Worth Publishers Inc.,U.S.; 6Rev Ed edition (9 Dec 2004).

Anderson, R., R. Anderson, et al. (2004). Experiences with a tablet PC based lecture presentation system in computer science courses. Proceedings of the 35th SIGCSE technical symposium on Computer science education. Norfolk, Virginia, USA, ACM.

Anderson, R. C. and J. W. Pichert (1978). "Recall of previously unrecallable information following a shift in perspective." Journal of Verbal Learning and Verbal Behaviour **17**: 1-12.

Austin, J. and P. F. Delaney (1998). "Protocol analysis as a tool for behavior analysis." Analysis of Verbal Behavior(15): 41-56.

Baddeley, A. (1998). Human Memory: Theory and Practice, Allyn & Bacon.

Baddeley, A. D. (1982). "Implications of Neuropsychological Evidence for Theories of Normal Memory." Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, The Neuropsychology of Cognitive Function **298**(1089): 59-72.

Baddeley, A. D., G. Hitch, et al. (1974). Working Memory. Psychology of Learning and Motivation, Academic Press. **Volume 8:** 47-89.

Bartlett, F. C. (1932). Remembering. Cambridge, Cambridge University Press.

Bellotti, V., B. Dalal, et al. (2004). What a to-do: studies of task management towards the design of a personal task list manager. Proceedings of the SIGCHI conference on Human factors in computing systems. Vienna, Austria, ACM.

Bellotti, V., N. Ducheneaut, et al. (2003). Taking email to task: the design and evaluation of a task management centered email tool. Proceedings of the SIGCHI conference on Human factors in computing systems. Ft. Lauderdale, Florida, USA, ACM.

Benton, S., K. Kiewra, et al. (1993). "Encoding and external-storage effects on writing responses." Journal of Educational Psychology **85**: 267-280.

Berntsen, D. (2002). "Tunnel memories for autobiographical events: Central details are remembered more frequently from shocking than from happy experiences." Memory & Cognition **30**(7): 1010-1020.

Berque, D., T. Bonebright, et al. (2004). Using pen-based computers across the computer science curriculum, ACM. **36:** 61-65.

Berque, D., D. Johnson, K. , et al. (2001). Teaching theory of computation using pen-based computers and an electronic whiteboard, ACM. **33:** 169-172.

Berry, E., N. Kapur, et al. (2006). "The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis." Neuropsychological Rehabilitation.

Beyer, H. and K. Holtzblatt (1998). Contextual Design: Defining Customer-Centered Systems. SAn Francisco, San Francisco: Morgan Kaufmann.

Bluck, S., T. Herbermas, et al. (2002). "Self-reporting functions of autobiographical memory in young adults: self-continuity, problem-solving and development of intimacy."

Bodner, G. E. and D. S. Lindsay (2003). "Remembering and knowing in context." Journal of Memory and Language **48**(3): 563-580.

Bransford, J., A. L. Brown, et al. (1999). How People Learn: Brain, Mind, Experience and School, Washington DC: National Academy Press.

Brewer, W. and C. Sampaio (2006). "Processes leading to confidence and accuracy in sentence recognition: A metamemory approach." Memory **14**: 540-552.

Brewer, W. F. and J. C. Tenyens (1981). "Role of schemata in memory for places." Cognitive Psychology **13**: 207-230.

Brin, S. and L. Page (1998). "The anatomy of a large-scale hypertextual Web search engine." Computer Networks and ISDN Systems **30**(1-7): 107-117.

Brockmeier, J. and D. Carbaugh (2001). Narrative and Identity. Studies in Autobiography, Self and Culture
John Benjamins Publishing Company.

Brotherton, J. A. and G. D. Abowd (2004). Lessons learned from eClass: Assessing automated capture and access in the classroom, ACM. **11:** 121-155.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms, New Jersey: Lawrence Erlbaum Associates.

Brown, A. S. (1991). "A review of the tip-of-the-tongue experience." Psychological Bulletin **2**(109): 204-223.

Brown, N. R., L. J. Rips, et al. (1985). "The subjective dates of natural events in very long-term memory." Cognitive Psychology **17**: 139-177.

Brown, R. and J. Kulik (1977). "Flashbulb memories." Cognition **5**: 73-99.

Bush, V. (1945). As we may think. The Atlantic Monthly. **1**: 101--108.

Carroll, J. M. (2000). Making Use: Scenario-Based Design of Human-Computer Interactions, MIT Press.

Cavanaugh, J. C. (1988). The place of awareness in memory development across adulthood. Everyday cognition in adulthood and later life, Cambridge University Press.

Chalfen, R. (1987). "Snapshot Versions of Life." Bowling Green Ohio, Popular Press.

Chalfen, R. (1997). Family photography: One album is worth a 1000 lies. Sociology: Exploring the architecture of everyday life D. M. Neuwman, Thousand Oaks, CA.: Pine Forge Press: 269-278.

Cohen, G. (1996). Memory in the Real World, The Open University, Psychology Press.

Cohen, G. (1998). The effects aging on autobiographical memory. Autobiographical memory: theoretical and applied perspectives. C. P. Thomspson, D. J. Herrmann, D. Bruceet al. NJ, NJ: Lawrence Erlbaum Associates Inc: 105-123.

Cohen, G., G. Kiss, et al. (1993). Memory: Current Issues (Open Guide to Psychology), Open University Press.

Conway, M. A. (1990). Autobiographical Memory – An Introduction, Open University Press.

Conway, M. A. (1995). Flashbulb Memories.

Conway, M. A., S. J. Anderson, et al. (1994). "The formation of flashbulb memories." Memory and Congnition **22**: 326-343.

Conway, M. A. and D. A. Bekerian (1987). "Organization in autobiographical memory." Memory & Cognition(15): 119-132.

Conway, M. A. and C. W. Pleydell Pearce (2000). "The construction of autobiographical memories in the self memory system." Psychological Review(107): 261-288.

Cowan, N. (1984). "On Short And Long Auditory Stores." Psychological Bulletin **96**(2): 341- 370.

Crabtree, A., T. Rodden, et al. (2004). Collaborating around collections: informing the continued development of photoware. Proceedings of the 2004 ACM conference on Computer supported cooperative work. Chicago, Illinois, USA, ACM.

Craik, F. I. M. (2002). Levels of processing: Past, present ... and future?, Psychology Press. **10:** 305 - 318.

Csikszentmihalyi, M. and E. Rochberg-Halton (1981). The meaning of things - Domestic symbols and the self., Cambridge University Press.

Cutrell, E., D. Robbins, et al. (2006). Fast, flexible filtering with phlat. Proceedings of the SIGCHI conference on Human Factors in computing systems. Montr\&\#233;al, Qu\&\#233;bec, Canada, ACM.

Czerwinski, M., E. Horvitz, et al. (2004). A diary study of task switching and interruptions. Proceedings of the SIGCHI conference on Human factors in computing systems. Vienna, Austria, ACM.

Davis, R. C., J. A. Landay, et al. (1999). NotePals: lightweight note sharing by the group, for the group. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit. Pittsburgh, Pennsylvania, United States, ACM.

DeVaul, R. W., S. A. Pentland, et al. (2003). The Memory Glasses: Subliminal vs. Overt Memory Support with Imperfect Information. Proceedings of the 7th IEEE International Symposium on Wearable Computers, IEEE Computer Society.

Dib, L. and V. Kalnikaitė (2008). Sketching and Contextualizing Sonic Family Memories. Workshop on Designing for Families, part of CSCW 2008 conference. San Diego, California, USA: 4.

Dickie, C., R. Vertegaal, et al. (2004). Augmenting and sharing memory with eyeBlog. Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences. New York, New York, USA, ACM.

Dieberger, A., P. Dourish, et al. (2000). Social navigation: techniques for building more usable systems, ACM. **7:** 36-45.

Dix, A., J. E. Finley, et al. (2003). Human-Computer Interaction.

Dourish, P., W. K. Edwards, et al. (1999). "Presto: An Experimental Architecture for Fluid Interactive Document Spaces." ACM Transactions on Computer-Human Interaction **6**(2): 133-161.

Dumais, S., E. Cutrell, et al. (2003). Stuff I've seen: a system for personal information retrieval and re-use. Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval. Toronto, Canada, ACM.

Durso, T. F. and K. M. Johnson (1980). "The Effects of Orienting Tasks on Recognition, REcall, and Modality Confusion of Pictures and Words." Journal of Verbal Learning and Verbal Behaviour(19): 416-429.

Easterbrook, J. A. (1959). "The effect of emotion on cue utilization and the organization of behaviour." Psychological Review **66**: 183-201.

Ekman, P. (1999). Basic emotions. Sussex, U.K, John Wiley & Sons, Ltd..

Ericsson, K. and H. Simon (1993). Protocol Analysis: Verbal Reports as Data Boston, Boston: MIT Press.

Fiske, P. A. (1995). "Social Schema for Remembering People: Relationships and Person Attributes in Free Recall of Acquaintances." Journal of Quantitative Anotropology(5): 305-324.

Foote, J., J. Boreczhy, et al. (1998). An intelligent media browser using automatic multimodal analysis. Proceedings of the sixth ACM international conference on Multimedia. Bristol, United Kingdom, ACM.

Freeman, E. and D. Gelernter (1996). Lifestreams: a storage model for personal data, ACM. **25**: 80-86.

Frohlich, D., A. Kuchinsky, et al. (2002). Requirements for photoware. Proceedings of the 2002 ACM conference on Computer supported cooperative work. New Orleans, Louisiana, USA, ACM.

Frohlich, D. and R. Murphy (2000). The Memory Box, Springer-Verlag. **4**: 238-240.

Gemmell, J., G. Bell, et al. (2006). "MyLifeBits: a personal database for everything." 49(1): 88-95.

Graesser, A. C. and G. V. Nakamura (1982). The impact of a schema on comprehension and memory. The Psychology of learning and motivation: Advances in research and theory. B. G. (Eds.). NY, New York: Academic Press. **16.**

Greene, J. and M. D'Oliveira (2006). Learning to use statistical tests in psychology, Open University Press.

Greenfield, A. (2006). Everyware: the dawning age of ubiquitous computing. New Riders: 11–12.

Grudin, J. and D. Bargeron (2005). Multimedia Annotation: An Unsuccessful Tool Becomes a Successful Framework. Communication and Collaboration Support Systems. T. H. a. T. I. E. K. Okada. Ohmsha.

Harper, R., D. Randall, et al. (2008). The past is a different place: they do things differently there. Proceedings of the 7th ACM conference on Designing interactive systems. Cape Town, South Africa, ACM.

Harvel, L. D., C. Scheibe, et al. (2005). Technology design for connecting student notes to online course content. INTERACT.

Hayes, G., S. N. Patel, et al. (2004). The Personal Audio Loop: Designing a Ubiquitous Audio-Based Memory Aid. Mobile HCI 2004.

Haynes, R. S., M. J. Carroll, et al. (2009). Design research as explanation: perceptions in the field. Proceedings of the 27th international conference on Human factors in computing systems. Boston, MA, USA, ACM.

He, L., J. Grudin, et al. (2000). Designing presentations for on-demand viewing. Proceedings of the 2000 ACM conference on Computer supported cooperative work. Philadelphia, Pennsylvania, United States, ACM.

Healey, J. and W. R. Picard (1998). StartleCam: A Cybernetic Wearable Camera. Proceedings of the 2nd IEEE International Symposium on Wearable Computers, IEEE Computer Society.

Heuer, F. and D. Reisberg (1990). "Vivid memories of emotional events: The accuracy of remembered minutiae." Memory & Cognition **18**: 496-450.

Hiltz, S. R. and R. Goldman (2005). Learning Together Online: Research on Asynchronous Learning Networks, Lawrence Erlbaum Associates.

Hodges, S., L. Williams, et al. (2006). SenseCam: A Retrospective Memory Aid. UbiComp 2006: Ubiquitous Computing: 177-193.

Hori, T. and K. Aizawa (2003). Context-based video retrieval system for the life-log applications. Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval. Berkeley, California, ACM.

Hoskins, J. (1998). Biographical objects - how things tell the stories of people's lives., Routledge.

Hughes, J. A., J. O'Brien, et al. (1997). Designing with ethnography: a presentation framework for design. Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques. Amsterdam, The Netherlands, ACM.

James, W. (1890). "The Principles of Psychology." London: MacMillan.

Johnson-Laird, P. N. (1988). The Computer and the Mind, Harvard University Press, Cambridge Mass. .

Johnson, D. W. (1981). "Student-student interaction: The neglected variable in education." Educational Research **10**(1): 5-10.

Jones, W., H. Bruce, et al. (2001). Keeping found things found on the web. Proceedings of the tenth international conference on Information and knowledge management. Atlanta, Georgia, USA, ACM.

Joslyn, S. L. and M. A. Oakes (2005). "Directed forgetting of autobiographical events." Memory & Cognition **33**(4): 577-587.

Kalnikaite, V. and S. Whittaker (2007). Software or wetware?: discovering when and why people use digital prosthetic memory. Proceedings of the SIGCHI conference on Human factors in computing systems. San Jose, California, USA, ACM.

Kalnikaite, V. and S. Whittaker (2008). Cueing Digital Memory: How and Why Do Digital Notes Help Us Remember? . British HCI, Liverpool, UK, ACM Press.

Kalnikaite, V. and S. Whittaker (2008). Social Summarization: Does Social Feedback Improve Access to Speech Data? CSCW, San Diego, USA, ACM Press.

Kalnikaitė, V. and S. Whittaker (2008). MoodyPie: Emotional Photology for Autobiographical Memories. Workshop on Emotion in HCI, part of British HCI 2008 conference. Liverpool John Moores University, Liverpool, UK.

Karger, D. R. and D. Quan (2004). Haystack: a user interface for creating, browsing, and organizing arbitrary semistructured information. CHI '04 extended abstracts on Human factors in computing systems. Vienna, Austria, ACM.

Kawamura, T., T. Fukuhara, et al. (2007). "Ubiquitous Memories: a memory externalization system using physical objects." Personal and Ubiquitous Computing **11**(4): 287-298.

Kazman, R., W. Hunt, et al. (1995). Dynamic Meeting Annotation and Indexing. DMS, HonoluluallHI.

Kidd, A. (1994). The marks are on the knowledge worker. Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence. Boston, Massachusetts, United States, ACM.

Kiewra, K. A. (1985). Investigating Notetaking and Review: A Depth of Processing Alternative, Routledge. **20:** 23 - 32.

King, G. (1986). "Say "Cheese"." The Snapshot As Art And Social History, London, Collins.

Kirk, D., A. Sellen, et al. (2006). Understanding photowork. Proceedings of the SIGCHI conference on Human Factors in computing systems. Montr\&\#233;al, Qu\&\#233;bec, Canada, ACM.

Kleinsmith, L. J. and S. Kaplan (1963). "Paired-associate learning as a function of arousal and interpolated interval." Journal of Experimental Psychology: Learning, Memory, & Cognition(65): 190-193.

Knez, I. (2006). "Autobiographical memories for places." Memory **14**(3): 359-377.

Kopytoff, I. (1986). The cultural biography of things: commoditization as a process. The social life of things --- commoditization and cultural perspective. A. e. In Appadurai, Cambridge University Press.

Koriat, A. (2000). "The feeling of knowing: Some metatheoretical implications for consciousness and control." Consciousness and congnition **9**: 149-171.

Koriat, A., M. Goldsmith, et al. (2000). "Toward a psychology of memory accuracy." Annual Review of Psychology **51**: 481-537.

Lachman, J. L., R. Lachman, et al. (1979). "Metamemory through the adult life span." De- velopmental Psychology **15**: 543-551.

Lachman, M. E. and R. Leff (1989). "Perceived control and intellectual functioning in the elderly: a 5 year longitudinal study." Developmental Psychology **25**: 722-8.

Lamming, M. and M. Flynn (1994). "Forget-me-not" Intimate Computing in Support of Human Memory. FRIEND21, International Symposium on Next Generation Human Interface. Meguro Gajoen, Japan.

Landay, J. A. and T. R. Kaugmann (1993). User Interface Issues in Mobile Computing. WWOS-IV, IEEE Press.

LeeTiernan, S. and J. Grudin (2001). Fostering engagement in asynchronous learning through collaborative multimedia annotation. INTERACT, IOS Press.

Lewis, C., P. G. Polson, et al. (1990). Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people. Seattle, Washington, United States, ACM.

Linton, M. (1982). Transformation of memory in everyday life. Memory Observed: Remembering in natural context. U. E. Neisser. SF, San Francisco: Freeman.

Loftus, E. F. and W. H. Calvin. (2001). "Memory's Future. Retrieved 5th March 2009." from http://williamcalvin.com/2001/PsychToday.htm.

MacWhinney, B., J. M. Keenan, et al. (1982). "The role of arounsal in memory for conversation." Memory & Cognition(10): 308-317.

Malone, T. W. (1983). How do people organize their desks?: Implications for the design of office information systems, ACM. **1**: 99-112.

Mann, S., J. Fung, et al. (2005). Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences. Conference on Computer Human Interaction, ACM Press.

McCarthy, J. and P. Wright (2004). What is enjoyment doing to HCI? 11th European Conference on Cognitive. European Association of Cognitive Ergonomics, Le Chesney, France.

McNally, R. J., S. A. Clancy, et al. (2001). "Directed forgetting of trauma cues in adults reporting repressed or recovered memories of childhood sexual abuse." Journal of Abnormal Psychology **110**: 151-156.

Miller, G. A. (1956). "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information." Psychological Review(63): 81-97.

Monk, A., P. Wright, et al. (1993). Improving your human-computer interface: A practical technique, Prentice Hall International (UK) Ltd..

Moran, T. P., L. Palen, et al. (1997). I'll get that off the audio: a case study of salvaging multimedia meeting records. Proceedings of the SIGCHI conference on Human factors in computing systems. Atlanta, Georgia, United States, ACM.

Mori, J., N. Basselin, et al. (2007). SharedLife: Sharing of Augmented Personal Memories in Ubiquitous Environments. 21st Annual Conference on the Japanese Society for Articificial Intelligence, Miyazaki, Japan.

Munteanu, C., R. Baecker, et al. (2008). Collaborative editing for improved usefulness and usability of transcript-enhanced webcasts. Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. Florence, Italy, ACM.

Munteanu, C., G. Penn, et al. (2006). Automatic Speech Recognition for Webcasts: How Good is Good Enough and What to Do When it Isn't. In Proceedings of the Eight International Conference on Multimodal Interfaces - ICMI, Banff, Alberta, Canada.

Nickerson, R. S. (1977). Some comments on human archival memory as a very large data base. . 3rd International Conference on Very Large Data Bases, Tokyo, Japan.

Nielsen, J. (1994). Heuristic evaluation. Usability Inspection Methods. J. Nielsen, and Mack, R.L. (Eds.). New York, John Wiley & Sons.

Norman, D. (2002). The Design Of everyday Things.

Norman, D. A. (1986). Cognitive engineering. In User Centered System Design. NJ, Erlbaum, Hillsdale, NJ: 31-61.

Oldfield, R. C. and O. L. Zangwill (1944). "Head's Concept of the Schema and its Application in Contemporary British Psychology, Bartlett's Theory of Memory." British Journal of Psychology **Part III**(33): 113-129.

Oleksik, G., D. Frohlich, et al. (2008). Sonic interventions: understanding and extending the domestic soundscape. Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. Florence, Italy, ACM.

Paivio, A. (1986). Mental representations: a dual coding approach. Oxford, England: Oxford University Press.

Paivio, A. (2006). Mind and its evolution: a dual coding theoretical approach, Routledge.

Petrelli, D., E. van den Hoven, et al. (2009). Making History: Intentional Capture of Future Memories. CHI 2009. Boston, US, ACM Press.

Petrelli, D., S. Whittaker, et al. (2008). Autotopography: what can physical mementos tell us about digital memories? Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. Florence, Italy, ACM.

Pimentel, M. G. C., G. D. Abowd, et al. (2000). Linking by interacting: a paradigm for authoring hypertext. Proceedings of the eleventh ACM on Hypertext and hypermedia. San Antonio, Texas, United States, ACM.

Piolat, A., T. Olive, et al. (2005). "Cognitive effect during note taking." Applied Cognitive Psychology **19**(3): 291.

Ratto, M., R. B. Shapiro, et al. (2003). The ActiveClass Project: Experiments in Encouraging Classroom Participation. CSCL.

Rekimoto, J. (1999). TimeScape: a time machine for the desktop environment. CHI '99 extended abstracts on Human factors in computing systems. Pittsburgh, Pennsylvania, ACM.

Rhodes, B. J. (2000). Margin notes: building a contextually aware associative memory. Proceedings of the 5th international conference on Intelligent user interfaces. New Orleans, Louisiana, United States, ACM.

Rhodes, B. J. and P. Maes (2000). "Just-in-time information retrieval agents." IBM Syst. J. **39**(3-4): 685-704.

Ringel, M., E. Cutrell, et al. (2003). Milestones in Time: The Value of Landmarks in Retrieving Information from Personal Stores
Proceedings of Interact.

Robertson, M., M. Lane, et al. (2000). "International students, learning environments and perceptions: A case study using the Delphi technique." Higher Education Research and Development **19**(1): 89-102.

Rodden, K. and K. R. Wood (2003). How do people manage their digital photographs? Proceedings of the SIGCHI conference on Human factors in computing systems. Ft. Lauderdale, Florida, USA, ACM.

Rosenschein, S. J. (2004). Quindi meeting companion: a personal meeting-capture tool. Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences. New York, New York, USA, ACM.

Rubin, C. D. (1988). Autobiographical Memory, Cambridge University Press.

Rubin, D. C., R. W. Schrauf, et al. (2003). "Belief and recollection of autobiographical memories." Memory and Cognition(31): 887-901.

Sachs, J. S. (1967). "Recognition memory for syntactic and semantic aspects of connected discourse." Perception & Psychophysics(2): 437-42.

San Pedro, J., V. Kalnikaite, et al. (2009). You can play that again: exploring social redundancy to derive highlight regions in videos. Proceedings of the 13th international conference on Intelligent user interfaces. Sanibel Island, Florida, USA, ACM.

Schacter, D. L. (1987). "Implicit memory: history and current status." Journal of Experimental Psychology: Learning, Memory, and Cognition **13**: 501-518.

Schacter, D. L. (1997). Searching For Memory: The Brain, The Mind, And The Past, Basic Books.

Schacter, D. L. and R. L. Buckner (1998). "On the Relations among Priming, Concious Recollection, and Intentional Retrieval: Evidence from Neuroimaging Research." Neurobiology of Learning and Memory **70**: 284-303.

Schank, R. C. (1982). Reminding and memory organisation. Strategies for natural language processing. W. G. Lehnert and M. E. Ringle. NJ, NJ: Lawrence Erlbaum Associates Ins.

Sellen, A. J., A. Fogg, et al. (2007). Do life-logging technologies support memory for the past?: an experimental study using sensecam. Proceedings of the SIGCHI conference on Human factors in computing systems. San Jose, California, USA, ACM.

Sellen, A. J., G. Louie, et al. (1997). "What Brings Intentions to Mind?" Memory **5**(4): 483-507.

Sharp, H., Y. Rogers, et al. (2007). Interaction Design: Beyond Human-Computer Interaction, Wiley.

Shea, P., E. Fredericksen, et al. (2001). Measures of learning effectiveness in the SUNY Learning Network. In J. Bourne & J. Moore (Eds) Online Education: Proceedings of the 2000 Sloan Summer Workshop on Asynchronous Learning Networks, Sloan-C Press.

Sheskin, D. (2007). Handbook of parametric and nonparametric statistical procedures, Boca Raton, FL: Chapman & Hall. CRC.

Shneiderman, B. (1997). Designing the User Interface: Strategies for Effective Human-Computer Interaction, Addison-Wesley Longman Publishing Co., Inc.

Shum, M. S. (1998). "The role of temporal landmarks in autobiographical memory processes." Psychological Bulletin **124**: 423-442.

Smith, S. M. (1979). "Remembering in and out of context." Journal of Experimental Psychology: Human Learning and Memory. **5**(5): 460-471.

Smith, S. M. and E. Vela (2001). "Environmental context-dependent memory: A review and meta-analysis." Psychonomic Bulletin & Review **8**: 203-220.

Sperling, G. (1960). "The information available in brief visual presentations." Psychological Monographs(74): 1-29.

Spiro, R. J., P. J. Feltovich, et al. (1992). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. Duffy & D. Jonassen (Eds.), Constructivism and the Technology of Instruction, Hillsdale, NJ: Erlbaum.

Stafford, L. and J. A. Daly (1984). "Conversational Memory: The Effects of Recall Mode and Memory Expectancies on Remembrances of Natural Conversations." Human Communication Research (10): 379-402.

Sternberg, R. J. (2006). Cognitive Psychology (4th ed.).

Stevens, M. M., G. D. Abowd, et al. (2003). "Getting into the Living Memory Box: Family archives & holistic design." Personal and Ubiquitous Computing **7**(3): 210-216.

Stifelman, L., B. Arons, et al. (2001). The audio notebook: paper and pen interaction with structured speech. Proceedings of the SIGCHI conference on Human factors in computing systems. Seattle, Washington, United States, ACM.

Stifelman, L. J., B. Arons, et al. (1993). VoiceNotes: a speech interface for a hand-held voice notetaker. Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems. Amsterdam, The Netherlands, ACM.

Tan, D., J. Stefanucci, et al. (2001). The Infocockpit: providing location and place to aid human memory. Proceedings of the 2001 workshop on Perceptive user interfaces. Orlando, Florida, ACM.

Tauscher, L. and S. Greenberg (1997). "How people revisit web pages: Empirical findings and implications for the design ofhistory systems." International Journal of Human Computer Studies, Special issue on World Wide Web Usability **47**(97): 138.

Thomsen, D. K. and D. Berntsen (2005). "The end point effect in autobiographical memory: More than a calendar is needed." Memory **13**(8): 846-861.

Tucker, S. and S. Whittaker (2006). Time is of the essence: an evaluation of temporal compression algorithms. Proceedings of the SIGCHI conference on Human Factors in computing systems. Montr\&\#233;al, Qu\&\#233;bec, Canada, ACM.

Tulving, E. (1972). Episodic and semantic memory, New York: Academic Press.

Tulving, E. (2002). "Episodic memory: from mind to brain." Episodic memory: from mind to brain **53**: 1-25.

Tulving, E. and D. L. Schacter (1990). "Priming and human memory systems." Science **247**(4940): 301-6.

Turkle, S. (2007). Evocative Objects - Things we think with, MIT Press.

Tversky, B. (2004). Narratives of Space, Time, and Life. **19**: 380-392.

Vaidya, C. J. and J. D. E. Gabrieli (2000). "Picture superiority in conceptual memory: Dissociative effects of encoding and retrieval tasks." Memory & Cognition **28**(7): 1165-1172.

Vemuri, S., C. Schmandt, et al. (2006). iRemember: a personal, long-term memory prosthesis. Proceedings of the 3rd ACM workshop on Continuous archival and retrival of personal experences. Santa Barbara, California, USA, ACM.

Walker, T. and J. Moore (2005). Work in progress - instant distance learning course development using Silicon Chalk&trade. Frontiers in Education, 2005. FIE '05. Proceedings 35th Annual Conference.

Watson, J. B. (1919). "Behaviorism." New York: Holt.

Weldon, M. S. and K. C. Coyote (1996). "Failure to find the picture superiority effect in implicit conceptual memory tests." Journal of Experimental Psychology: Learning, Memory, & Cognition **22**: 670-686.

Weldon, M. S. and H. L. Roediger (1987). "Altering retrieval demands reverses the picture superiority effect." Memory and Cognition **15**(4): 269-280.

Wellman, H. M. (1977). "Tip of the tongue and feeling of knowing experiences: A developmental study of memory monitoring." Child Development, **48**: 13-21.

Wenyin, L., S. T. Dumais, et al. (2001). Semi-automatic image annotation. INTERACT 2001, IOS Press.

Whittaker, S., O. Bergman, et al. (2009). "Easy on That Trigger Dad: A Study of Long Term Family Photo Retrieval." To appear in Personal and Ubiquitous Computing.

Whittaker, S., O. Bergman, et al. (2009). "Easy on that trigger dad: a study of long term family photo retrieval." Personal and Ubiquitous Computing.

Whittaker, S. and J. Hirschberg (2003). Look or listen: Discovering effective techniques for accessing speech data. CHI.

Whittaker, S., J. Hirshberg, et al. (2002). SCANMail: a voicemail interface that makes speech browsable, readable and searchable. Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves. Minneapolis, Minnesota, USA, ACM.

Whittaker, S., P. Hyland, et al. (1994). FILOCHAT: handwritten notes provide access to recorded conversations. Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence. Boston, Massachusetts, United States, ACM.

Whittaker, S., Q. Jones, et al. (2004). ContactMap: Organizing communication in a social desktop, ACM. **11:** 445-471.

Whittaker, S. and C. Sidner (1996). Email overload: exploring personal information management of email. Proceedings of the SIGCHI conference on Human factors in computing systems: common ground. Vancouver, British Columbia, Canada, ACM.

Whittaker, S., J. Swanson, et al. (1997). "TeleNotes: managing lightweight interactions in the desktop." 4(2): 137-168.

Whittaker, S., S. Tucker, et al. (2008). "Design and evaluation of systems to support interaction capture and retrieval." Personal and Ubiquitous Computing **12**(3): 197-221.

Wilcox, D. L., N. B. Schilit, et al. (1997). Dynomite: a dynamically organized ink and audio notebook. Proceedings of the SIGCHI conference on Human factors in computing systems. Atlanta, Georgia, United States, ACM.

Wilkerson, M., G. W. Griswold, et al. (2005). Ubiquitous presenter: increasing student access and control in a digital lecturing environment, ACM.

Wilson, A. E. and M. Ross (2003). "The identity function of autobiographical: Time is on our side." Memory **11**(2): 137-149.

Winograd, E. (1988). Some observations on prospective remembering. Practical Aspects of Memory: Current Research and Issues. M. M. Gruneberg, P. E. Morris and R. N. E. Sykes, Chichester: Wiley. **2**: 348-353.

Wright, C. P. and F. A. Monk (1991). The use of think-aloud evaluation methods in design, ACM. **23**: 55-57.

Wright, D. B. (1993). "Recall of the Hillsborough disaster over time: Systematic biases of "flashbulb" memories." Applied Cognitive Psychology **7**: 129-138.

Wright, D. B. and G. Gaskell (1992). The construction and function of vivid memories. Theoretical perspectives on autobiographical memory. M. A. Conway, D. C. Rubin, H. Spinnler and W. A. Wagenaar, Dordrecht: Kluwer Academic Publishers: 275-292.

Zerubavel, E. (1996). "Social memories: Steps to a sociology of the past." Qualitative Sociology **19**(3): 283-299.