



University  
of Glasgow

Macknet, David Taylor (2012) *Decisions to delete: subjectivity in information deletion and retention*.

PhD thesis

<http://theses.gla.ac.uk/3286/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

# **Decisions to Delete:**

**Subjectivity in Information Deletion and Retention**

University of Glasgow

Humanities Advanced Technology Information Institute

## **David Taylor Macknet**

**Bachelor of Arts, British and American Literature, Pacific Union College**

**Master of Science, Information Systems, University of San Francisco**

**Master of Letters, Philosophy, University of Glasgow**

Submitted in fulfilment of the requirements for the  
Degree of PhD, Humanities Computing

Department of Humanities Advanced Technology Information Institute  
Faculty of Arts  
University of Glasgow

March, 2012

## Abstract

This research examines the decision-making process of computer users with reference to deletion and preservation of digital objects. Of specific interest to this research is whether people provide different reasons for deleting or preserving various types of digital object dependant upon whether they are making such decisions at home or at work, whether such decisions are to any extent culturally determined, and whether they consider others in the course of making such decisions.

This study considers the sociological implications of such decisions within organisations, and various psychological errors to be expected when such decisions are made. It analyses the reasons given for these decisions, within the contexts of home and work computing. It quantifies the frequency with which these activities are undertaken, the locations in which such objects are stored, and what aids the user in making such decisions.

This research concludes that, while computer users generally desire their digital objects to be organised, they are not provided with adequate support from their computer systems in the decision to delete or preserve digital objects. It also concludes that such decisions are made without taking advantage of metadata, and these decisions are made for the same reasons both at home and at work: there is no discernible difference between the two contexts in terms of reasons given for such decisions. This study finds no correlation between subjects' culture and reasons given for deletion / preservation decisions, nor does it find any correlation between age and such reasons. This study further finds that users are generally averse to conforming to records management policies within the organisation.

For archivists and records managers, this research will be of particular interest in its consideration of the usage of and attitudes towards records management systems. Specifically, in organisations possessing formal records management systems, this research investigates the frequency with which individuals violate records management procedures and why they consider such violations to be necessary or desirable. This research also argues towards a more proceduralised decision-making process on the part of the ordinary user and a deeper integration between records management systems and computer operating systems.

Designers of formal information systems should consider this research for its implications regarding the way in which decisions are affected by the context in which those decisions are made. Information systems design may be best suited to understanding—and ameliorating—certain types of cognitive error such that users are enabled to make better deletion and preservation decisions. User interface designers are uniquely positioned to address certain cognitive errors simply by changing how information is presented; this research provides insight into just what those errors are and offers suggestions towards addressing them.

For sociologists concerned with institutional memory, this research should be of interest because the deletion and preservation decisions of members of an organisation are those which shape the collection of digital artefacts available for study. Understanding the reasons for these decisions is likely to inform what interpretations can be drawn from the study of such collections. Also of interest to sociologists will be the variety of reasons given for deletion or preservation, as those reasons and decisions are what shape, to some extent, institutional memory.

---

# Contents

---

<b>Contents</b>	<b>i</b>
<b>List of Tables</b>	<b>iv</b>
<b>List of Figures</b>	<b>v</b>
<b>1 Introduction and Background</b>	<b>1</b>
1.1 Terminology . . . . .	3
1.2 Background . . . . .	4
1.3 Thesis Plan . . . . .	8
<b>2 Related Work</b>	<b>9</b>
2.1 Concepts of Age / Youth . . . . .	9
2.2 The Computer as (an amoral) Extension of Self . . . . .	10
2.3 Conceptualising Digital Objects . . . . .	13
2.4 Sociology Including Objects . . . . .	15
2.5 Computer / Cyborg as Mirror of Self / Role Model . . . . .	17
2.6 What brings about moral behaviour? . . . . .	19
2.6.1 Social Control . . . . .	19
2.6.2 Dissemination of Ethical Standards . . . . .	21
2.6.3 Normativity? Or Merely Nash Equilibria? . . . . .	22
2.6.4 Ethical Relativism? . . . . .	23
2.6.5 Thinking Individuals within Organisations . . . . .	26
2.7 Difference between Ought and Action . . . . .	29
2.8 Context and Selection . . . . .	30
2.9 Verbal vs. Pictorial Presentation yields differences in decision . . . . .	33
2.10 Finding Patterns Where None Exist . . . . .	34
2.11 Psychological Issues: Common Errors in Judgement . . . . .	36
2.11.1 Psychological Issues: Unconscious Decision-making . . . . .	36
2.11.2 Errors in Comparison . . . . .	36
2.11.3 More Errors in Comparison . . . . .	40
2.11.4 Errors in Probability Estimation . . . . .	43
2.12 Consistent Errors in Calculating Probability . . . . .	47
2.12.1 Errors in Calculating Risk . . . . .	50
2.12.2 Ambiguity and Comparison of Risks . . . . .	53
2.13 Psychological Issues: Confusion in Concepts . . . . .	54

2.13.1	Confusion between Causal and Indicative . . . . .	54
2.13.2	Irrelevant Information Affects Decisions . . . . .	56
2.14	Psychological Issues: Normativity of Errors in Judgement . . . . .	57
2.15	Not “Records” Behaviour, “Social” Behaviour . . . . .	58
<b>3</b>	<b>Research Methodology and Design</b>	<b>61</b>
3.1	Research Questions Arising from Literature Review . . . . .	62
3.2	Study Aims . . . . .	69
3.3	Different Means of Fulfilling the Study Aims . . . . .	70
3.3.1	Quantitative Study . . . . .	70
3.3.2	Qualitative Study . . . . .	71
3.3.3	Mixed Methods Study . . . . .	71
3.3.4	Methodological Selection . . . . .	72
3.3.5	Grounded Theory, Described . . . . .	74
3.4	Ethical Considerations . . . . .	75
3.5	Alternate Methods Considered . . . . .	77
3.5.1	Participant Observation . . . . .	77
3.5.2	Ethnography . . . . .	78
3.5.3	Interviews . . . . .	79
3.5.4	Focus groups . . . . .	80
3.6	Research Method: Surveys . . . . .	80
3.6.1	Survey Composition . . . . .	81
3.7	Potential Weaknesses of Surveys as Study Method . . . . .	86
3.7.1	Context . . . . .	86
3.7.2	Randomness of Subject Selection . . . . .	86
3.7.3	Terms Used for ‘Employment’ . . . . .	86
3.7.4	Transformation of Digital Objects . . . . .	87
3.7.5	Frequency of Reasons Used in Decisions . . . . .	88
3.7.6	Reasons Used for Different Object Types . . . . .	89
3.7.7	Object Presentation . . . . .	90
<b>4</b>	<b>Implementation of Study</b>	<b>92</b>
4.1	Data Analysis Methodology . . . . .	93
4.2	Notes on Data Consolidation . . . . .	95
<b>5</b>	<b>Interpretation of Results</b>	<b>104</b>
5.1	Analysis of Consolidated Data . . . . .	104
5.1.1	Demographic Findings . . . . .	104
5.1.2	Filing Activities Home . . . . .	105
5.1.3	Use Default Stores Home . . . . .	108
5.1.4	Clean-up Home . . . . .	110
5.1.5	Deletion Affordances Home . . . . .	111
5.1.6	Deletion Affordances Work . . . . .	114
5.1.7	Computer Skills Training . . . . .	115
5.1.8	Record Generating Activity . . . . .	116
5.1.9	Reasons Given for Deletion and Preservation . . . . .	118
5.1.10	Home / Work Mixing . . . . .	123
5.1.11	Work Records-Management Environment . . . . .	125

<b>6 Findings</b>	<b>131</b>
6.1 Organisation of Digital Objects Desired . . . . .	131
6.2 Operating Systems' Storage Locations Accepted . . . . .	133
6.3 Poor Metadata Usage Despite Training . . . . .	135
6.4 Computer Usage Not Context-Dependent . . . . .	137
6.5 Records Management Systems Usage . . . . .	141
6.6 Summary Conclusions . . . . .	141
6.6.1 Object Disposition . . . . .	141
6.6.2 Context . . . . .	142
6.6.3 Records Management Systems . . . . .	142
6.6.4 Research Implications . . . . .	143
<b>Bibliography</b>	<b>146</b>
<b>A Survey</b>	<b>156</b>
A.1 Subjectivity in Information Retention . . . . .	156
<b>B Data Analysis Scripts</b>	<b>179</b>
<b>Index</b>	<b>217</b>

---

# List of Tables

---

5.1	Breakdown of ‘Profession’ of study participants . . . . .	105
5.2	Age Range as Percent of Respondents within ‘Home Country’ . . . . .	105
5.3	‘Profession’ and ‘Use Default Stores Home’ . . . . .	110
5.4	Frequency of “clean-up” of user’s documents . . . . .	111
5.5	Comparison of ‘Deletion Affordances Work’ and ‘Deletion Affordances Home’ . . . .	114
5.6	Given ‘Regards Work IP As Their Property’, analysis of ‘Keeps Work IP’ . . . . .	129

---

# List of Figures

---

2.1	Stated Probability vs. Willingness to Gamble [FT04b, p. 749]	48
2.2	Proposition Stated as Gain / Loss vs. Willingness to Gamble [TK04c, p. 601]	50
5.1	Filing Activities Home	106
5.2	Age3 and Filing Activities Home	107
5.3	Profession and Filing Activities Home	107
5.4	Profession and Filing Activities Home 2	108
5.5	Use Default Stores Home	109
5.6	Clean-up Home	110
5.7	Deletion Affordances Home	112
5.8	Deletion Affordances Home 2	112
5.9	Profession and Deletion Affordances Home 2	113
5.10	Deletion Affordances Work	114
5.11	Training Organisation Pie	116
5.12	Record Generating Activity Home	117
5.13	Record Generating Activity Work Area	117
5.14	Functional Reasons Home	119
5.15	Functional Reasons Work	119
5.16	Personal Reasons Home Moving Average	120
5.17	Personal Reasons Work Moving Average	121
5.18	Reasons Involving Others Home Pie	122
5.19	Reasons Involving Others Work Pie	123
5.20	Home Work Mixing	124
5.21	Work RM System	126
5.22	RM Violation	127

## Chapter 1

---

# Introduction and Background

---

With the use of the general- or multi-purpose computing machines, people have been given responsibility for maintaining digital objects; these digital objects may be of historic or evidentiary value to the individual, to an organisation, or to some future as-yet-unidentified individual or group. However, because of the nature of these computers, decisions regarding the preservation or destruction of digital objects have been largely left up to the individual. Rather than certain classes of object being granted privileged status and actively managed, the general-purpose computer presents the user with a myriad means of managing all digital objects held upon the system. Because digital objects are necessarily mediated through the computer system, these digital objects are not necessarily presented in such a manner as would facilitate active management practices—digital objects are in some cases automatically retained, automatically destroyed, or are presented to the user via a generic interface which allows the user to elect, based upon vague information, to take some action upon those digital objects.

There is great concern within the fields of records- and archive-management [see Bai08] that “ordinary users” are granted such power over objects within their control, as such decisions would traditionally and historically have devolved to a professional within the fields. While these fields are generally concerned with very specific classes of object (in the case of records-management) or are concerned with objects which may prove to be of an historic interest (in the case of archivists), the concerns seem to be centre upon a broader set of issues, namely that the general-purpose computer provides both little framework for decisions to be made while simultaneously allowing a wide variety of individual action upon objects which may have evidentiary or historical interest. In order to address some of these concerns, it is necessary to understand not only what decisions are encouraged or discouraged by the various user-interfaces by which such decisions are affected, but also what motivations lie behind the committing such

decisions to action.

This research seeks to understand certain aspects about decisions made by human agents as they decide to keep or erase digital objects. The digital objects considered include such items as email, documents, movies, sound files, text messages, and email contacts. Study subjects were asked to provide information for both deletion and preservation of each object type, within the contexts of both home and work. Of interest were the reasons provided for these decisions and the frequency of performing each type of decision<sup>1</sup>. Of specific interest were whether participants considered others in their actions. Broadly speaking, the concepts of ‘duty’ and ‘obligation’ were examined: whether in the decision-making process participants considered that they had some responsibility towards others, themselves, towards their organisation, or towards their family members. This research also attempts to determine whether there is a relationship between decisions made in the home context vs. the work context—whether different reasons are given for actions in one context vs. the other.

In order to understand the decision-making process, we need to understand: how people conceptualise their relationship with the computer, how people make moral judgements versus what they state *qua* moral judgement, how people make decisions, and to have an understanding of cognitive errors people are likely to make in the course of making decisions. We must also understand whether there are gender-, generational- or culturally-based differences in decision-making, and whether any such differences may play a part in the responses given. This research is concerned with issues surrounding moral judgment or normativity, duty and responsibility, as applied to digital objects.

This research begins by examining a range of research thought to apply to the subject area. This literature review spans a wide variety of subject matter, generally concentrating within the areas of sociology and psychology but also touching upon issues in moral philosophy and—very briefly—neurology.

Research then progresses through the formalisation of research methodology. This chapter includes a discussion of the study aims, framed as questions possibly answerable via a study of human subjects. From there, the chapter proceeds to a consideration of different means to answer these questions and an examination of possible ethical issues which may affect the study. Also included in this chapter is a post-facto examination of assumptions which were overlooked during the initial framing of the study.

---

<sup>1</sup>It should be noted that this study does not examine the frequency of *applying* each reason to a particular decision: rather, this study enquired as to the frequency of decisions being made of a particular type, and as to whether particular reasons were explanatory of *any* of these decisions having been made. This was an explicit choice upon the part of the researcher, due to doubts regarding the accuracy of participants’ memory in estimating the frequency with which they have historically applied individual reasons to particular decisions.

The final two chapters provide a detailed analysis of the data generated during the course of the study, and a discussion of findings believed to be significant. Included in the findings are ‘summary conclusions,’ which was provided to study participants.

## 1.1 Terminology

It should be noted that this research uses several terms in a particular manner which is not in accordance with standard English usage. Additionally, there are certain terms which have no privileged meaning, but which may be used interchangeably.

### Necessary

This paper utilises the term ‘necessary’ in its logical or philosophical sense: if something is ‘necessarily’ the case, then the world could not be possibly be otherwise. Any use of the term is a strong assertion, distinct from any casual use of the term.

### Duty

‘Duty’ may entail moral, prudential, or fiduciary duty. ‘Duty’ should also be taken in the broadest sense of the term, meaning that the agent recognises some sense of obligation either to her present or future self, to another person, to posterity, to society, etc.

### Record

The term ‘record’ is given privileged use, where used, and is herein mostly taken to mean those objects which are managed by records managers, or in its formal meaning within such a context; the term ‘record’ also encompasses any document or digital object which is regarded by an individual or organisation to be somehow privileged or special, or is managed as being so. This primary usage of the term will be denoted as ‘record’. There is a secondary usage of the term ‘record’, however, which refers specifically to *database* records. Where this computer-science term is under discussion, effort has been made to specify that ‘database records’ are the item under discussion.

### Digital Object

This research considers the term ‘digital object’ to indicate an electronic artefact perceived by the computer user as a discrete unit. This could be a file present within a file system, a file folder, or otherwise; the term may encompass such objects as are formally recognised to be ‘of record’ (i.e., financial, medical, business, criminal, or other types of record).

### Home Culture

The term ‘home culture’ is used to indicate not a person’s residence but rather the

geographical location which has contributed most to a person’s cultural identity. For example, if a person grew up in Scotland yet currently lives in England, this paper will identify them as ‘Scottish’. This is significant because the study is concerned with identifying behaviour and attempting to draw conclusions regarding several different cultures and how those cultures may influence retention/destruction decisions.

### **Affordance**

The term ‘affordance’ is borrowed from product design engineering [Nor02]. An ‘affordance’ is a feature which indicates, to the user of an object, that actions should be performed in a certain manner. An example of affordances in the physical world would be push-bars on doors which open away from the individual passing through the door as compared to loop-type handles which are clearly meant to be pulled towards the door user. The term will be extended, in this paper, to include aspects of the computing world which serve a similar purpose (although not necessarily because they were designed to function in such a manner).

### **To Disposition**

The term ‘to disposition’ (past tense, ‘dispositioned’) should be taken to mean ‘to have decided whether to delete or preserve’. This term should be interpreted as being distinct from any formal assessment process—‘to disposition’ a document is *merely* to decide upon a particular course of action with regards a particular document; it should not be taken as implying any particular knowledge of archival or records-management practise such as appraisal, nor any particular decision-making process. ‘To disposition’ should not be taken as synonymous with ‘to dispose of’ or ‘to delete’.

## **1.2 Background**

The Records Management (‘RM’ *sic passim*) community has expressed concern that decisions made by “lay records managers” result in a poor quality of decision [see Bai08]. The RM community is concerned that there seem to be entire domains of record-like digital objects which are kept only by the users and consumers of those objects; such content is never subjected to “proper” records management processes. This concern has largely remained unexpressed outside of the records management community<sup>2</sup>.

---

<sup>2</sup>Cohen [Coh07] provides several examples wherein subject-matter experts are included in the records-management process for the purposes of determining which objects are of significance. Cohen is mostly concerned with institutional library management, however

This particular area seems to be one of great tension. On one side of the discussion lies argument that documents, in today’s world, are simply another manifestation of a form which has continued for thousands of years. In making this argument, Levy [Lev01] attempts to define the field as unchanged: people are simply behaving as they always have, which is to say that people, by and large, do not concern themselves with preserving their own artefacts against the possible interest of future generations. On another side of the discussion lies Bailey [Bai08], arguing that the change to digital documents is such a radical change that records managers must “buck up,” “bite the bullet,” descend into managing folksonomies and involve themselves in mastering technologies, lest the world of records management be consigned to the dust-bin of history, or be relegated to small islands of practice amongst civil servants and the like. Along with Bailey [Bai08] come the proponents of the “Big Bucket” school of thought, advocating some mix of formal taxonomy and folksonomy whereby users are given “big buckets” into which to “throw” their documents, which documents will be formally organised by the records manager at some later date [DX04; EN98; Wan+09]. On the third side of this discussion seems to be the general public: happily trundling along, erasing their own history, blissfully unaware of any cause for concern except when they are unable to locate an email, or when their computer crashes, etc. To them, the only possible solutions are simply to keep as many copies as possible, in as many locations as possible, or to trust that a single copy will be adequate and hope that nothing goes wrong. Users, for the most part, simply are not very concerned about what they choose to erase—they are concerned that *they* erase what they like, when they like, based upon their own (heretofore unknown) criteria.

This is our area of interest: the unsettled ground of ongoing digital history, wherein exist multiple originals, original copies, vastly growing stores of unmanageable information. This area is of interest not only because of the chaos and conflicting opinion—indeed, that chaos makes the area *less* interesting, if anything—but because it seems to be an area in which we find an odd disconnection between users of a technology (i.e., computers) and an understanding *of* that technology<sup>3</sup>, and who end up acting against their own best interests. This behaviour is of interest because it would seem to be something unparalleled in the analogue world, at least in the sheer scope and volume of the problem.

Within this general area lie certain concerns regarding human behaviour. We already know, to some extent, what the end result of their behaviour comes to, which is to say that the

---

<sup>3</sup>Certainly, it can be said that except for the simplest of technologies, users of the technology are to some extent ignorant of the workings of the technology—no strong assertion is being made, here, with regards to the digital world and its relationship to humanity; rather, the assertion is that the technology is of such vastly complex nature that even experts are necessarily ignorant of the underlying processes when engaged in using such artefacts in mundane ways. This ignorance results in a relationship not with the computer directly, but with the system “at hand,” mediating their relationship and affecting their decisions [Hei93].

concerns of records managers, historians, and information professionals are valid: people do not manage information very well, resulting in vast stores of digital *junk* on the one hand, and sorely depleted stores of articles of importance. Understanding what informs the decision to delete or retain digital objects is significant because so much depends upon such decisions—if the decision-making process is not fully understood, then we may expect the status quo to prevail, rather than addressing the very valid concerns of those who study the artefacts left behind in the aftermath of such decisions. What this study attempts to ask are questions about the *causes* of deletion or retention decisions. What *values* influence decisions to delete? What cultural phenomena influence users in their treatment of digital materials? What do users consider, when they save a copy of something or consign it to the “recycle bin?” *Why* do people delete or preserve digital objects?

For the purposes of this study, research questions were formalised in terms which could be answered by means of a human study. These questions served to direct the literature review portion of the research.

**Question 1.** What reasons do people give for deleting or preserving digital content?

**Question 2.** Do these reasons vary for each individual when operating in a different context (i.e., at work vs. at home)?

**Question 3.** Do particular reasons correlate with different professions, with particular demographic factors, different cultures, and / or different levels of technical skill?

**Question 4.** Are others considered in deletion / retention decisions?

**Question 5.** What types of “hygiene practises”<sup>4</sup> do participants engage in, and with what frequency?

**Question 6.** How do such hygiene practises correlate with reasons given for deletion / retention actions?

**Question 7.** To what extent do people engage in normatively aberrant actions<sup>5</sup> with regards to deletion / retention decisions and how do they rationalise these actions?

---

<sup>4</sup>“Hygiene practices” should be taken to include such activities as deleting, archiving, filing, organising—in general, hygiene practices are practices which maintain some semblance of order within the computer systems in question, but are particular to the management of digital objects rather than practices such as operating-system or hardware maintenance.

<sup>5</sup>Normatively aberrant actions should be understood to be those actions for which people would ordinarily censure themselves, e.g., behaviours which, when judged by their own moral values, would be considered wrong in some way.

**Question 8.** Does the existence of an RM system within the organisation correspond to any particular behaviour?

**Question 9.** Does having received RM or computer skills training affect the manner in which deletion / retention decisions are made, as determined by the reasons given for such decisions?

Three broad areas were considered as possibly influencing retention / destruction decisions, and as potentially being measurable for the purpose of determining correlation: differences in the individual or workplace, properties of the objects about which decisions were made, and perceptual factors which may influence such decisions.

Differences in employment context were considered such as whether their organisation has a strong concept of “record,” operates within a particular regulatory or legal environment, and whether their place is within a simple or complex organisational structure. Also considered was whether their home culture has a tradition of records management. The general computer literacy of the individual might have some bearing upon the types of decisions made, as might whether the individual routinely manipulated such content as part of their primary job function (e.g., a computer programmer might routinely manipulate software code files).

With regards to the digital objects themselves it was thought that decisions could be affected by such things as the physical location of the content (i.e., on a department or network drive, held by an application service provider, etc.), whether the content was shared or private (i.e., whether their digital objects are accessed or used by others), or differences in file-structure (e.g., is the content highly organised, or is it “kept loose” in a “My Documents” folder or on their “Desktop”).

One perceptual factor which may affect different decisions was thought to be whether the individual perceived the digital objects as partial objects (e.g., blocks of text within some larger context such as within a web-site or wiki page), as discrete objects (e.g., an individual file or document), or as groups of objects (e.g., file folders containing numerous, discrete files). It was also thought that any perception of “ownership” might have some influence upon retention / destruction decisions, in both the positive sense that any feelings that the participant had ownership might influence the process and also the negative sense that any feelings that someone else had ownership might affect such decisions. Finally, it was thought that such decisions might be influenced by whether the participant had in mind the concept of “a legacy” or “posterity” as part of their conscious intention regarding their curation process.

### 1.3 Thesis Plan

Chapter 2 provides an in-depth examination of literature believed to have bearing upon this research. It considers a number of sociological and psychological factors including how perceptions of age and youth (section 2.1) might affect decision-making; a wide variety of psychological issues such as how context affects selection (section 2.8) and errors in judgement (section 2.11); and also considers issues of morality such as how morality arises (section 2.6) and the normativity of judgement errors (section 2.14). The literature reviewed in this chapter was selected based upon whether said literature was thought to have potential bearing upon the research questions as detailed on page 6.

Chapter 3 (page 61) details the research methodology and design of the human study portion of this research. The literature review serves to refine the field of study from those research questions outlined on page 6 into a series of questions which were thought to be logically related and could be answered by human participants in a reasonable period of time; these questions are stated as study aims (section 3.2). Chapter 3 describes different research methods considered and provides rationale for the selection of electronic survey as the study method. Chapter 3 also details the means by which the surveys were constructed and addresses some potential weaknesses in this method of study (section 3.7 on page 86).

Chapter 4 (page 92) details the implementation of the survey<sup>6</sup>. It also discusses the data analysis methodology, and provides an explanation of the way in which data was consolidated in order to facilitate data analysis. This chapter, in combination with the survey and data analysis scripts<sup>7</sup> should allow for this research to be replicated.

Chapter 5 (page 104) provides an analysis of data generated by the survey. It is not exhaustive, but presents those findings which were thought to be of the most significance and from which theories as to human behaviour could be derived<sup>8</sup>.

Chapter 6 (page 131) builds upon the data analysis presented in chapter 5 and, together with literary sources discussed in chapter 2, draws conclusions with regards to the study findings and puts forth hypotheses for further study. Chapter 6 represents the evaluation of chapter 5 in light of literature examined in chapter 2, presenting possibilities for where human behaviour with regards to digital objects may merit deeper examination.

---

<sup>6</sup>See appendix A on page 156 for the complete survey.

<sup>7</sup>Appendix B on page 179.

<sup>8</sup>One of the aims in conducting qualitative research is the generation of hypotheses. See section 3.3.2 on page 71 for a discussion of qualitative research methodology and aims.

## Chapter 2

---

# Related Work

---

This chapter details a wide range of literature related to decision-making. It considers psychological issues such as how perceptions of age may affect how people regard digital objects, sociological issues such as what brings about moral behaviour, and a wide range of cognitive errors which may affect the disposition of digital objects. This chapter serves as the intellectual foundation for the investigation conducted in the study<sup>1</sup>. This chapter also serves to support the findings and conclusions reached, and provides the basis from which further hypotheses are stated<sup>2</sup>. This chapter is largely an investigation into decision-making, and will be referred to frequently throughout this research.

### 2.1 Concepts of Age / Youth

Into any discussion of deletion / preservation must come some consideration of the perceptions and prejudices brought about by how people conceive of “the old” versus “the new.” The very age of something, in technological terms in particular, affects any decision regarding the technology. While this may be particularly easy to demonstrate with regards to technological hardware (computers, mobile phones, etc.), prejudices in this area are also likely to be true of software artefacts; this may be demonstrated most easily, perhaps, with attitudes towards “old” file formats, or older versions of software. It may also be possible to demonstrate by observing how (in)frequently files are accessed within a file system, or how frequently older articles or blog posts are consulted.

People value new technological artefacts for their newness more than for their function. This way of valuing distorts the role played by technology, in that people do not, then, regard

---

<sup>1</sup>See chapter 3 on page 61 and chapter 4 on page 92 for details of the study.

<sup>2</sup>See chapter 6 on page 131.

technology as a tool first; rather, they regard possession of the tool as something which places them at risk of ageing; they are unable to fully incorporate it as an extension of themselves, as they are constantly forced to discard it lest they be perceived as ageing; they constantly focus on their own fear of ageing, through their decisions which are driven by this fear; in this arena their decisions are not made rationally. Woodward asserts that, “[f]or a [western] society such as ours that is built on the technological values of efficiency, cost-control, and innovation (the ‘new’), what is perceived as ‘old’ is understood as not only antithetical to our dominant values, but dangerous. Our language in everyday life confirms that a technology that is ‘getting old’ is suspect per se [...] [Woo99, p. 288].”

This skewed emphasis upon the age of technological items is likely to spill over into perceptions of digital objects, as well: if a digital object is “old,” or in an “old format,” it is more likely to be selected for destruction. The age of the item may, then, outweigh any other considerations. This is particularly unique to digital records, as they fall into the category of items for which this type of discrimination is particularly common merely because they are digital. Thinking about technological items in the digital arena seems to be dominated by worries about ageing.

## 2.2 The Computer as (an amoral) Extension of Self

Proceeding from considerations of age-based prejudice, and the ageing associated with “old” items of technology, we will consider the computer as an extension of the self; incorporation of technology items into one’s self to “stay young” is certainly a motivator, but is by no means the most significant cause of our close relationship with our machines.

Turkle[Tur84; Tur97] writes about the relationship between people and their computers<sup>3</sup>. She has studied this relationship across the axes of age, sex, profession. She examines the ways in which people become confused about the computer (in terms of its state of aliveness, consciousness, intelligence, etc.), and ways in which people choose to nurture their misconceptions (in which they knowingly treat the computer as if it were alive, conscious, intelligent, and avoid anything which might disturb these conceptions). She shows us that computers are different than other technologies, not only because they inhabit a strange, indeterminate space in our thinking, but because they provide us with a means to think: they

---

<sup>3</sup>Turkle has revised Turkle [Tur84] since its original publication, adding an additional introduction and epilogue. This revision does not address issues pertinent to this research, nor does it alter the original content of the work. There was an “anniversary” edition of this work published in 2004, which also maintained the author’s original points. Turkle may very well have changed her stance on some aspects of computing, particularly to do with the means in which the computer serves as a device for connecting individuals and through which individuals relate; this research is not concerned with such interactions *per se*, although it does consider whether interaction with digital objects may also be viewed as social behaviour (see section 2.15 on page 58).

are not thinking machines, they are thinking-with machines. We use them to examine ourselves, incorporate them into ourselves, get aspects of ourselves into the external world in a way which is impossible otherwise, so that we may consider our own thinking, reflect.

Turkle states clearly that the relationship with the computer is not gender-based, but that females (in Western society) are socialised to be “soft masters” and males to be “hard masters.” Soft masters relate to the computer task as artistry; they relate to the end result, rather than to the programming needed to achieve some end result; they relate to the computer as a tool for artistry, rather than as a tool through which to control. Hard masters relate to the computer task as if dictating actions to an avatar; programming is the important aspect to them, rather than the end result; they relate to the computer more as engineering than as artistry. For both styles of mastery, however, the computer is an extension of the self.

Turkle examines the types of community which have arisen surrounding the practise of computing, and considers which types of mastery are included / excluded. She examines the psychology of the individuals who are valued within these societies, and how these communities reinforce their values and protect themselves from outsiders. Particularly with the hacker community, she analyses them as being neurotic in their relationships with others and the machine, fearful of intimacy, and of finding solace and intimacy in the machine which, after all, cannot reject them. The depth of relationship with something which is without morality may contribute to the issue: having a deeply intimate relationship with what amounts to an amoral agent would seem to be only slightly different from having a deeply intimate relationship with an immoral agent, and to cause the user to accept the lack of morality; if one’s partner is immoral / amoral, the burden of morality is not only left with only one partner, but is a perhaps unwelcome addition to the partnership.

Considerations of duty to others may not only be influenced by the fact that one partner is amoral, but also because the relationship to the computer is inherently that of an intimate relationship between two: what one does with one’s own computer may be felt to be more personal than communal, excluding consideration of duty or responsibility to others by the very intimate nature of the relationship. As Turkle states, “[w]hen nineteenth-century Romantics looked for an alternative to the mechanism and competition of society, they looked to a perfect society of two, ‘perfect friendship,’ or ‘perfect love.’ This desire for fusion has its echo today, although in a new and troubling form. Instead of a quest for an idealized person, now there is the computer as a second self [Tur84, p. 307].”

Turkle talks of the taboo nature of relating to the machine, in terms of (Western) societies’ having drawn a hard line between science and sensuality. This, coupled with the degree of intimacy with the machine, would seem to place pressure upon the user to approach

computerised tasks from a distant, morally-disengaged perspective, particularly if the moral sense is at all perceived as an emotion. Turkle asserts that people, “Terrified of being alone, yet afraid of intimacy, we experience widespread feelings of emptiness, of disconnection, of the unreality of self. And here the computer, a companion without emotional demands, offers a compromise. You can be a loner, but never alone. You can interact, but need never feel vulnerable to another person [Tur84, p. 307].” For Turkle, the seduction of the computer acts to change the user, to strip away moral obligation because the user is relating to the machine. This is true whether the user is embracing the relationship (and, so, becoming more like the computer), or if the user is rejecting the intimacy of the relationship (and, so, becoming cold towards the computer, treating it as an object rather than anthropomorphising it).

Turkle also points to changes in the way we think of ourselves: as having been “programmed,” of needing to “shut down,” to “have some down time,” to “reboot.” These perceptions of ourselves as akin to an amoral agent may also skew the decisions made away from considerations of morality, duty, and responsibility; although this effect is not explicitly stated, Turkle [Tur84, pp. 271-2] strongly asserts that such ways are not simply a difference in ways of speaking about ourselves, but are differences in conceptualisation *of* our selves.

If people see the computer as an extension of themselves, then they relate to it in a manner which is very intimate, and possibly very worrisome: particularly, if they conceptualise the computer as some better-organised storage of their own thoughts, then saving/erasing things takes on a much greater significance. We cannot, ordinarily, choose to remember something in exact detail (a conversation, for example), yet we can do so with an instant-message chat session; we cannot ordinarily choose *not* to remember a conversation, yet if our greatest memory of that conversation is kept in a computer, we may elect to “forget” the conversation permanently. That this remembering / forgetting is conducted in a fundamentally different way from human memory is perhaps not so important as that it is conducted in a context which is particularly skewed away from normative and emotional involvement: the context within which such actions are carried out consists in an environment wherein emotions are suppressed, according to Turkle [Tur84, p. 307]. Extending the self into an amoral sphere, a domain in which moral agency is excluded on multiple levels, would seem to be to suppress certain very human capacities while working in this domain.

If the machine is also held to be capable of agency, and thus potentially culpable, it would seem that even more human agency is diminished: we are, after all, not responsible for our actions if “the system” has had the final say. It is difficult to locate moral agency in this. We have assigned agency to the machine, which is only capable of accepting agency in a very limited sense; it is not capable of making a moral decision, yet moral agency has been shared with it.

Moral agency is given to the machine, but the machine is incapable of accepting that agency in an appropriate sense. We are wrong for ascribing this agency, and wrong for relinquishing it . . . yet we feel differently: we feel perfectly justified in asking the machine to take responsibility for the maintenance of electronic records, and do not perceive (or care) that this delegation necessarily places any destruction / retention on a wholly amoral basis.

Stone [Sto99] takes this worry a step further, in examining just how deeply people's perceptions have been changed by their relationship with technology items. "The change in the permeability of the boundaries between nature and technics [...] does not simply mean that nature and the technics mix—but that, seen from the technical side, technics become natural [...]. In technosociality, the social world of virtual culture, technics is nature [Sto99, p. 92]." We do not relate to items of technology as created items in an important sense; rather, we relate to them as if they are *right* in the way a tree or a bird is *right*. We, therefore, do not consider our relationship with technology artefacts as if they were artefacts; we privilege technology artefacts by considering them *natural*, therefore beyond scrutiny for their behaviour: they are "just doing as they were programmed." We have ascribed moral agency to an amoral agent (the computer), and have simultaneously absolved it of any responsibility to carry out that moral agency. It appears that we have undergone a net loss of moral agency.

## 2.3 Conceptualising Digital Objects

"Many of us hold on to objects having special personal significance, tucking them away in our sock drawer, or in shoe-boxes in the closet or the attic. in an informed and possibly unselfconscious way, we maintain a personal archive, a treasure chest of cherished artefacts and the memories they hold for us [Lev01, p. 96]."

Just as we delegate more decisions and more agency to the computer, we also possess more digital objects which are precious to us, or which are critical to our businesses, or merely need to be accessed for some reason or other. These objects are not different *in content* than objects in the physical world; rather they are different in that we are required to go through a series of delegates in order to access them, with each viewing being more akin to a performance, with the computer following the script to generate something which resembles a familiar form.

Levy [Lev01] argues that the form of a thing conveys a whole series of cues to the perceiver as to that object's purpose.

"What's more, this conventional form carries a conventional content [...]. Nowhere on the [document], however, does it actually state that this *is* a [document of a

certain type] [...] nor does it explain any number of other facts that are immediately available to us. It doesn't have to. It can simply rely on the fact that we, literate members of the culture, have already acquired the skills needed to recognize them [...] [Lev01, p. 18].”

This argument—that people respond to things based upon their form as well as their content—has strong implications for digital objects, in particular because the *form* of digital objects is not fixed, nor has it been long-established; those digital objects which are most clearly viewed as being *of type: receipt*, for example, are viewed as such via mediation of some rendering process; this perception is not immediately accessible to the user of the computer, but that user must open the digital object using some appropriate software package, and trust that the software has rendered it in such a fashion as to allow for the judgement that the object is *of type: receipt*. This is perhaps tangential or in direct opposition to Levy's argument later on in the work, but is important for our understanding of the individual's relationship to their computing environment.

Levy draws upon Bruno Latour's term 'delegation' to explain “this process of handing off jobs or tasks to others[Lev01, p. 26].” For Latour, this delegation process involves handing off responsibility to both animate and inanimate objects; the example Levy likes to call upon is that of the sales receipt. For Levy, this process is not as simple as it was for Latour. “We humans may make artefacts and delegate certain tasks to them, but they in turn shape our behavior, in effect delegating to us as well[Lev01, p. 27].” Levy recognises the reciprocal relationship we have with our objects: we pass responsibility to them, and they in turn ask that we understand the nature of that delegation. It is a process which obligates (for common items such as the sales receipt) each person who relates to such items to learn at least the rudiments of their form, to understand the conventions used in their own subtle languages. We understand their form and function merely by glancing at the object, even if we are too far away to read the thing itself.

For digital objects, the same cannot be true, at least not with the same rigidity. Let us take the layout and form of a thing to be the object's metadata—data which shouts out, “I am a sales receipt!”—and its content to be some number of lines consisting of brief descriptions, some figures, and a total, possibility with a business name, a date, etc. It is possible, in the analogue world, to directly apprehend that an item is a sales receipt merely by accessing its metadata from across the room. Digital objects, conversely, usually only present to us the bare rudiments of metadata: they may show us their file name, perhaps a few details about their creation, and their location upon whatever media is being examined. Digital objects only present to us the metadata when we request it in some manner, the most usual manner being to actually open

the file and look at it, in which case we are accessing its content (including any such metadata) simultaneously.

“The simple fact is, we humans have found a way to delegate the ability to speak to inanimate objects, and have become deeply dependent on them for an endless array of services[Lev01, p. 33].” Levy is concerned that we have delegated our ability to speak to these objects; perhaps we should wonder why these objects have not been constructed in such a way as to allow these objects to speak to us in a more coherent manner, in accordance with the way in which we are accustomed to apprehending things in the physical world.

Why are these objects—only perceptible via some technological mediator—not organised in some coherent manner for us, rather than our having to struggle to recall where we placed an item of interest? Levy asserts that “[i]t is the rare person who isn’t somewhat traumatized by the state of his or her desk[Lev01, p. 122].” It is also possibly the rare person who isn’t traumatized by the state of his or her computer file system. The point, though, is that although our desks may be incredibly messy, we are still able to locate *a sales receipt* easily amongst the clutter, primarily because the physical world has evolved to grant us such affordances<sup>4</sup> and conventions as we require to be able to perform such basic tasks. The digital world seems to be lack some of the same type of affordance (or, at least, the affordances provided by the digital world are simply the equivalent of those found in the analogue world), while simultaneously allowing for us to delegate even more of our agency to it than we had before. “For there is, it seems to me, a degree of anxiety embedded in all our attempts to order and organize, to control the world around us[Lev01, p. 128].” Perhaps we are correct to be anxious.

## 2.4 Sociology Including Objects

Latour [Lat05] argues for a sociology of *things*, or at least for a sociology which considers things as actors. Latour states that, “[...] *[A]ny thing* that does modify a state of affairs by making a difference is an actor [...] [Lat05, p. 71].” Latour argues that, “[i]n addition to ‘determining’ and serving as a ‘back-drop for human action’, things might authorize, allow, afford, encourage, permit, suggest, influence, block, render possible, forbid, and so on [Lat05, p. 72].” Thus, we may wish to include “things” within our consideration of the social context within organisations, because the particular “things” under consideration in this study (computers and operating systems) most certainly do affect decisions in precisely the ways cited by Latour.

---

<sup>4</sup>The term ‘affordance’ is borrowed from product design engineering [Nor02]. An ‘affordance’ is a feature which indicates, to the user of an object, that actions should be performed in a certain manner. An example of affordances in the physical world would be push-bars on doors which open away from the individual passing through the door as compared to loop-type handles which are clearly meant to be pulled towards the door user.

We run into a problem, however, when we consider computers and operating systems as “actors,” in that, while they certainly do affect decisions, they cannot be said to be actors *in the same way* as human actors. In particular, Latour states that “. . . [A]ctors are always engaged in the business of mapping the ‘social context’ in which they are placed. . . [Lat05, p. 32].” Turkle proposes that people relate to computers *as if* they were proper actors<sup>5</sup>, making allowances for the limited way in which said actors are able to function within the broader social context. Turkle resolves this lack on the part of computer-as-actor by offering that the computer is an amoral actor—merely an extension of the self. This would seem, however, to lessen the agency posited by Latour: if the computer is merely an extension of the self, we should expect that its agency would be subsumed by that of the individual using the computer, rather than affecting decisions in the way that they do; we would expect that computers would not “authorize, allow, afford, encourage, permit, suggest, influence, block, render possible, forbid,” etc.

But, of course, computers and operating systems most certainly are possessed of agency in this way, while simultaneously failing to engage with the social context as would an actor possessed of true agency. Stone argues that we relate to items of technology as if they were agents in the same way that nature is an agent: computers and operating systems affect decisions without engaging with the moral aspects of the social context; they behave according to their own rules, within the social context, yet without considering anything beyond their “built-in” nature. The problem with this, however, is that there seems to be some loss of agency: if computer users are relating to their computer systems *as if* relating to an agent found in nature (e.g. a guard dog, say, or a horse), we do not expect that natural agent to be possessed of agency in the same way as computer systems have agency, i.e., in just such a way as proposed by Latour. Rather, we expect those truly natural agents to “map the social context” of their situations, yet we do not expect them to engage with objects and the content of those objects in the way in which computers do, nor do we delegate to those animals in the same domains and to the same extent as we do computers. We may expect to delegate our agency in guarding property to the guard dog, but we do not view the dog as an extension of the self; further, we do not expect the dog to differentiate between objects which may be legally or ethically sensitive, we do not expect the dog to faithfully carry our communications to distant locations, we do not use the dog to conduct financial transactions, etc.

This is particularly problematic when considering computer systems to be an extension of the self *a la* Turkle. If human agents regard computers as extensions of their selves, yet computer systems possess agency, then human agents are not only delegating some of their agency to the computer system but are accepting that the computer systems’ mediation are

---

<sup>5</sup>See section 2.2 on page 10

somehow correct or desirable; in extending the self into objects possessed of agency, the human being deliberately accepts a modification of the self in ways which may or may not be desired. Indeed, this modification of the self via the incorporation of technology—resulting in a *cyborg* conceptualisation of the self—will be the focus of the next section.

## 2.5 Computer / Cyborg as Mirror of Self / Role Model

To accept the computer as an extension of ourselves is to conceptualise ourselves as cyborg: a mixture of humanity and machine. Whether we perceive ourselves as cyborg or as merely augmented humanity depends upon the degree to which we have accepted the change to ourselves, upon the degree of integration of the changes, and whether we perceive these changes as threatening.

Cyborgs provide a representation of the world which is specific to a moment in time (synchronic), and appear “during periods of radical social and cultural change [Gon99, p. 270].” Gonzalez argues that, when people’s self-conception cannot accommodate new concepts or technology, the cyborg arises in art and literature as a hybrid between humanity and technology. Both extremes are represented *in extreme*, pointing out the differences between the two elements of humanity: the exaggerated form of “true humanity,” and the extreme “non-humanity.” Gonzalez’ analysis considers artistic representations of the cyborg, moreso than media / entertainment representations, or technological implementations / imitations of humanity.

“The image of the cyborg has historically recurred at moments of radical social and cultural change. From bestial monstrosities, to unlikely montages of body and machine parts, to electronic implants, imaginary representations of cyborgs take over when traditional bodies fail. In other words, when the current ontological model of human being does not fit a new paradigm, a hybrid model of existence is required to encompass a new, complex and contradictory lived experience. The cyborg body thus becomes the historical record of changes in human perception. One such change may be reflected in the implied redefinition of the space the cyborg body inhabits [Gon99, p. 267].”

So, modern representations of the cyborg should be encouraging us to examine which changes, in particular, are being proposed in representations of the cyborg, in order to identify current areas of psycho-ontological conflict.

For today’s technology marketers, “cyborg” is a positive term<sup>6</sup>. Cyborgs in the marketplace, then, point out that it is desirable to be machinelike in (at least) one’s playing of games, engagement with sports, or interaction with other machines. To be able to perform repetitive tasks with speed and dexterity, to augment human capacities by adding machinelike attributes or machine enhancements: these are areas of tension, where humanity is wishing it were different / better / more other / less flesh. It seems that, at least in certain areas of interaction with technology, people wish they were more like machines or were endowed with machine capabilities.

Balsamo [Bal99] explores the role played by the cyborg in fiction, concluding that, in blurring the distinction between human and machine, cyborgs act to reinforce *differences* between human and machine. She finds this also true with regards to gender roles, with the cyborg producing / reproducing strong gender stereotypes. Cyborgs also tend to present as either purely good or purely evil; thus, cyborgs are polarising figures, rather than unifying figures. Rather than viewing the cyborg as humanity seeking to incorporate some change, she sees cyborgs as actively resisting change, as backlash against technology.

The distinction between human and machine parallels the distinction between human and computer, as discussed by Turkle [Tur84]<sup>7</sup>. This distinction reflects fears about the nature of humanity, that humanity seems to want to define itself in opposition to technology, rather than as incorporating it as a part of human nature. We make technology, yet we fear it to the point where we want to exclude it from our true selves. This is a paradox, surely, and one which seems emergent: the cyborg points not *simply* to a fear of technology, but a fear of ourselves being *affected* by the technology, and on a massive scale. Simultaneously, the cyborg reflects taboo desires: to be a better killing machine would certainly be perceived as an unacceptable trait in a person, but for a person to wish to be more machine-like in general, particularly when playing games, puts the desire at one remove; it is self-deception, on some level, a mind-trick to allow the games player to feel better about socially-unacceptable desires.

Whether we’re resisting or attempting to incorporate technology into our conceptions of self, examining the tension surrounding the image of the cyborg would seem to provide any number of questions as to how people interact with technological artefacts. For the current study, most relevant would be, perhaps, to find instances in which cyborgian ideations bleed over into behaviour with regards to relating to computers in general.

---

<sup>6</sup>A quick internet search for products yielded the following *cyborgs*: the Duratec Cyborg bicycle; Casio’s Cyborg 009 wristwatch; Cyko’s Cyborg wheel-rims; the Mitsubishi Mirage Cyborg-ZR automobile; Fox Racing’s Cyborg Backpack; the Saitek Cyborg Mouse, Keyboard, Gaming Headset, Rumble Pad, and Joystick; the Virtue Cyborg Board [a circuit-board for rapid-firing of paint balls from Makdev’s Cyborg paint ball gun]; the Black Diamond Cyborg clip crampons and ABS plates; the M2R Genesis Cyborg motocross helmet; the Cyborg terminal extensions project at cyborg-terminal.SourceForge.net; the Green Cyborg Area 51 laptop from Alienware; and countless robotic children’s toys

<sup>7</sup>See section 2.2 on page 10

## 2.6 What brings about moral behaviour?

Let us step away from questions of identity and into questions of practical morality. In particular, let us consider what causes moral behaviour; more particularly, what induces people to engage their moral sense in a given context. This is relevant to the study because morality is at some level concerned with duty to oneself and to others. If we are to understand how and when agents engage their concern for others in the course of decisions to delete or preserve records, we must understand something of morality, of agency, and of will. We need to understand whether morality is flexible, engaged in some contexts and not in others. We need to understand what, within an organisation, causes people to consider questions of duty, and to engage on that very human level with their co-workers and with their responsibilities.

We are not as concerned, here, with whether agents have surrendered their will to some political structure<sup>8</sup>. Although related, philosophers considering choice and free will are primarily concerned about whether one *ought* to be bound by political (or organisational) strictures; they do not consider whether people *feel* themselves to be bound by rules and whether, feeling bound by a particular rule in a particular circumstance, they do not feel any need to evaluate the situation further. Those concerned with the use of the legal system to affect control over individuals or groups of people<sup>9</sup> are not concerned with *exercising* one's will within a social context, nor are they concerned with understanding the phenomenology of subjecting one's will to authority<sup>10</sup>, both of which are of more interest to this study. First, let us consider the sociological aspects of action within an organisation.

### 2.6.1 Social Control

Vaughan [Vau98] is concerned with *control* of individuals within an organisation: she wants to understand what it is that causes people to behave in a certain way, ethically, within an organisation. Vaughan demonstrates that, essentially, *punishment does not work* to control individuals within organisations. She presents the example of the manager as *amoral calculator* as an inevitable result of the application of rational choice theory; this model assumes rational intent in any decision to act—an unsupported premiss in the argument. Vaughan argues that there is an insufficient understanding of the decision-making process in cases of normatively blameworthy action. This study suggests that right records-management decisions will be shaped by the social structures of the organisation, and that the norms within the

---

<sup>8</sup>For a more detailed philosophical investigation into submission of the will to authority see Edmundson [Edm93]; Murphy [Mur97]; Raz [Raz75]; Raz [Raz85]; Raz [Raz90].

<sup>9</sup>For a more detailed discussion of the affects of legislation upon digital objects and, through regulation of digital objects, regulation of individual freedoms see Lessig [Les06, p. 31-138].

<sup>10</sup>See Kim [Kim00] for a discussion of digital phenomenology and ontology of digital being.

organisation override (most) other normative commitments. The question becomes, How does the organisation shape the decisions, and how does the technology shape the decisions?

“One project showed that legalistic deterrent strategies can have negative unintended consequences: [...] in some cases a deterrent regulatory posture actually reduced compliance [Vau98, p. 28].” This reduced compliance may be in part because those violating the regulations are actively resisting the imposition of control from outside the organisation, or it may be because transgressions are very exceptional and have little effect upon the careers of those violating the regulation or policy [Bov04, p. 63]. The paradox of regulation causing a decrease in compliance may also be due to the fact that individuals within a regulated context are no longer making judgements of the same kind, but are applying an external decision-making system rather than an internalised normative system.

“[T]he laws, rules, and administrative regulations designed to guide organizational behavior are likely to be *mala prohibita* [wrong only because it is prohibited] rather than *mala in se* [wrong in and of itself]; thus the standards to which organizations are expected to adhere and the consequent punishments are not clear to either organization decision-makers or the public [Vau98, p. 29].” What is the psychology involved in relating to a system in which the norms are not those of morality, but are those of regulated behaviour? If most rules of the organisation are *mala prohibita*, it would seem that behaviour would be governed less by one’s own conceptions of morality than by an understanding of the applicable regulation. This would seem to argue that, within an organisation, individuals are behaving much more in a manner most beneficial to themselves within a rule-dictated system (in pursuit of a Nash equilibrium) than they are behaving as moral agents.

Bovens [Bov04] makes an even stronger argument against individuals being seen as acting in normatively blameworthy ways within complex organisations<sup>11</sup>:

“There are strong indications that the moral barriers that in private life help prevent injury to others, function much less stringently when natural persons are operating within the framework of a complex organisation. Natural persons seem to be much less sensitive to those internalised moral norms when, and in so far as, they function within an institutional and hierarchical framework (Milgram 1974; Robbins 1980: 237-9; Finney and Lesieur 1982: 275-6; Meeus and Raaijmakers 1984; Jackall 1988; Kelman and Hamilton 1989; Vaughan 1996) [Bov04, p. 68].”

---

<sup>11</sup>For Bovens, complex organisations are simply organisations which are large enough and which involved shared responsibility for decision-making. They are organisations which present the “paradox of shared responsibility. As the responsibility for any given instance of conduct is scattered across more people, the discrete responsibility of every individual diminishes proportionately[...] [Bov04, p. 46]”

It is unclear whether the “hierarchical framework” is causative of this lack of sensitivity to “internalised moral norms,” or whether relating to a complex organisation is sufficient to cause such lack of sensitivity. Bovens allows that simply functioning within the norms of a complex organisation may be sufficient to cause such lack of sensitivity to both internal norms and to external regulation, whether or not the organisation is structured as a hierarchy [Bov04, pp. 71,101-2]. Bovens does point out, though, that the mobility of executives between different organisations allows them to escape responsibility for their decisions [Bov04, p. 110] and that hierarchy generates social and peer-group pressure to conform [Bov04, p. 125] contrary to individually held norms.

## 2.6.2 Dissemination of Ethical Standards

Glinow and Novelli [GN82] made a study of professionals involved in academic publishing; specifically, they considered editorial staff and review board members of a series of journals, all of which journals are academic publishers within the domain of business management and Organisational Behaviour. Their study examines the ethical structure within academic publishing as a profession. They determined that there is “(1) substantial dissensus about whether standards exist; (2) vague or nonexistent mechanisms for communicating standards; and (3) few institutionalised penalties for violating ethical standards [GN82, p. 417].”

There was a degree of consistency in responses within the individual publishing houses, reflecting the unique cultures present within each house. These responses, however, while consistent on the surface, broke down under scrutiny. For example, when considering whether it would be ethical to select reviewers with or without consideration for race, responses tended to break down by publishing house and were fairly consistent. However, in a detailed analysis of related questions, the responses were such as to indicate that there was no real consistency within the organisation: individual responses varied significantly with respect to issues of affirmative action, discrimination, and other topics involving issues of race.

For this research, then, the normative aspects of information deletion / retention must be regarded as individual, rather than common to the organisation, at least in areas which are without strong, external, professional organisations (such as may be found amongst Human Resources professionals, for example). It seems also that, while there may be a common consensus that certain behaviours are of normative significance, dissemination of ethical standards is in no way consistent. This is of vital importance to the research topic, because it means that, while there may be “industry best practise,” the application of that practise must necessarily end up subject to massive variance according to individual knowledge and ethical

norms.

Glinow and Novelli [GN82] point out that there is no educational mechanism for ethical norms, nor is there any consistency in their application. Vaughan points out that external controls do not work for enforcing norms [Vau98, p. 29], yet points to self-interest and the establishment of Nash equilibria as perhaps a controlling mechanism at play. This brings us to Games Theory, and the examination of the formation of just such equilibria.

### 2.6.3 Normativity? Or Merely Nash Equilibria?

Rapoport [Rap63], using an extremely long sequence of individual “plays” of the Prisoners’ Dilemma game (300-700 individual plays), arrives at a behaviour pattern in which players tend to lock into one of two strategies: players either cooperate (by choosing that which guarantees some degree of success to both players), or players engage in some series of playing the riskier (yet probabilistically preferable) choices. 65% initially chose to cooperate. The probability that the players would cooperate on the next play, following a “chance” instance of cooperation, was 0.6, as was the second play, with a steady rise in this probability as the number of cooperative responses grew. A similar pattern emerged amongst those which chose to compete. This “locking in” effect continues up until a probability of 0.9.

Lock-in is quite interesting to the topic at hand, primarily because it may tend to shape behaviour in records management, and within an organisational structure: for example, in organisational cultures which have established an adversarial relationship with an audit group, or with their regulatory body, we might see this type of lock-in behaviour, wherein each party is engaged in an activity which is either cooperative or competitive. Thus, records management behaviour would tend to be carried out accordingly, with those in the adversarial relationship perhaps presenting us with the most interesting cases for study, if only because they might be so aberrant. These types of cases present us with an opportunity to observe the game theorists’ Nash equilibria in effect in a real-world situation.

A “Nash equilibrium”<sup>12</sup> is a state in which decision-makers have each reached an optimal decision with regards to their own desired outcome—a state in which no individual may obtain a more desirable outcome simply by changing her behaviour. This does not necessarily entail an optimal outcome for *all* participants (and may, in fact, entail suboptimal outcomes for *all* participants). The presence of a Nash equilibria also does not entail the establishment of any normative state—it merely indicates that the individuals involved in the situation have reached

---

<sup>12</sup>“Nash equilibria” are usually referred to within the context of games theory, as they involve competitive situations. They have broader implications, however, than merely games—for empirical purposes, however, games are often used to test various theories with regards to behaviour under such situations.

a stable state of behaviour.

Nash equilibria are important to our study primarily because we may encounter such equilibria with regards to decisions to delete or retain: individuals may shape their behaviour to optimise their own desired outcome, based upon a desire to maintain the status quo with regards to whatever balance is in place within their organisation. In doing so, their decisions would be formed from a desire to help or hinder others, perhaps, or in order to maintain their position relative to their peers. This type of a decision-making process would be expected to generate different retention decisions: decisions would be made based upon how one's position would be served by the presence or absence of digital objects, or upon whether the destruction or retention were thought to result in criticism.

Rapoport made another observation in the course of his study, which was that the participants tended to become “contaminated” by whichever game they had played first, carrying over their behaviour into subsequent variations of the game. They had “learned how to play the game” and are bypassing that learning (and non-automatic) portion of their minds. For our study, this might indicate that retention behaviour would be shaped by the individuals' earlier experiences of the workplace—that their knowledge had become unconscious.

#### 2.6.4 Ethical Relativism?

Jones [Jon91] asserts that basically organisations serve to insulate from the outside world, so that agents feel more distance from the effects of their moral judgements, and that organisations alter the moral field in that the agents are part of whatever culture is present within the organisation and, consequently, may make different moral judgements than they would in other contexts. For Jones, as for Vaughan [Vau98], normativity is necessarily relative to the culture in which the agent is acting. For Jones, though, a ‘moral issue’ holds special status, in that it exists independent of whether the agent is aware of it *qua* moral issue, and extends across different cultures. In this way, Jones would have some non-relativistic measure of morality, while still allowing that agents recognize moral issues differently (if at all), dependent upon the (organisational) culture. However, he excludes the organisation from having any role in the agent's perception of moral issues, having any role in the formation of motivation to behave in a moral manner, and from playing a part in defining moral issues *qua* moral issues. So, while Vaughan was concerned at the pragmatic aspects of bringing about moral behaviour within an organisation, Jones is interested in what affects the ‘moral intensity’ in a given situation, although he does not go very far in developing this. It seems that both Vaughan and Jones are interested in the same thing, in the long run: bringing about moral behaviour; they simply

differ in their approach to the issue.

Jones points to previous research which demonstrates that “[...] managers will use lower levels of cognitive moral development in actual work environments compared to hypothetical situations, such as those found on tests designed to measure moral development. These authors have suggested that moral development, or at least the levels at which people actually reason, may be context dependent [Jon91, p. 384].” Is the observation, rather, that people are capable of making a moral judgement without actually committing themselves to that moral judgement; whether “the right thing” is known to them, along the lines of something which is learned to be wrong, rather than something which they have reasoned out to be wrong? This would seem a neatly related issue to earlier considerations of rules superseding moral judgement, and to provide much for future research into practical morality. For our study, however, we are not concerned with whether the moral agent *feel committed to* a particular moral judgement, nor whether they have evaluated the situation using learned rules or standards, but whether they *consider* the situation in terms of duty or morality *whatsoever*. Whereas with Rapoport and with Vaughan the agent is operating according to the norms of the local culture, Jones provides us with a situation in which the agent is morally accountable to behave in a certain way, but just does not.

Webster and Trevino [WT95], like Jones, attempts to characterize ethical decisions taking place within the context of an organisation. Unlike Jones, however, Webster and Trevino separates individual variables from situational variables. She asserts that the agent (depending upon their moral development) thinks about moral issues in a stable manner, independent of context, at least in their initial responses to such issues. Because the agent’s moral development is not sufficient to predict behaviour, Webster and Trevino expands the field to incorporate context as a factor affecting the decision to act. In a way, Webster and Trevino is concerned with action, where Jones is concerned with motivation. These stances may not be entirely incompatible, but are simply concerned with different aspects of the same problem.

“The individual’s cognitive moral development stage determines how an individual thinks about ethical dilemmas, his or her process of deciding what is right or wrong in a situation. However, cognitions of right and wrong are not enough to explain or predict ethical decision-making behavior. Additional individual and situational variables interact with the cognitive component to determine how an individual is likely to behave in response to an ethical dilemma. Three individual variables, ego strength, field dependence, and locus of control, are proposed to influence the likelihood of an individual’s acting on cognitions of what is right or wrong [WT95,

p. 602].”

Regardless of whose model is correct (whether that of Webster and Trevino or Jones), each attempts to explain why individuals operating within an organisational context appear to alter their moral behaviour.<sup>13</sup> Each would attempt to preserve some fixed-ground, as far as the rightness or wrongness of an issue being consistent regardless of the context. They both regard the failure of the individual to recognize or act upon that moral rightness as something to be explained, rather than considering the morality of action to have changed based upon the context.<sup>14</sup> Both Jones and Webster and Trevino have referred to Kohlberg, 1969. The paper referred to is one which outlines the process of an individual’s moral development. Kohlberg asserts, in this paper, that an individual progresses through the stages of moral development and that each progression is a one-way transition: individuals do not revert to an earlier stage of moral development. This is a very strong assertion, and might be called into question by the presentation of different normative standards depending upon context. Such possibility is not presently at consideration, however. Of interest to this study, Kohlberg developed a means of assessing the moral development of individuals. This might have been of interest in the development of survey questions for this study: i.e., how morally-developed is each survey respondent, and is that reflected in their responses? Webster and Trevino mentions a standard method of assessing moral development, developed by Kohlberg, called “Standard Issue Scoring;” Webster and Trevino also points out that Gibbs & Widaman developed a “Social Reflection Questionnaire” which is “easier and less time consuming to administer and score while retaining the qualitative nature of interview responses [WT95, p. 606].” This questionnaire may prove valuable should this study be extended to include a stronger emphasis upon moral judgements.

Habermas [Hab03] expands upon the relatively simple conception of moral development offered by Jones and Webster and Trevino by including the concept of a norm needing to be “warranted” for a particular situation: “[T]he norm that is ‘appropriate’ to the situation is selected from the plurality of warranted norms that might be applied in any given case. [Hab03, p. 245]” Further, Habermas argues that “[W]e discover the rightness of moral judgments in the same way as the truth of descriptions: through argumentation. [Hab03, p. 247]” Thus, not only

---

<sup>13</sup>Certainly, individual behaviour changes from one context to another—that is the problem! The question, though, is whether individuals are setting aside their morality, whether their morals change from context to context, or whether they have somehow had their moral sense paralysed by the features of their environment: that they *would* behave morally, they just do not perceive the requirement to *make* moral judgements.

<sup>14</sup>Consider the Humean Theory of Motivation [Hum89], Rule Utilitarianism *a la* Hospers [Hos89], and Virtue Ethics *a la* Foot [Foo89]. Hume would argue that an agent, having made a moral judgement, would necessarily feel motivated by that moral judgement

If, in the workplace, agents are not motivated to act according to their own professed moral standards, then something must be blocking their exercise of moral judgement Consider Raz and his belief in *blocking reasons*, or Murphy in his belief that individuals surrender some portion of their judgement when consenting to participate in a political structure. [Mur97]

are there potentially multiple applicable norms potentially applicable to a given situation, but such norms must be examined in order to determine their rightness. Normativity, then, is not a simple matter of applying one set of rules to every context, nor it is a simple matter of knowing which set of rules is active within a context and applying those rules. Rather, the making of moral judgements require a different level of intellectual engagement than, say, the application of an organisational policy; this may go some way towards explaining where the social control worried over by both Vaughan and Bovens causes the paradoxical result of a lack of normative engagement.

### 2.6.5 Thinking Individuals within Organisations

We cannot ignore the expectation that there be someone to somehow be held to account for decisions made within complex organisations. Arendt [Are03] considers one of the more horrendous sides of issue in considering Adolf Eichmann's seemingly contradictory personal life, in which he attended church and was apparently a good husband and father, while simultaneously being "the architect of the holocaust" as his vocation. Arendt found this to be a common feature of those who worked within the Third Reich: "[T]here were very few people even in the Third Reich who wholeheartedly agreed with the late crimes of the regime and a great number who were perfectly willing to commit them nevertheless [Are03, p. 35]." For Arendt the resolution to this conflict is to reject collective guilt or innocence [Are03, p. 29] and to point to bureaucracy as one cause of such a situation:

"[I]f one wishes to define bureaucracy in terms of political science, that is, as a form of government—the rule of offices, as contrasted to the rule of men, of one man, or of the few, or of the many—bureaucracy unhappily is the rule of nobody and for this very reason perhaps the least human and most cruel form of rulership. [Are03, p. 31]"

Arendt might certainly find the same type of management in the complex organisation as was found in Nazi Germany; certainly the large, multi-national organisation of today, while not even approaching such massive numbers as were employed by the Third Reich, are likely to be similar in terms of organisational complexity and opaqueness in the decision-making process. Likewise, large organisations are typically not managed by democratic process, but by fiat.

That said, however, we are still left with the feeling that somehow someone ought to be in a position to be held responsible for corporate decisions; at the very least, we expect the individuals within complex organisations to be aware of their actions, the possible results of

those actions, and to act responsibly. Arendt would agree with this feeling as, for her, individual responsibility was the only responsibility which made logical sense [Are03, p. 35].

The question for us, however, must concern whether the individual within the organisation is operating within the legal system or within some system of morality. “[T]he almost unanimous assumption of moral philosophy throughout the centuries stands in curious contradiction to our current belief that the law of the land spells out the essential moral rules upon which all men agree, either because God told them so or because they can be derived from the nature of man [Are03, p. 102].” For Arendt, the answer is clear: that the “law of the land” comprises the moral system within which individuals operate. This is difficult to accept, however, in light of, e.g., Bovens [Bov04]; Glinow and Novelli [GN82]; Jones [Jon91]; Rapoport [Rap63]; Vaughan [Vau98]; Webster and Trevino [WT95]. That is not to say that Arendt is wrong, *per se.*, but that she appears to be approaching the issue from the perspective of legality—despite examining what is, certainly, a moral issue—rather than from the perspective of human psychology. While she addresses Eichmann and the horrors of the holocaust, she considers the individuals within the Third Reich to have been almost universally defective in some manner, rather than as individuals operating within a complex organisation which served to alter the moral landscape in such a way as to allow for the commission of great evil.

Arendt does grasp something central to the discussion here, however, in her consideration of conscience:

“Conscience supposedly is a way of *feeling* beyond reason and argument and of knowing through sentiment what is right and wrong. What has been revealed beyond doubt, I think, is the fact that such feelings indeed exist, that people *feel* guilty or feel innocent, but that alas, these feelings are no reliable indications, are in fact no indications at all, of right or wrong [...] . [T]hese feelings indicate conformity and nonconformity, they don’t indicate morality. [Are03, p. 107]”

This is a very critical point when considering the behaviour of individuals operating within complex organisations, in particular if feelings of guilt are brought about in response to the predominant normative structure within the organisation, rather than as a response to some external normative structure: individuals may indeed be operating within a moral structure (i.e., that of the organisational culture), and may respond to that moral structure with feelings of guilt or innocence. Those feelings, however, are not an indicator as to whether, on some absolute scale or on appeal to an external scale (perhaps legal, perhaps ethical) the individual will be judged as having done *right*.

The question of conscience becomes confused within complex organisations, however, due to the diffuse effect of action and any rewards for unethical behaviour. Wiltermuth [Wil11] demonstrates that individuals are more likely to engage in unethical behaviour if the benefits are shared with others: “If people are able to attribute a part of their motivation for acting unethically to benefiting others, they may reduce how unethical they perceive their behavior to be. As a result, they may be less likely to condemn themselves for behaving dishonestly [Wil11, p. 159].” Within a complex organisation, then, we might imagine that we would find that unethical behaviour which benefits the entire organisation (and benefits the individual only in proportion to the others within the organisation) is likely to be not uncommon whatsoever, particularly under what Wiltermuth terms ‘elastic’ situations, i.e., situations in which the facts of a situation are “vague and uncertain.” Wiltermuth maintains that individuals wish to perceive themselves as ethical, yet will behave unethically if the situation is ambiguous and if the benefits of behaving so are shared with others. Thus we may expect to find unethical behaviour even within complex organisations with an organisational culture which encourages ethical behaviour.

According to Wiltermuth, such unethical behaviour appears to be something which is at least partially determined by each individual’s personal moral framework, with “people who score highly on Machiavellianism, possess an external locus of control, and hold a relativistic instead of an absolutist moral philosophy all have been identified as people who are particularly likely to engage in unethical behavior [Wil11, p. 157].” With regards to Arendt [Are03] we may wish to consider that the individuals working within the organisation of the Third Reich certainly may have possessed such an “external locus of control” in the sense that they were operating within an organisation which controlled the actions of its members through a great variety of means; however, we may also consider that these individuals would have been operating within an ethical framework which constituted such an “external locus of control” in the form of the deontological system put forth by Kant. Thus, within the Third Reich, not only were individuals *encouraged* by their organisational culture to commit unethical actions, but their individual tendencies towards unethical behaviour were supported by any ambiguities in the situation, by the desire to share benefits with others within the organisation, and by multiple external loci of control.

Arendt examines a special—and especially horrendous—case, and argues from that case that something was fundamentally flawed within the individuals who acted in support of the Third Reich. This, however, overlooks the influences upon the individuals faced with such a confluence of circumstances; that is not to argue that such individuals *ought* to have acted differently, but that such a conclusion overlooks the larger picture—it concludes too quickly, and so misses the possibility that something about the situation could provide insight into how to prevent such

from occurring again. Wiltermuth provides the starting point for the consideration of such possibilities.

For the purposes of this study, we must consider that individuals may be making decisions within complex organisations. The organisational culture of these organisations may encourage unethical behaviour, or not—it would be quite difficult to evaluate without extensive study of an individual organisation. However, in accordance with Wiltermuth, we may regard individual decisions with regards to at least certain digital object as being carried out within an “elastic” situation; thus, the individual may be encouraged by ambiguity and possible gains to behave in an ethically dubious manner. Wiltermuth [Wil11, p. 159] further complicates our evaluation by presenting a case wherein individuals justified their behaviour by stating an alternate reason for their behaviour; thus, any evaluation of reasons given for a particular action may be suspect should the individual being studied consider that such a reason would be perceived as negative by the researcher.

## 2.7 Difference between Ought and Action

As an alternative interpretation of what’s going on when people appear to violate their own standards of behaviour, Redelmeier and Tversky [RT04] find that it appears physicians make *ought* statements which routinely deviate from actual practice; that is, their stated beliefs as to the appropriate treatment do not coincide with their actual practice. Redelmeier and Tversky maintain that this is generalisable to other situations and to other professions. This is slightly worrisome, as it may mean that any study results considered will necessarily deviate from actual practice to some greater or lesser degree; study results may give us an idealized version of what the individuals would have us believe about their actions [see also Wil11].

Redelmeier and Tversky further find that, physicians “make different judgments in evaluating an individual patient as compared with considering a group of similar patients [...] and that lay people also make this distinction [RT04, p. 887].” This is a special case, applicable to the study of duty to others, because it would seem that people actively desire that a certain standard of moral behaviour be in place for everyone else, while remaining free to violate this standard as individuals. Redelmeier and Tversky provide no answer for this discrepancy: “The discrepancy between the aggregate and individual perspectives demonstrated in these experiments cannot be attributed to differences in either medical information or economic incentives; hence it is difficult to explain on normative grounds [RT04, p. 891].”

## 2.8 Context and Selection

Shafir, Simonson, and Tversky [SST04] may not have been able to reconcile this difference between what people say they ought to do and what they actually do, but they certainly pointed the way towards resolution, in their studies of choice. In particular, some of their studies of the selection process might resolve this difference.

In these studies, Shafir, Simonson, and Tversky found that people select based upon positive features and reject based upon negative features; because of this behaviour, an *enriched* option will be selected more often *and also rejected* more often<sup>15</sup>. People seek for more options when they are unable to choose between the present options, suggesting that they do not simply order the options and select the one which is more appealing. People will more readily make a choice when there is a “clear loser” among the options. Further, people will choose an *inferior* option if the superior option has been endowed with some feature which is intended to make it appealing, but does not appeal to the subject.

This selection behaviour may present a way forward, in considering both the discrepancies between what people say ought to be done and what they do, and also in considering the differences in behaviour between different social contexts. People necessarily make decisions between presented options, or between known options, rather than evaluating a decision sans context and proceeding to seek out the option which would most closely match their pre-selected ideal. Restated, people do not formulate an ideal prior to selection, against which to compare all options; rather, individuals evaluate only the options presented and base their decisions upon those options. Thus, if an individual is situated in a particular social context, or is making a decision using a particular software package, that individual is choosing between a very limited set of options, and is unable to behave otherwise.

To take Shafir, Simonson, and Tversky’s study of selection further, this type of behaviour suggests that digital objects which are *enriched* will be selected both for retention and destruction more frequently than *impoverished* digital objects. Thus, management activity would seem to be more focused upon a certain class of digital object. If this is the case, it becomes even more important to determine exactly *how* the presentation of the digital object affects the decision to delete or retain, as the factors which are enriching the object may not

---

<sup>15</sup>An example of an “enriched option” as compared to an “empoverished option” might best be considered by presenting an example. Suppose one were considering the purchase of a computer and are presented with two choices. One computer is described as having a certain amount of memory, video performance, and storage. The other computer is described similarly, but also details the various peripheral ports, bus speed, cooling capacity, and so forth. The second computer in this example is the “enriched option”, the first the “empoverished option.” The systems *may* be identical in every way, but the descriptions of the two differ in the details provided for evaluation; the systems *may not* be identical, however, in which case the options are enriched or empoverished in fact, not simply in description—the critical point, however, is that there is a difference in quantity of available information to the decision-making process which results in a difference in perception of the objects.

be desirable in terms of being criteria for selection. It also becomes much more important to understand the different choices presented by different software systems, as those systems are affecting the *possible* decisions, thus dictating the outcome of those decisions.

This is not to say that people do not make decisions in a rational manner: “people do not choose between the equated alternatives at random. Instead, they resolve the conflict by selecting the alternative that is superior on the more important dimension, which seems to provide a compelling reason for their choice [SST04, p. 940].” People, then, choose a particular alternative based upon whichever dimension is more important to them, in any given situation. This, also, provides some explanation as to why people may behave differently when they are situated in different social contexts: the more important dimension, in selecting objects to retain in one’s own life, might be to select those documents which are important to recording a personal history; the more important dimension, in a business context, might be issues of perceived legal requirements. There may certainly be an overlap between obligations, as in individuals retaining files at work because they are important on a personal level, or retaining documents at home because of some legal requirement. However, the “more important dimension” is likely to differ, from context to context.

Shafir, Simonson, and Tversky’s studies on selection and rejection, however, add another problem to our study: the affect upon behaviour of *enriched* and *impoverished* objects. They ask us to “consider two options, an enriched option, with fewer positive and fewer negative features, and an impoverished option, with fewer positive and fewer negative features. If positive features are weighed more heavily when choosing than when rejecting and negative features are weighted relatively more when rejecting than when choosing, then an enriched option could be both chosen and rejected when compared to an impoverished option [SST04, p. 941].” This would certainly seem to present a problem for decision-making in general, and records management behaviour in particular: records management is both selection and rejection (if one takes a decision to delete as “rejection”). Thus, decisions to delete are highly influenced, also, by the information presented to the individual. This ought to produce differences in decisions across records formats, as records held within one type of management system will present different attributes than those within another system. This ought, also, to produce differences in decision in records maintained (or presented) in similar formats: records presented in a common container (e.g., a web browser) are likely to minimize differences between records, thus affecting the outcome of the decision.

Shafir, Simonson, and Tversky [SST04] address difficulties in deciding between very similar options by presenting two cases: that in which the individual defers the decision indefinitely, and that in which the individual’s decision is pushed in one direction or another because of the presence of some clearly superior or inferior option (“asymmetric dominance”).

“[T]here are situations in which people prefer each of the available alternatives over the status quo but do not have a compelling reason for choosing among the alternatives and, as a result, defer the decision, perhaps indefinitely [SST04, p. 947].” This is another case in which we ought to find behaviour affected by context as well as by irrelevant features of the records themselves. This affect ought to produce an increased tendency to retain records, in cases where records present few distinguishing features. It is also likely that decisions made within a work environment—where decisions are, theoretically, more restricted—ought to err towards retention even more frequently than in the home environment. At least one of the aspects which will affect the decision is that of the features presented to the individual; if features presented do not present a compelling reason for considering the possibility of a duty to others, it is likely that such a consideration will not be made.

Alternatively, when an agent is presented with a clear choice, whether that choice is one they would have made in isolation or otherwise, the agent will be influenced by the presence of a clear alternative, and will choose. “The ability of an asymmetrically dominated or relatively inferior alternative, when added to a set, to increase the attractiveness and choice probability of the dominating option is known as the asymmetric dominance effect (Huber, Payne, & Pluto, 1982) [...] . Subjects’ tendency to delay [...] is much greater when they lack clear reasons for [decision], than when they have compelling reasons to [decide] [SST04, p. 948].” In a social context, we are more likely to see this type of a decision-making process, as “asymmetric dominance is enhanced when subjects anticipate having to justify their decisions to others [SST04, p. 951].” We would expect, then, to see decisions to delete or retain being made in significantly different ways, depending upon the social context of those decisions, as well as depending upon the relative availability of information about which to structure the decision.

A variant of asymmetric dominance may be seen in that “the tendency to prefer an alternative is enhanced or hindered depending on whether the tradeoffs within the set under consideration are favorable or unfavorable to that alternative. A second cluster of context effects, called extremeness aversion, refers to the finding that, within a set of options, options with extreme values are relatively less attractive than options with intermediate values [SST04, p. 950].” We will return to extremeness aversion later on, in considering errors in comparison, although it is a closely related concept to asymmetric dominance: it is included there because it fits more closely with *errors in comparison* than it does with simple considerations of how people make choices<sup>16</sup>; extremeness aversion, also, falls into a class of behaviours which have been exploited in one way or another, in order to artificially induce a particular decision, whereas asymmetric

---

<sup>16</sup>See section 2.11.3 for a discussion of useless features (page 40), compromise effects (page 41), and contrast effects (page 42).

dominance is merely a fact about the way people perceive available choices.

## 2.9 Verbal vs. Pictorial Presentation yields differences in decision

In keeping with the idea that some portion of a retention / deletion decision is unconscious, we must also consider that the presentation of a record determines—at least to some extent—the end result. This is perhaps most well-documented when considering the differences between comparisons of similarity and difference when subjects were given pictorial evidence and when they were given verbal evidence. “These findings suggest the presence of two different modes of comparison of objects that focus either on their common or on their distinctive features. In the first mode, the differences between the stimuli are acknowledged and one searches for common features. In the second mode, the commonalities between the objects are treated as background and one searches for distinctive features. The near-perfect separation between the verbal and the pictorial stimuli [...] suggests that conceptual comparisons follow the first mode that focuses on common features while perceptual comparisons follow the second mode that focuses on distinctive features [GT04, p. 123].” We may expect, then, that operations involving classification of digital objects are most optimally performed via the use of conceptual constructs such as taxonomies, while operations involving comparison are most optimally performed via the use of pictorial representations of the object.

Detecting differences between digital objects, then, may be best—or most comfortably—performed utilizing some pictorial representation of the object in question. If this is the case, we might expect the selection of single objects for deletion to rely more upon the pictorial representation of the object or record, if available. Therefore, we may expect to find any outliers in terms of naming convention, file type, etc. to be most likely to be selected, merely because these features are most readily distinguished utilizing the visual centre of the brain. Likewise, we would expect any multiple-object selection to be performed on the basis of locating similar objects, which selection would be best facilitated by the conceptual centre of the brain, and is thus more likely to be accomplished via the use of a taxonomical structure rather than any pictorial attribute. Single records may be more likely to be selected based upon visual attributes, groups of records by verbal or conceptual attributes.

Gati and Tversky state that, “it is conceivable that the difference between pictorial and verbal stimuli observed in the present studies is due, in part at least, to an inherent difference between pictures and words. In particular, studies of divided visual field [...] suggest that ‘the right hemisphere is better specialized for difference detection, while the left hemisphere is better specialized for sameness detection’ (Egath & Epstein, 1972, p. 218) [GT04, p. 127].” Regardless

of the cause of this difference in operation, it is obvious that decisions made based upon visual properties will differ from those made based upon conceptual properties. We may expect to find, further, that objects which are rich in one type of property while simultaneously poor in the other type of property will tend to be selected for retention or destruction in a vastly different manner from one another.

The question becomes, “Are different types of duty or responsibility distinguished via conceptual or perceptual means?” More broadly, “Are conceptions of duty or responsibility generally perceived via visual or linguistic means?” If perceptions of duty to others are emotionally-based, we may be able to make the argument that pictorial representations would be more likely to trigger these types of perceptions to arise. This is, perhaps, a vast jump in supposition; it is, however, possibly true, because of the fact that pictorial perceptions are centred in the same hemisphere of the brain as those of the emotions. If perceptions of duty to others are conceptually-based, we may be able to make the argument that linguistic representations would be more likely to trigger these types of perceptions to arise, for similar reasons: the hemisphere associated with linguistic processing is also closely associated with conceptual processing. In either event, if perceptions of duty to others are based in either centre more so than the other, we might be able to perceive some difference in the types of duty awarenesses given rise to by different stimuli, along the visual / linguistic axis. Thus, it may be possible to demonstrate that different types of records engender different judgement responses (and, perhaps, entirely different *classes* of judgement), based solely upon the availability of different visual / linguistic stimuli.

It should be noted that, while verbal vs. pictorial presentation might be a rich field for study, the study method chosen was unable to address such concerns; this aspect of decision-making deserves study, certainly, but was omitted in order to better serve the aims of the larger study<sup>17</sup>. It is expected that there will be differences in the selection process for deletion and preservation, but this has not been tested in the present study.

## 2.10 Finding Patterns Where None Exist

On the other side of the issue of visual versus linguistic perception, we must consider properties or attributes which are added by the perceiver: properties which do not exist, or which are weighted more heavily than they ought to be, or to which an interpretation has been granted which is not only unsupported by the facts, but is unsupportable by them. We will consider a study by Tversky and Gilovich [TG04]: that of the “hot hand” in basketball. The phenomena

---

<sup>17</sup>see section 3.7 on page 86.

of the “hot hand” is common to basketball, gambling, and no doubt many other cases in which sequences of independent operations are evaluated in an attempt to divine the future. Tversky and Gilovich provide both a statistical examination of how the “hot hand” is false, as well as an exploration of why it might be that people perceive it to be true (desires for patterns, beliefs about statistics in which things have to “balance out,” etc.).

“[P]eople [...] tend to ‘detect’ patterns even where none exist, and to overestimate the degree of clustering in [...] sequential data [TG04, p. 264].” This phenomenon may have more of an effect upon other areas of records analysis than simply on judgements to retain or delete: while the user may be searching for patterns in data or wishing for patterns to exist between files or records, this does not seem to be something which fits very well with the process of deciding to delete or preserve these files or records; it may be an operation more commonly performed during analysis of information for some other purpose (e.g., judgements of risk). We would expect to see a higher incidence of this type of error in the analysis of relational data, in which the desire for there to be discernible patterns is certainly very high and in which the sample size tends to be relatively low (resulting in a high standard deviation); yet, because the data is “all of the data,” the difference between *population standard deviation* and *sample standard deviation* tend not to be taken into account, leading the analyst to an even deeper error with regards to standard deviation, if standard deviation is considered at all. This type of an error, though, would seem to be related more to the *use* of relational data than to its retention, so it may not play any significant role in retention.

We might surmise, though, that perceiving a pattern in data may lead the user to stop gathering data, in which case this cognitive error may lead to less data in quantity *and* in quality. Perceiving data as useful for some diagnostic purpose may also contribute to its being retained; in this way, an error in perception may lead to data being retained when it otherwise might be discarded. If the pattern perceived is viewed as a negative pattern, it may be more likely that the data be erased, particularly if the analyst believes that others may be able to perceive the same pattern. This type of decision—one in which an agent retains or destroys data based upon some perception of a pattern, and the belief that others may perceive the pattern also—is one in which the agent is considering herself and others. Whether the agent considers others within the context of duty or responsibility, however, would indicate that the agent felt motivated to act upon the pattern in such a way as to protect or involve others, which seems to be a different type of situation than simply perceiving some pattern in the data, erroneous or not.

## 2.11 Psychological Issues: Common Errors in Judgement

### 2.11.1 Psychological Issues: Unconscious Decision-making

There are different neural processes involved in learning to perform a task and in performing that task, once learned. The learning process appears to be a generic process within the brain, repeated in different areas depending upon the type of task. However, the process of the learning taking place in one area, to be handed off to another area when the learning is complete, is clearly the way the brain functions. This allows for tasks to be taken out of the conscious realm—requiring attention—and shifted into the non-conscious realm of neural activity, wherein performance of that task does not require any significant amount of attention. The learning centres may be re-engaged in certain circumstances; namely, they may be reactivated if something about the task at hand is novel, requiring conscious attention (See Ito, Miyashita, and Rolls [IMR97]).

It may be that many decisions to delete are made at a non-conscious level, or that they at least take place without much conscious deliberation. Some level of conscious thought is likely to be engaged when examining a digital object for retention or destruction; it is the depth of deliberation which is at question, simply because so much of the process may be “automated” already. Our task with reference to unconscious decision-making must be to adequately characterise this decision process, to understand which portions of it are made on a conscious level, and which portions are left to the unconscious; which portions of the decision have been made in advance of the actual carrying through of the action, perhaps at the time of the creation of the record, or of the filing of the record, etc..

### 2.11.2 Errors in Comparison

Similarity is not simply the inverse of dissimilarity: they are separate concepts, with some overlap, giving cases in which a subject will attribute both similarity and dissimilarity. Evaluation of similarity / dissimilarity in figures may be regarded as a discrete mental task from the same evaluation of words, as addressed in the section on pictorial vs. conceptual decision-making. Evaluations are also context-dependent, sensitive to the ordering of stimuli, and sensitive to the quality / quantity of features under consideration. “Similarity or dissimilarity data appear in different forms: ratings of pairs, sorting of objects, communality between associations, errors of substitution, and correlation between occurrences. Analyses of these data attempt to explain the observed similarity relations and to capture the underlying structure of the objects under study [Tve04, p. 7].”

To illustrate the directionality of choice, people are comfortable with saying “North Korea is like Red China,” but are uncomfortable in saying “Red China is like North Korea.” Another example of the directionality of comparison: the Son may be like the Father, but not vice versa. Likewise, in testing, judgements of difference yield substantively deviant results as compared to judgements of similarity, despite the two being reciprocal concepts (mathematically). “[T]his asymmetry in the choice of similarity statements is associated with asymmetry in judgments of similarity [...] . Apparently, the direction of asymmetry is determined by the relative salience of the stimuli; the variant is more similar to the prototype than vice versa [Tve04, p. 8].” Also, people utilize different criteria for assessing difference and similarity, with difference focusing upon distinctive features and similarity focusing upon common features.

Generally, we need to consider cases in which the evaluator of digital information is deciding whether a discrete item “is a” something, and also cases in which the evaluator is comparing items to one another within the same class of items. In either of such cases, depending upon what information is presented to them by their computing environment, they will be deciding with greater or lesser quantity and quality of information, will be deciding based upon verbal or pictorial cues, and will be engaged in comparing against items in sets which are changing during the course of their comparison (provided they choose to delete, move, or reorganise any items).

“Two preliminary comments regarding feature representations are in order. First, it is important to note that our total data base concerning a particular object (e.g., a person, a country, or a piece of furniture) is generally rich in content and complex in form. It includes appearance, function, relation to other objects, and any other property of the object that can be deduced from our general knowledge of the world. When faced with a particular task (e.g., identification or similarity assessment) we extract and compile from our data base a limited list of relevant features on the basis of which we perform the required task. Thus, the representation of an object as a collection of features is viewed as a product of a prior process of extraction and compilation [Tve04, p. 10].”

This provides an even more complex field for consideration, as the agent is perhaps not comparing on visible properties, but is making judgements based upon some personal agglomeration of properties which has been subsequently winnowed down based upon criteria unknown to us, all of which takes place prior to the agent’s making any decision about deletion or retention.

“When faced with a set of objects, people often sort them into clusters to reduce information load and facilitate further processing. Clusters are typically selected so as to maximize the similarity of objects within a cluster and the dissimilarity of objects from different clusters. Hence, the addition and/or deletion of objects can alter the clustering of the remaining objects. A change of clusters, in turn, is expected to increase the diagnostic value of features on which the new clusters are based, and therefore, the similarity of objects that share these features [Tve04, p. 29].”

This process—of grouping objects together so as to maximize object similarity within the group and to minimize similarity between groups—is subject to error, in particular when considering that certain objects may be property-enriched while others are property-poor.

“[I]f the common features are weighed more heavily in judgments of similarity than in judgments of difference, then a pair of objects with many common and many distinctive features may be perceived as both more similar and more different than another pair of objects with fewer common and fewer distinctive features [...] . Moreover, on the average, the prominent pairs were selected more frequently than the nonprominent pairs in both the similarity and the difference tasks. For example, 67% of the subjects in the similarity group selected West Germany and East Germany as more similar to each other than Ceylon and Nepal, while 70% of the subjects in the difference group selected West Germany and East Germany as more different from each other than Ceylon and Nepal. These data demonstrate how the relative weight of the common and the distinctive features varies with the task and support the hypothesis that people attend more to the common features in judgments of similarity than in judgments of difference [Tve04, p. 27].”

This study neatly demonstrates not only that similarity and difference are not psychologically complementary, as is the case mathematically, but also demonstrates that the relative availability or accessibility of a given property dictates its use by an agent in making a decision. Thus, we are presented with errors of this type, wherein the accessibility of a property (in this case, any number of properties regarding the similarity and difference between West and East Germany) dictates the weight given to this property, in ranking exercises. This is similar, in some ways, to a conjunction error, in that the sum of the conjuncts is greater than 100%. The conjunction error may be due to overestimating low probability events; this is based in the problem of intellectual or emotional accessibility giving undue weight to a particular fact.

For a user making a decision to delete or retain, she will have to involve herself in a number of grouping and sorting operations, in order to assess similarity of digital objects as well as difference of digital objects. However, similarity and difference are not additive. Similarity and difference are each more likely to be attributed to objects with a greater set of properties under consideration. Similarity is *directional*: e.g., “North Korea is like Red China” [Tve04, p. 81]; the choice in attributions of similarity in this directional manner has an influence upon the judgement of similarity, despite the fact that, in logical terms, it is the same operation:  $((A \cap B) = (B \cap A))$ . All of these features with regards to similarity and difference will contribute to any end decision made, whether to delete or retain. As such, the sorting and grouping process becomes as important in determining the outcome of the operation as the content of the objects themselves. We may see, perhaps, similar effects to that of cases wherein the agent is loss-averse or loss-seeking, depending upon how a particular statement is phrased<sup>18</sup>; if operations are based upon similarity moreso than difference—as in a case where digital objects are diverse and require extensive categorization prior to disposition—we may see our end decision made more as a result of conceptual properties than pictorial properties<sup>19</sup>. Conversely, if operations are based upon difference moreso than similarity—e.g., a case wherein digital objects are uniform, pre-categorized—we may see decisions made more on the basis of difference than upon similarity, on the basis of pictorial rather than conceptual properties.

Tversky introduces the concept of ‘*effective context*’, which is salient to this research. In thinking about how someone would manage, say, documents kept in their *My Documents* folder as compared to their behaviour towards files kept in their *Project Specifications* folder, the simple change in *effective context* implies that the behaviour would be different. Setting aside for a moment the fact that the user might be presented with a subset of files in the one case, simply having the files presented within the more specific container would seem to alter the context significantly enough to influence the decision. Taking the idea of effective context one step further, it seems that unless different media present the same (or equivalent) properties, digital objects will be evaluated differently depending upon the effective context of the media environment. Different applications present different effective contexts, each of which context affects the decisions made. We may wish to consider whether the use of online applications (e.g., Google Docs) acts to minimize changes in effective context, providing the user of these applications with a more homogeneous environment and, thus, a more consistent behaviour towards their documents; we must ask whether this is desirable, in particular because this grants the same effective context to documents managed for personal use and for professional

---

<sup>18</sup>See section 2.12.1, on page 50, for a discussion of framing and its influence on risk aversion and risk seeking.

<sup>19</sup>See section 2.9, on page 33, for a discussion of pictorial selection vs. conceptual selection.

use as a member of some corporate institution. Likewise, we must ask whether it is desirable to have, for example, medical records granted the same effective context as marketing materials for the medical office.

### 2.11.3 More Errors in Comparison

In considering errors in comparison it is important to differentiate between those effects which are produced under any circumstance; and those effects which are truly errors and which are frequently exploited in order to influence agents to make a decision which is not in their best interest, or to decide in a manner which diverges from how they would ideally decide. In the previous section we discussed issues of comparison, context, similarity and difference, and such issues of comparison which are present in nearly all cases of deciding between options. In this section we will consider some of the latter class of comparison error: those which would be regarded as error by the agent engaged in the decision process, if only that agent were aware of what was taking place and how her decision was being manipulated.

#### Valueless Features Influence Decisions

Choice can “be influenced by features that have little or no value [SST04, p. 956].”

“[E]ndowing an option with a feature that was intended to be positive but, in fact, has no value for the decision maker can reduce the tendency to choose that option, even when subjects realize that they are not paying for that added feature [SST04, p. 957].” Thus, if an agent were to be interested in retaining an electronic log of Internet traffic throughout their organization because that log might prove useful to a companion in determining resource allocation, they might decide to retain that traffic log. If, however, they were told that the log could *additionally* tell them how much greater volume of Internet traffic was approaching a holiday period, or during certain times of the day, the decision might be to destroy the log—provided that the agent did not regard such additional “features” of their decision as being positive. This type of decision is *an error in judgement*, because the decision was made not based upon some negative attributes present in the decision, or because of some negative result; nor is it a cognitive error like those in section 2.11.2 (page 36). Certainly, in the interest of privacy they might argue that keeping such logs might be unethical or what have you. No, absent the ethical dimension from consideration: the decision to erase the logs could be made simply because the decision to keep the logs had been endowed with a *non-valued* feature, *simpliciter*. Making such a change in decision based upon perception of a negative feature would certainly not be an error in judgement, but perceiving additional, unimportant features should not alter the decision. Yet

the act of presenting something as having value which does not have value to the agent causes the agent to attribute negative value to an option.

“Providing a context that presents compelling reasons for choosing an option apparently increases people’s tendency to opt for that option, whereas comparing alternatives that render the aforementioned reasons less compelling trends to increase people’s tendency to maintain the status quo or search for other alternatives [SST04, p. 958].” To state this differently: people like to be given reasons to choose, and dislike being given reasons not to choose. When given reasons to choose, they will do so. When given reasons not to choose the other options, people will prefer to maintain the status quo at the very least. At worst, when given negative reasons to choose certain options, people will abandon all choices in favour of seeking a different set of options. Anything presented as a list of simply pro- statements, then, is likely to engender a choice of one of the options. Anything presented as a list of simply con- statements is likely to cause no choice to be made, or all options to be discarded.

### Compromise Effects

Choices are often dictated by *compromise effects*, wherein the range of choices presented prompts the selection of “*the medium-sized cola*,” for example. Compromise effects are “ubiquitous.” Thus, e.g., in legal situations attorneys may or may not present the jury with a “compromise verdict,” depending upon their judgement of the likelihood that the jury will convict or acquit on a more serious charge.

“‘Compromise’ effects are well known to both district attorneys and defenders: as a tactical matter, one side and/or the other might choose not to request that judges instruct jurors to consider lesser included charges, hoping to force the jury to elect between acquittal and conviction of a serious charge, believing that the jury will otherwise be unduly prone to pick the compromise judgement, even if that judgement would attract little support in a two-option set (against acquittal alone, or against conviction of the serious offense alone, assuming one could decide to convict and then, sequentially, grade) [...] . [T]he decision to offer a compromise verdict diminishes the probability of acquittal and is, as a result, a tactical judgement that the client must make in consultation with his attorney [KRT04, p. 976].”

Attorneys routinely take advantage of the fact that juries will fall victim to compromise effects, in order to “win” the best outcome at trial, whether or not their client is guilty. Lawyers are susceptible to compromise effects, themselves, despite being aware of and taking advantage of the effect. Compromise effects are used most often in order to encourage the selection of a

middle-of-the-road choice, with such selection being where the most profit lies in any given set of options.

Compromise effects, though, are also present in many everyday situations, not subject to manipulation by clever lawyers or marketing people; when we choose to select the a extreme option in between keeping everything and utilizing a rigid information-management system (formal software application or simply a manually-implemented policy), we are selecting “the medium-sized cola,” even though rational consideration might otherwise tell us that our best decision might be one of the more extreme options. In this example, in-between solutions might be to allow the operating system to help us to “clean up” digital objects, or to automatically archive emails—both are cases wherein either extreme represents a strong commitment to a particular course of action, leading to selection of a compromise. This type of compromise between extremes is how we select between choices wherein no obvious benefit lies with the extreme choices<sup>20</sup>.

### Contrast Effects

Similar to compromise effects are contrast effects:

“Two types of violations of context-independence—compromise effects and contrast effects—have recently been demonstrated. ‘Compromise effect’ refers to the finding that the same option is evaluated more favorably when it is seen as intermediate in the set of options under consideration when it is extreme [...] . ‘Trade-off contrast,’ or simply ‘contrast,’ refers to the observation that the same option is evaluated more favorably in the presence of similar options clearly inferior to it than in the absence of such options. Contrast effects, more generally, are ubiquitous in perception in judgement [SST04, p. 963].”

A clear example of a contrast effect would be the presence of an absolutely horrible choice amongst other, semi-acceptable choices. For example, an agent might not select a particular vehicle for purchase if that vehicle had a few scratches, a few blemishes, and did not run very smoothly. If, however, the agent had been offered a hideously damaged vehicle immediately prior to evaluating the slightly damaged vehicle, the agent is more likely to view the blemishes as insignificant. The contrast between the two vehicles has changed the agent’s perception and influenced selection.

---

<sup>20</sup>This is related to the “imaginability” of circumstances, as discussed in section 2.11.4 on page 44. It is also closely related to extremeness aversion, as discussed in section 2.7 on page 29

“There are two separate questions policy makers might pose if they believed context-dependent decisions were commonplace. First, they must ascertain whether such decisions are problematic, harming some cognizable interests of some party. Second, they must determine whether the decisions they deem harmful can be avoided by reducing the decision makers’ authority, ‘educating’ them about their tendencies, or designing procedures to reduce context-dependence [KRT04, p. 975].” For our research, we must follow this suggestion: we must actively seek to observe compromise / contrast effects in action within decisions to delete or preserve digital information, and must attempt to classify such effects such that we may report them and seek to have them addressed in the future. As to whether we can capture such effects and, additionally, determine whether they play any role with regards to conceptions of responsibility or duty, that is a difficult question.

## 2.11.4 Errors in Probability Estimation

### Ignoring Prior Probability

People repeatedly, consistently, predictably, and systematically misjudge the probability of certain events. Introduce a description to any test in which the subject knows the prior probability and the subject will allow that description to influence their prediction of the current probability of a similar event, even when the information can have no possible bearing upon the probabilities under consideration. As Tversky and Kahneman say, “people respond differently when given no evidence and when given worthless evidence [TK04b, p. 205].”

One experiment informed a group of subjects that a certain percentage of a population were engineers ( $p = .7$ ) or lawyers ( $p = .3$ ). The subjects were then asked to predict whether a given individual was an engineer or a lawyer. In the absence of any information, predictions were based upon the base probability of ( $p = .7$ ). Given any amount of description about the person, however, the subject discarded prior probability: they “went with their gut.” Further tests followed along this same manner, demonstrating that, when given any descriptive information whatsoever, subjects discard known probabilities. “The subjects correctly utilized prior probabilities when they had no other information [...] . However, prior probabilities were effectively ignored when a description was introduced, even when this description was totally uninformative [TK04b, p. 205].” The uninformative nature of prior probability was established by repeating the experiment with the probabilities reversed, but the descriptions unchanged. The study concludes that people prefer to estimate probability based upon descriptive information rather than upon known mathematical certainties, constructing interpretations of the descriptive text when clearly it was not in their best interest to do so—they ignored prior probability in

order to imagine whether a person fitting the description was likely, in their experience, to be an engineer or a lawyer. Their estimation of probability was based not upon facts, but upon whether the description fit the participant's stereotypical picture of an engineer or lawyer.

In considering deletion / retention decisions which are motivated by risk, certain authors have maintained [see Jon07] that the experts in a given field should be the ones to evaluate certain types of information for its retention or destruction, rather than records managers or librarians. Tversky and Kahneman's research would seem to argue in direct opposition to that assertion: this study determined that even psychologists aware of this effect tended to engage in the same behaviour, and to fall victim to the same temptation to make decisions based upon some descriptive material, rather than based upon the prior probability. The expert in the field is just as susceptible to this affect as everyone else; therefore, someone is needed who can make decisions based upon the science—someone who is aware of this temptation and can hopefully formalise the decision-making process to avoid falling victim to this cognitive error.

### **Imaginability of Circumstances**

“People tend to overestimate the probabilities of representative (or available) events and/or underestimate the probabilities of less representative events. The violation of the conjunction rule demonstrates this tendency even when the ‘true’ probabilities are unknown or unknowable. The basic phenomenon may be considerably more common than the extreme symptom by which it was illustrated [TK04a, p. 250].” If people can bring to mind an example of an event taking place, they overestimate the probability of that event's taking place again. Thus, the probability of an earthquake would be overestimated by someone who had experienced an earthquake recently, despite the fact that the probability remains static over time. This is true in evaluating probabilities in general, but in estimating risk in particular:

“Imaginability plays an important role in the evaluation of probabilities in real-life situations. The risk involved in an adventurous expedition, for example, is evaluated by imagining contingencies with which the expedition is not equipped to cope. If many such difficulties are vividly portrayed, the expedition can be made to appear exceedingly dangerous, although the ease with which disasters are imagined need not reflect their actual likelihood. Conversely, the risk involved in an undertaking may be grossly underestimated if some possible dangers are either difficult to conceive, or simply do not come to mind [TK04b, p. 213].”

Thus, any circumstance which is easy to bring to mind becomes more probable in the mind of the imaginer, and those events which are traumatic or shocking are much more easily brought

to mind. Therefore, not only do people weight the probabilities in favour of extraordinary occurrences, they also overestimate the probability of negative outcomes more than they do positive outcomes.

### **Misunderstanding of “Chance”**

As well as simple errors in estimation based upon imaginability, or the availability of a particular item to memory, people also experience errors in estimation which are more fundamental, particularly with regards to how they understand the workings of chance. “Chance is commonly viewed as a self-correcting process where a deviation in one direction induces a deviation in the opposite direction to restore the equilibrium. In fact, deviations are not ‘corrected’ as a chance process unfolds, they are merely diluted [TK04b, p. 207].” The thought process in question is, essentially, that if tossing a coin yields a probability of ( $p = 0.5$ ) then a toss of “heads” ought to, in the subsequent toss, yield “tails.” It is not so concrete as this example, however, because people will readily admit that the probability remains ( $p = 0.5$ ) for each and every toss of the coin, yet feel that chance out to “even out” after some period of time; this feeling that the “evening out” ought to produce a certain value is where the error enters the thought process: the sentiment is that the *probability* ought to change, to produce appropriately balanced results.

### **Anchoring**

“In many situations, people make estimates by starting from an initial value which is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formulation of the problem, or else it may be the result of a partial computation. Whatever the source of the initial value, adjustments are typically insufficient. That is, different starting [points] yield different estimates, which are biased towards the initial values. We call this phenomenon anchoring [TK04b, p. 214].”

In Tversky’s study of anchoring, he found that anchoring takes place in a large variety of circumstances, but is especially prevalent when presented with a conjunctive situation. The subject will tend to form an estimation of the probability of the first element in the conjunct and will then *modify* that estimate of probability to incorporate the second element of the conjunct. This is not conjunction error, exactly, but is related: people use one fact as the an anchor, the second as a modifier, similar to the manner in which people use one element of a conjunct as an explanatory fact. In this case, however, one element dominates the estimation process so as to leave the estimate fixed about a certain, initial value. This is different than

conjunction errors, where both probability estimates change, based upon the explanatory value of the other element in the conjunct.

### Conjunction Errors

“Studies of choice among gambles and of judgments of probability indicate that people tend to overestimate the probability of conjunctive events and to underestimate the probability of disjunctive events. These biases are readily explained as effects of anchoring. The stated probability of the elementary event (e.g., of success at any one stage) provides a natural starting point for the estimation of the probabilities of both conjunctive and disjunctive events. Since adjustment from the starting point is typically insufficient, the final estimates remain too close to the probabilities of the elementary events in both cases. Note that the overall probability of a conjunctive event is lower than the probability of each elementary event, whereas the overall probability of a disjunctive event is higher than the probability of each elementary event. As a consequence of anchoring, the overall probability will be overestimated in conjunctive problems and underestimated in disjunctive problems [TK04b, p. 215].”

This article provides the classic “Linda is a bank teller and a feminist” example, wherein the probabilities of the conjunction of the two facts is judged to be more probable than either fact by itself, despite the fact that there must certainly be instances of female bank tellers who are not feminists, and that there are female feminists who are not bank tellers. Logically, the probability of someone being both a bank teller *and* a feminist is necessarily less than or equal to the lower of the two probabilities:  $[pB \cap F \leq (pB \vee pF)]$ . The mathematical operation  $(pB \cap F)$  is the *intersection* operator; in evaluating conjunctions, however, people evidently feel the need to express  $(pB \cup F)$  (to use a *union* operator) or to express  $(pB \cup pF)$  (to use a mathematical OR operator). In everyday language, ‘or’ may mean {union / mathematical OR / intersection}.

This is apparently not quite so common of an error when subjects think of elements of the conjunction as properties rather than classes, despite the fact that properties and classes are logically equivalent. Tversky and Kahneman suggest that this phenomena may be avoided, then, by structuring elements of conjunctions as objects with properties rather than as classes of objects. Again, the Anchoring effect is seen. “Conjunctions involving hypothetical causes are particularly prone to error because it is more natural to assess the probability of the effect given the cause than the joint probability of the effect and the cause [...] . [W]e propose that the higher conditional estimate serves as an anchor that makes the conjunction appear more probable [TK04a, pp. 244-245].”

### Errors in Probability Estimation: conclusion

People's estimates of probability tend to be biased based upon the phenomena of anchoring, the gambler's fallacy, and the imaginability of the negative consequence under consideration. People commonly overestimate the probability of a conjunctive event and underestimate the probability of a disjunctive event. Estimation of probability is formed by some unconscious guess or rough calculation which then dominates further estimates of probability. People feel that things should "balance out" to meet their estimate, rather than revising their estimation of probability.

These types of error should be found in decisions to delete or retain digital objects in particular when considering whether objects will be needed in the future or will be detrimental if not destroyed. The availability of cases in which destruction took place and which were followed by the desire not to have destroyed the object will be one type of an event which is easily available to the imagination, and which may dominate behaviour with respect to destruction decisions: subjects may have only ever needed a vanishingly small percentage of the digital objects which have been destroyed, yet that small percentage serves as an anchor for them in their decisions. Likewise, cases in which the subject did *not* destroy a digital object which decision subsequently caused them some trouble or distress may also serve as an anchor.

Conjunction errors are likely to be found more when considering risk than in considering benefit, although it is certainly possible to imagine cases on either side. The formulation of conjunctions, however, would seem to be an operation likely to take place in a more formal setting, or in a group setting, e.g., a risk assessment meeting. When considering decisions to delete or retain, under those conditions, risks are expressed along with some estimate of their probability, and their probabilities subsequently summed in some manner as to produce a risk score, indicating which decision ought to be taken. A formal risk assessment, however, is performed under conditions which we would hope might mitigate against such affects. Each individual element of the extended conjunction which comprises a risk assessment, though, is likely to be subject to the anchoring effect.

## 2.12 Consistent Errors in Calculating Probability

Fox and Tversky [FT04a] examined the subjective judgement of probability. Their study concludes that subjects attribute greater initial probability to an event when it has reached some threshold of possibility, overvaluing the probability as probability approaches ( $p = 0.4$ ), and that subjects attribute lesser probability to an event as it approaches the threshold of certainty, underestimating the probability as probability approaches ( $p = 1.0$ ). This yields in

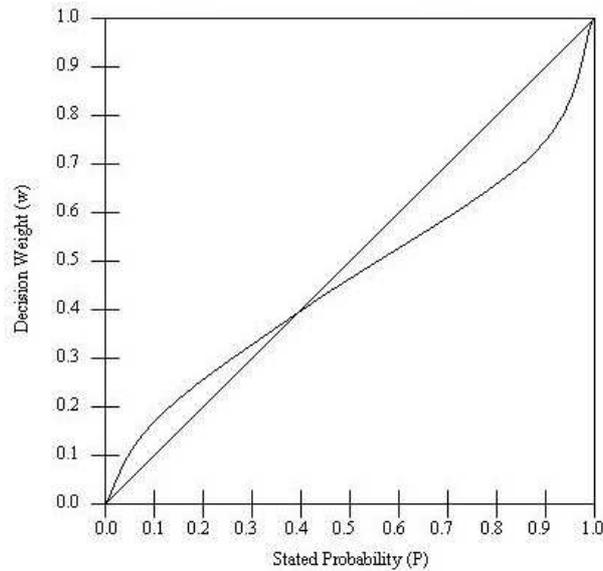


Figure 2.1: Stated Probability vs. Willingness to Gamble [FT04b, p. 749]

an S-shaped probability estimation, rather than a linear estimation. What this means, initially, is that people estimate probabilities poorly—but very consistently, and very predictably.

The initial finding, in which subjects overestimate events with ( $p < 0.4$ ) and underestimate events with ( $p > 0.4$ ), *should* mean that subjects will err slightly towards the destruction of digital objects with low probability of being harmful to the organisation, on the one hand, and should err towards the destruction of digital objects which have high probability of being useful to the organisation, on the other: assume that the decision to delete is made on the basis of how damaging a digital object might be ( $D$ ) vs. how useful that same object might be ( $U$ ). If ( $pD < 0.4$ ), then ( $pD$ ) is being overestimated. Likewise, if ( $pU > 0.4$ ) then ( $pU$ ) is being underestimated. This gives us the case wherein digital objects are selected for destruction improperly, and wherein fewer digital objects are selected for retention properly.

In a twist on the idea that disjunctions are perceived to be less probable than each of the members of the disjunct, Fox and Tversky briefly touch on the fact that the description of an event, if unpacked “into an explicit disjunction of constituent events,” is judged to be of greater probability [FT04a, p. 806]. Thus, the probability of someone dying within a certain time period may be estimated to be ( $p = 0.2$ ). Yet, the same person may estimate that the probability of that person dying of cancer is ( $p = 0.05$ ), the probability of them dying from a fall is ( $p = 0.05$ ), them dying from a heart attack is ( $p = 0.05$ ), them dying from drowning is ( $p = 0.05$ ), them dying from a stroke is ( $p = 0.05$ ), etc. All of these individually low estimations of individual elements in the disjunct sum to greater than the estimation ‘that they will die, full-stop’.

The unpacking principle is interesting as well, of course, because it fits in with some of the

other, similar effects such as the imaginability of an event increasing its perceived probability<sup>21</sup>. This area applies to single-factor risk assessment as opposed to multiple-factor risks (the risk of, e.g., the primary software developer dying vs. the risk of catastrophic flooding leading to generator failure at the off-site data storage facility causing destruction of data). The single-factor risks, because of imaginability and the unpacking principle, are undervalued, while the multiple-factor risks are overvalued. Stated differently, the aggregate risk (death from heart attack, death from stroke, hit by car, plague, etc.) exceeds the simple risk (death), which should not be the case (because, certainly, one cannot enumerate and quantify the risk of death from every possible cause).

These phenomena persist in circumstances wherein we might expect the agent to be very careful in making a judgement “Johnson et al. (1993), for example, reported that subjects were willing to pay more for a health insurance policy that covers hospitalization for all diseases and accidents than for a policy that covers hospitalization for any reason [FT04a, p. 807].” The subjects were obviously aware of the distinction between the two health insurance policies, and could consider the elements of the unpacked, explicit disjunction. They also had opportunity, one supposes, to consider that there might be circumstances under which they would not be covered for hospitalization; intentional acts of self-harm would seem to be excluded from coverage, for instance, as would injuries caused by brawling in the pub (which may be taken to be intentional self-injury, and thus not accidental). Yet they were willing to pay more for the more descriptive (and more limited) of the two policies.

Probability calculations, as calculation, come into play when considering the possible future role of individual digital objects, or of classes of digital objects. Calculation error is subject to some of the same errors as estimation error, because it is based in some of the same cognitive deficiencies. Calculation error, though, affects decisions to delete in a more concrete manner: where estimates may affect individual decisions, the *class* of decision would seem to be different where estimates of probability are involved. Calculation of probability takes place in formal decision-making processes, wherein people are aware that they are making a decision which will have a significant impact upon the future; they are attempting to make a judgement which may incorporate aspects of normativity, legality, pragmatics.

Where both estimation and calculation errors affect digital objects, they affect them differently than physical objects primarily because the balance of pressures is different: storage is of concern with physical objects, as is maintenance and physical preservation. These aspects have all been downplayed in people’s consideration of digital objects, so other aspects of the decision become more important. It would seem that the decision ought to, in this way, result

---

<sup>21</sup>See section 2.11.4, page 44.

in a more human-centred decision. We will see if that is the case during our research, hopefully.

### 2.12.1 Errors in Calculating Risk

#### Prospect Theory

More specifically than simple considerations of probability, evaluation of risk is subject to its own cognitive errors.

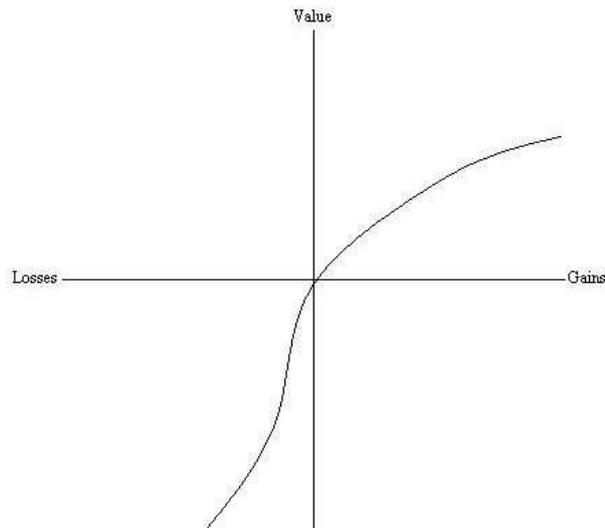


Figure 2.2: Proposition Stated as Gain / Loss vs. Willingness to Gamble [TK04c, p. 601]

Depending upon how something is phrased (as a possible loss, or as a possible gain), the person being asked to gamble will make a different decision. Loss aversion is a stronger motivator than possible gain: the displeasure of losing a sum of money is greater than the potential pleasure of winning the same amount of money, so if the bet is framed as one with potential gain it is more likely to be taken. The figure shown here demonstrates willingness to bet, based solely upon the manner in which a proposition was framed: in terms of possible gain or possible loss along the Y axis, and in terms of willingness to gamble, along the X axis. The situations stated were merely phrased so as to be difficult to reframe from the domain of possible loss into that of possible gain. Subjects were much more willing to bet upon the possibility of a possible gain (willing to bet 2 against someone else's 1) than they were willing to bet when the problem was phrased as possible loss (they required a win of 3 to their bet of 1).

“Prospect theory distinguishes two phases in the choice process: a phase of framing and editing, followed by a phase of evaluation (Kahneman and Tversky 1979). The first phase consists of a preliminary analysis of the decision problem, which frames the effective acts, contingencies, and outcomes. Framing is controlled by

the manner in which the choice problem is presented as well as by norms, habits, and expectancies of the decision maker. Additional operations that are performed prior to evaluation include cancellation of common components and the elimination of options that are seen to be dominated by others. In the second phase, the framed prospects are evaluated, and the prospect of highest value is selected [TK04c, pp. 598-599].”

The actual decision, then, is several cognitive steps from the posed problem: the agent must first come to some understanding of the situation, in terms of potential loss or potential gain; must then evaluate the situation in search of possible “acts, contingencies, and outcomes;” and only then can the agent begin to formulate a decision consciously. However, the manner in which the situation was framed and edited has a very significant impact upon the outcome. Framing the issue is tremendously significant in terms of the end decision made, and is something over which the agent has no control.

Risk aversion is different than probability estimation errors because risk cannot be framed in the same manner as gambling: probability estimation errors concern a single factor (e.g., whether a digital object will become corrupt) whereas risk aversion effects are found when there is the possibility of both risk *and* loss, and both risk and loss are of the same content albeit of different magnitudes. Probability estimation errors are different than prospect errors because prospect errors are affected by the framing of the prospects whereas probability estimation errors are affected by the imaginability of a particular event—prospect errors may be presented as possible loss vs. possible gain whereas probability estimates are unidimensional in this manner and cannot be reframed in the same way.

Despite that probability estimates are not subject to framing in the same way as gambles, it is difficult to untangle the two phenomena: if the subject perceives a particular decision as a gamble (e.g., the possibility that something might cause legal problems if not destroyed vs. the possibility that something might be needed later), then that subject may go through a mental process similar to that as described by prospect theory. If, however, only a single dimension is considered as would be the case when estimating a probability (e.g., only considering whether something will cause legal problems if not destroyed), then the subject may not go through that mental process and may only be subject to errors in estimating probability. Whether this is significant has not been determined; however, the decision-making process in each case is different enough to affect decisions to delete or preserve. The question for this study, then, is whether participants perceive deletion and preservation decisions to be probability estimations or gambles; the decision-making process—and the cognitive errors—will be different depending

upon the subjects' perception of the situation.

### Competence

People's perception of their own competence (or lack thereof) has a significant impact upon whether they are willing to gamble on a given outcome versus simply "playing the odds" of the situation. If people believe themselves to be knowledgeable in a given area, they are more likely to gamble—even though their probabilities of gambling correctly are lower than if they were to gamble based on the odds alone. "Because the chance level [in the gamble] was known to the subject whereas the skill level was not, ambiguity aversion implies that subjects would shift as much uncertainty as possible to the chance component of the gamble. In contrast, 87% of the choices reflect a preference for skill over chance [HT04, p. 647]."

A decision to retain / delete may be perceived as a gamble: a gamble that the object will not be required at some future date on the one side and that the object will not prove harmful if kept. If gambling that the object will be required in the future, and the gamble is made by the person in whose area of competence the record was generated, it would seem that the gambler would over-retain; if gambling that the record will prove dangerous (and legal issues are outside of the gambler's area of competence), it would seem, also, that the tendency would be to over-retain (because the gambler would be more likely to "play the odds" than otherwise). The praise allocated for retention would tend to argue in favour of retention, while the relatively small quantity of blame would also tend to argue for retention or, at least, to some formal evaluation of the odds. (Interestingly, this is not the case where the professional records manager is the one gambling, as the professional records manager will receive blame in *either* case.)

"[T]he balance of credit to blame is most favorable for bets in one's area of expertise, intermediate for chance events, and least favorable for bets in an area where one has only limited knowledge. This account provides an explanation of the competence hypothesis in terms of the asymmetry of credit and blame induced by knowledge or competence [HT04, p. 648]." This hypothesis is stronger, though, than just that people will receive credit or blame, so they attempt to use their knowledge in order to determine the correct bet. Heath and Tversky found that "people prefer to bet on their high-knowledge predictions even when the predictions are unlikely to be correct [HT04, p. 654]." Further, when calculating the difference between playing the odds and playing based upon their own knowledge, Heath and Tversky determined that "people are paying a premium of nearly 20% for betting on high-knowledge items [HT04, p. 654]." Finally, Heath and Tversky found that "many decision makers do not regard a calculated risk in their area of competence as a gamble [HT04, p. 670]."

If we consider, then, the difference between an individual making a decision to delete or retain a digital object, and that same decision as undertaken by a records manager, we must conclude that the records manager would make a better decision simply because the records manager is not subject to blame in the same way as the individual within the organisation: the records manager is not expected to have knowledge of the content of the record in the same way as the individual expert. Thus, where an individual would be subject to praise or blame for not recognizing the importance of a record, and would be inclined to bet with their “knowledge” (and pay a 20% premium for doing so), the records manager would be expected to have made the decision in line with other properties of the object, and would be subject to praise or blame *not* based upon the content of the record, *per se*.

If we consider individuals making decisions over the domain of digital objects present within their computer systems, we must take into account that they are making decisions over areas within their domain of expertise as well as over other domains. Those which are subject to praise / blame are those objects which will be subject to more deletion / retention decisions, while other objects will be less so, and may be deleted / retained in a more systematic manner as a result. It would seem that the objects of most import to any individual will also be those most likely to be improperly destroyed *and* improperly retained, by a margin of (let  $f$  = frequency of expert’s error;  $f \rightarrow 20\%$ ). The likelihood of the expert making a poor decision is, (let  $pN$  = the probability of an incorrect decision by a non-expert simply playing the odds;  $pE = pN + 0.2$ ).

### 2.12.2 Ambiguity and Comparison of Risks

People are averse to ambiguity in a situation only when the ambiguous situation is compared with a clear situation. If they are presented with a one-off gamble, they’re happy to gamble, whereas if they’re given a choice between gambles where one is relatively clear or in which the variables are familiar and one in which the variables are relatively vague, people become sensitive to their own lack of knowledge and will either prefer the clear gamble or will prefer not to gamble at all. This is the “comparative ignorance hypothesis.” Comparative ignorance may be induced through a variety of means, even in situations which would ordinarily be regarded as clear, e.g. via the introduction of an “expert” who would also be gambling alongside of the subject.

“The main implication of this account, called the comparative ignorance hypothesis, is that ambiguity aversion will be present when subjects evaluate clear and vague prospects jointly, but it will greatly diminish or disappear when they evaluate each prospect in isolation [FT04b, pp. 778-779].” Comparative ignorance is similar to contrast effects; where contrast effects apply

to decisions about the similarity or difference between things and affect decisions because the agent perceives one situation to be preferable *in contrast to* another situation, comparative ignorance involves the subject evaluating her own knowledge as compared to that of an hypothetical other, or as compared to some idealized conception of what *ought* to be known. Comparative ignorance, then, is about epistemological doubt, where contrast effects are about subjective preference wherein the subject doubts not her own knowledge but her own desire.

Comparative ignorance tends to cause subjects to make *no* decision, rather than to cause them to make a *wrong* decision. In this way, comparative ignorance differs from areas in which the agent is deciding based upon knowledge within their area of expertise. Comparative ignorance tends to exert pressure towards suspension of decision; perceptions of competence exert pressure towards deciding based upon knowledge rather than statistical certainties. Both result in an agent making what is, statistically, a poor decision. Comparative ignorance brings about doubts about knowledge, result in the agent deferring to another or making no decision; feeling expertise in a particular area will cause the agent to feel pressured to bet based upon their own knowledge, rather than taking advantage of any known mathematical probability.

## 2.13 Psychological Issues: Confusion in Concepts

### 2.13.1 Confusion between Causal and Indicative

People choose actions which are associated with their desired outcome, despite knowing that the action is in no way causal of the outcome. People attempt to *magic* the outcome by imitating the conditions of the outcome: “people [...] select an action correlated with an auspicious outcome even if they believed that the action is only diagnostic of the outcome and in no way causal. Thus even if students were presented with compelling evidence that review sessions have no causal influence on their examination performance, and they accept the evidence, they might nonetheless be tempted to attend, so long as better grades are associated with attendance than with non-attendance [QT04, p. 827].” Similarly, people choose actions which support their existing beliefs, and “may in part be motivated by the individual’s attempt to convince himself that the belief is valid [QT04, p. 829].”

This effect is somewhat related to the comparative ignorance hypothesis, in that it relates to epistemic doubt. It is not, however, about comparing one’s own knowledge to that of others (real or imagined), but about recognizing a lack of knowledge and choosing to imitate those who are believed to have knowledge. This behaviour is somewhat like the child pretending to shave: they recognize that the behaviour is important, and may even know *why* it is important, but

they do not recognize that its importance does not derive from the performance of the action, but from the actual end result. Thus, the child's shaving may remove the shaving foam, but does not accomplish the same end result as that of the parent: removal of whiskers. This is, then, a cognitive error in that the importance has been attached to the behaviour in the wrong way.

We might see this at play in decisions to retain or delete where agents believe that retaining copies of digital objects is desirable because it demonstrates good practice within the field (e.g., "good businesses keep backup copies of things"). We might also see this phenomenon played out on the computer by people who compulsively file and organise their digital objects, telling themselves by this action that they are organised and in control of their digital content. The desire *to be organised* may motivate the subject to go through what they perceive as organising behaviour, but this behaviour may not be achieving the end goal of *being organised*. In confusing causal with indicative, the agent may purge huge quantities of digital objects in order to achieve what resembles neatness or organisation, but which is just a more manageable number of digital objects *to them*, not having modified their system in any significant manner whatsoever on a taxonomical or organisational level.

Shafir and Tversky [ST04] expand this confusal of causative and indicative to include something they call *Quasi-magical Thinking*:

"Magical thinking refers to the erroneous belief that one can influence an outcome (e.g., the role of a die) by some symbolic or other indirect act (e.g., imagining a particular number) even though the act has no causal link to the outcome. We introduce the term quasi-magical thinking to describe cases in which people act as if they erroneously believe that their action influences the outcome, even though they do not really hold that belief. As in the Prisoner's Dilemma, the pattern of preferences observed in Newcomb's problem, may be described as quasi-magical thinking. When the program's prediction is known, the outcome depends entirely on the subject's decision and the obvious choice is to take both boxes. But as long as the program's prediction is not known and the eventual outcome depends on the behavior of both subjects and the program, there is a temptation to act as if one's decision could affect the program's prediction [ST04, p. 716]."

This effect is not limited to those who are simply novices in a given field, imitating the behaviours of experts. This effect includes those who are experts in a given field, are aware that they cannot influence the outcome by altering their behaviour, yet continue to behave as if it were possible to do so. This is particularly true in cases wherein the behaviour of others is at issue—when to

compromise would provide both parties with some degree of happiness, rather than the winner taking all.

“[Q]uasi-magical thinking appears even more puzzling because it undermines the link between belief and action. Whereas magical thinking involves indefensible beliefs, quasi-magical thinking yields inexplicable actions. The presence of uncertainty, we suggest, is a major contributor to quasi-magical thinking; few people act as if they can undo an already certain event by performing an action that is diagnostic of an alternative event [ST04, p. 717].” The behaviour of the test subjects was such that, if they believed the other subject to have already made their selection, they ceased behaving as if they could somehow affect the outcome. Once uncertainty *that* a decision had been made was removed, people were comfortable making the best decision for their own benefit. So long as there was doubt *that* a decision had been made, people were confused, and played to their own disadvantage: they played as if everyone would cooperate so long as *they* cooperated. Once a decision had been made, though, they were free from the obligation to cooperate.

### 2.13.2 Irrelevant Information Affects Decisions

Shafir and Tversky explore the disjunction effect and how it plays out in the Prisoner’s Dilemma and also in other (more common) situations. This is similar to “quasi-magical thinking [ST04, p. 716],” in that the subject suspends their decision-making process because of some unknown factor; it is not quite the same, however, in that it becomes clear that the disjunction effect is not related to normativity in any way: the disjunction effect is not about praiseworthy or blameworthy action, but is simply a brute fact about the way people behave.

The introduction of an irrelevant factor can be made to bring about the disjunction effect and to alter the decision of the subject. For instance, a student is presented with the opportunity to buy a trip to Hawaii. The student wants to go to Hawaii. Therefore, the student would elect to purchase the trip. Introduce the fact that the student does not know her exam outcomes yet, and the student will elect to delay the choice to buy the trip, even though she would desire to go under both possible outcomes (she passes her exams, or she fails her exams), although for different reasons in each case. The disjunction effect is enough to put off the decision. “A disjunction effect occurs when people prefer  $x$  over  $y$  when they know that event  $A$  obtains, and they also prefer  $x$  over  $y$  when they know that event  $A$  does not obtain, but they prefer  $y$  over  $x$  when it is unknown whether or not  $A$  obtains. The disjunction effect amounts to a violation of [...] consequentialism [ST04, p. 705].” This is of particular concern because it would seem to indicate that people are unable not only to think logically about a decision, but are able

to be dissuaded from arriving at a conclusion merely by including an unresolvable proposition in their decision process. If this is the case, and people are unable to resolve logical questions when there is an unknown, it would seem that far more digital objects would be retained on the basis of the subject simply not being able to decide, based upon unknown future needs if nothing else.

More worrisome is the fact that “people are assumed to focus on items that have been explicitly mentioned, to apply pre-stored knowledge structures, or to remember relevant past experiences [. . .]. [P]eople find it relatively easy to reason logically about each isolated outcome, but a disjunction of outcomes leads them to suspend logical reasoning [ST04, p. 719].” This is more worrisome, because it is not just that people suspend judgement: they suspend logical thought and *may still make some judgement*.

“It is apparently difficult to devote full attention to each of several branches of an event tree. As a result, people may be reluctant to entertain the various hypothetical branches. Furthermore, they may lack the motivation to traverse the tree simply because they presume, as is often the case, that the problem will not be resolved by separately evaluating the branches. We usually tend to formulate problems in ways that have sifted through the irrelevant disjunctions: those that are left are normally assumed to involve genuine conflict [ST04, p. 720].”

Formal methodologies for decision-making, which set out risks and benefits, and attempt to assign some numerical quantifier to each, are an attempt to circumvent just this type of paralysis of logic.

## 2.14 Psychological Issues: Normativity of Errors in Judgement

Tversky and Kahneman [TK04a] have demonstrated, time and again, the inconsistencies of decision under adversity. Fox [Fox92] takes issue with their conclusion, picking at the premiss that logical thought is desirable over intuition. He proposes that intuition *just is* prejudice, which *just is* prejudgement, and that prejudgement is a logical, evolutionary advance over beings which would have to consider every situation on its own unique merits. He suggests that people, to be sane, need to ignore the pressure to be overly logical. He does not quite present an argument, really, against the facts as presented by Tversky and Kahneman . . . but an argument against the idea that we need to be logical in the way Tversky and Kahneman have proven we are not.

Fox [Fox92] carries this idea through an involved discussion of Hume, Kant, and Quine (who apparently held similar views as to the evolutionary root of such seemingly-illogical beliefs, quoted by Fox as stating that “Veridical expectation has survival value in the wild. Innate standards of subjective similarity that promote successful expectation will tend to be handed down through the survival of the fittest [Fox92, p. 143].”). This line of philosophical argument seems to be concerned with causation more than anything. He moves on to a discussion of moral philosophy, bringing in Peter Strawson, with questions as to whether it is blameworthy for someone to (mis)judge in the manner Tversky and Kahneman have demonstrated. Fox’s argument using Strawson is one concerning whether it is appropriate to ascribe blame if the subject does not have free will in the matter. This is the “meat” of his argument: because people have a moral sense, and this moral sense is internal to that person (innate), the question of ascription of blame must also be based in an internal moral sense (i.e., intuition). Thus the psychologists who are examining behaviour based upon *external* criteria are judging people by a standard which cannot apply: that of something outside morality, something which cannot play a part in the ascription of blame. We, as humans (maintains Strawson) cannot maintain objectivity for long, and certainly not indefinitely.

Fox describes the manner in which stereotyping plays a role in everything from our family relationships down to our understanding of linguistic structures, and maintains that the human brain *prototypes* everything, and measures against those prototypes. He concludes that thinking *just is* stereotyping, and that to ignore the stereotypes is not something attainable nor desirable.

It is clear that people behave in predictable, deterministic ways, and do so based upon prejudices, learned behaviour, and upon consistent logical mistakes. Sugden [Sug98] argues for this not being so bad, as “salience ultimately comes down to a non-rational (but *not irrational*) propensity to choose in certain ways *when reason gives out* [...] . Within a community [...] there is often something approaching common knowledge that certain properties are salient for the members of that community [Sug98, p. 385].” Somehow, according to Sugden, people are able to arrive at decisions which are good *despite* the fact that logic flew out the window at some point: people *just do decide*, and decide according to whatever is ‘good’ within that community.

## 2.15 Not “Records” Behaviour, “Social” Behaviour

Another facet of the problem is how people perceive their actions; if they perceive decisions to delete or retain as decisions about records, our case is a bit clearer than if people perceive their decisions as Stone [Sto99] would have us believe: as decisions about community-building activities in which digital objects are important to the community only so long as they are

actively part of discussion or inform group activities.

Stone considers information-centred behaviour from a social perspective, rather than the perspective of conflict or competition (as contrast Vaughan [Vau98] and Rapoport [Rap63]). She points to the construction of communities as a primary motivating factor in people's interactions (specifically, on the world wide web). Thus, if we consider information management practices on the web, we can frame them in sociological terms, rather than information management terms: these types of record are artefacts with a primarily social existence. Things such as blog posts and social networking statuses are viewed as socialization, rather than records-generating.

Regarded this way, it makes sense that there would seem to be no need for their management after a short period of time, as the digital objects have served their social purpose and are no longer important to the community; they have lost their social context, and that is what gave them their value to the creators and those exposed to them. It also makes sense from this perspective that, after a certain period of time, users of these types of service may feel free to erase old records, regardless of whether others have contributed to them, as in comments which are erased with the original posting<sup>22</sup>. The entire "conversation" has lost its social context, so the user perceives it as having little value, and thus perceives little duty to preserve it. For those who may have linked to it from outside the site, or participated in the commentary linked to the posting, little consideration is given because the information has served its purpose in the construction of the community. Such artefacts, constructed on the web and for a particular community, have a context which may stretch across national borders and which may persist for a period of days or weeks, but are something which are, by their very nature, only persistent in the manner of a play bill: after a few weeks, they may be kept around for sentimental reasons, but the performance is over and will be forgotten in time except to a few who may have a persistent interest in the content.

If we are to extend Stone and the sociology of digital objects to include a full range of digital objects, we need to consider that the objects she studied were fairly narrow in their scope: blogs and social networking statuses are highly restricted both in their form and in the manner in which they are shared throughout the community. This same does not hold true for emails and documents. Nor is the level of mediation of such things as database records standardised to the same degree simply because there may be multiple interfaces to a database, each revealing different aspects of any number of different data entities. The objects considered by Stone are both more highly-standardised and simpler than many of the digital objects to be considered by this study, and would seem also to occupy an unique position *as* social objects: their *primary purpose* could be said to be the participation in a social dynamic, even should they convey

---

<sup>22</sup>Although, see section 2.1 on page 9 for a discussion of how age and youth bias our perceptions.

information about a particular person or topic. The same is not true for, e.g., email sent to a co-worker regarding a particular project.

It is perhaps a bit of a stretch to extend the “social behaviour” of Stone into other technologies—but only a small stretch, considering the ephemeral and short-lived nature of some forms of communication (e.g., text messages and emails). Stone considered online interactions, which are conducted at least somewhat in the public forum. But when we consider that emails are frequently sent to vast distribution lists then we are able to see where, perhaps, email communications could be perceived as serving a similar, social purpose, rather than being perceived as having more of a functional purpose. Similarly, the transitory nature of text messages may be perceived as solely serving a social purpose (no matter that they also convey information).

The difficulty in incorporating Stone into this research, however, is that there is no hard and fast means of determining which purpose a particular form of digital object serves *primarily*; indeed, it is likely to be true that individuals attribute different primary purposes to different digital objects depending upon the content of the digital object or depending upon the context in which the digital object is considered. This presents perhaps one of the larger challenges when considering digital objects: not only do the objects lack any physical indications as to whether they are to be considered a particular type of object, but the boundaries between particular forms of digital objects are porous and variable.

Stone [Sto99] is concerned with the interaction between actors as mediated by certain types of digital object. This interaction must certainly inform decisions to delete or preserve digital objects, but is problematic when considering a full range of digital objects because we lack access to the complexity of the decision-making process involved in such mediated interaction. Social considerations are certainly an important aspect to decisions to delete or preserve; this study, however, should consider other aspects of such decisions—whether actors consider others in their decisions is merely one part of the decision process, and does not tell the entire story by any means. While Stone is concerned with social interaction, this study is concerned with understanding the act of deletion or preservation rather than understanding the way in which relationships are mediated by digital objects.

The following chapter addresses the overall research methodology and how the human study was designed. It includes a discussion of the aims of the study, how such study might alternatively been constructed, and goes on to analyse deficiencies in the study as conducted.

## Chapter 3

---

# Research Methodology and Design

---

### Introduction

This chapter details the research methodology and method used to conduct an investigation into the reasons people give for behaviour with regards to the disposition of digital objects. This chapter also narrows the focus somewhat from the breadth of possibilities examined in chapter 2 (page 9), providing a set of study aims which were felt to be cohesive and could be answered using the study method selected. It also details the manner in which the study instrument was designed and tested, and potential weaknesses of that study method.

This research is concerned with factors influencing decisions to delete or preserve digital objects. Beliefs about the future, about duty and obligation, ethics and morality, are all central to these decisions and necessarily influence the decisions made. The literature within the field of records management has very little to say concerning these beliefs, instead focusing upon issues surrounding use, misuse, or adoption of records management systems; where ethical considerations enter into the discussion, the literature tends to focus upon well-defined objects which are clearly records or record-worthy, rather than upon the actors' motivations where the record-worthiness of a digital object may be less than clear<sup>1</sup>. Nor do neighbouring fields such as business management, knowledge management, or information science consider what drives such decisions, choosing to focus upon the results of poor decisions—or focus upon enforcement of regulations and behaviour—rather than addressing the factors behind those decisions. Very little existing literature examines beliefs and motivations of agents within the work environment acting upon digital objects<sup>2</sup>, nor is there much literature which considers behaviour variances

---

<sup>1</sup>See section 2.6 on page 19 for a discussion upon ethical standards within organisations, social control, dissemination of ethical standards, etc.; and see section 2.15 on page 58 for a discussion upon behaviour as regarded as a social act rather than a records-management act.

<sup>2</sup>Although, see Binark and Sütçü [BS09] for an investigation of normativity within the context of a massively-

between home and work contexts<sup>3</sup>.

### 3.1 Research Questions Arising from Literature Review

This section provides a listing of questions which arose from various sections of the literature review and were formulated during the investigation into each topic area. These questions were consolidated into a single list and formed the basis both for further areas of study within the literature review (section 2 on page 9) and also for the construction of the survey instrument (see section 3.6.1 on page 81 for a discussion of the survey composition process, and appendix A on page 156 for the survey instrument itself). Included in this section are those research questions which *a*) could possibly be answered by a human study, *b*) were somewhat related to the study area as opposed to those which looked to deviate into entirely new branches of study (although these were retained for possible later investigation), and *c*) were somewhat thematically related to each other. While not all individual topic areas were investigated by the human study—e.g., the effects of verbal vs. pictorial representation and its effect upon the selection of objects (section 2.9 on page 33)—those areas *did* inform the survey composition process.

#### Research Questions for Section 2 on page 9

- Are age-based prejudices strong enough to outweigh prickings of conscience, or considerations of obligation?
- Is a digital record automatically devalued because it is old, to the point where considerations of duty are pushed aside?
- If society as a whole ascribes value in the same manner, then is it a bad thing to devalue digital records because of their age? This larger question is perhaps best left to someone else's consideration, in other research, but is important to point out here because we are not making an assertion in either direction in this paper.

#### Research Questions for Section 2.2 on page 10

- Do people incorporate new digital objects into their lives in a similar manner as they incorporate new technology items? I.e., as people obtain new hardware gadgets because of their “coolness” or simply *because* they are “the new thing,” do they follow the same behaviour with digital objects?

---

multiple, online, role-playing game.

<sup>3</sup>Although, see Dervin and Riikonen [DR09] for an argument against viewing context as significant with regards to digital technologies.

- Do people destroy digital objects because they are “old,” either in terms of having been created at some point in the past, or in terms of having been created by software which is perceived as “old” or “outdated?” In particular, do people privilege the new over the old in decisions to delete?
- Is the age of a digital object inappropriately incorporated into retention / destruction decisions?
- Do people correlate the age of a digital object with its usefulness?
- How do people arrive at the “age” of a digital object? Which “age” is used: e.g., “file creation date,” “last modified date,” “last accessed date?”
- Within classes of digital objects which do not provide any indication of “age,” do users feel that something is lacking? I.e., How fixated are people upon “age” as a decision criterion?

#### **Research Questions for Section 2.5 on page 17**

- Do people wish they were more machinelike in dealing with technology systems *in general*, or are these desires limited to certain, very specific arenas?
- Do desires to be machinelike carry over into use of other technologies, or are these desires emotionally confined to their “appropriate” place?
- Do people who idealise machinelike behaviour find that it influences their relationship to the computer in general?
- Are people who idealise machinelike behaviour more comfortable in a rules-governed, machine-mediated interaction than in interpersonal interactions?
- Do those who idealise machinelike behaviour find that it affects their interpersonal relationships? How?

#### **Research Questions for Section 2.6.5 on page 26**

- Does adherence to rules become the sole responsibility (or an overriding responsibility), in cases where agents are following rules regarding retention / destruction?
- Does following rules supersede normative concepts of rightness?
- Does the imposition of regulation bring about amoral behaviour?

- Is behaviour significantly different between organisations with information destruction / retention policies and organisations with no formal policy?
- In organisations with destruction / retention policies, do people think about duty or responsibility to *others*, or merely about duty or responsibility to *follow policy*?
- Are people “contaminated” by previous employment? Do they carry over their behaviour from one position to the next, with regards to their management of digital objects?
- Ought this study include something along the lines of the “Social Reflection Questionnaire” referred to by Trevino? [WT95, p. 606]

### Research Questions for Section 2.7 on page 29

- Can we observe people making *ought* statements with regards to their management of digital objects; can we observe a difference between their stated intention and their actual behaviour?
- To what degree is this effect at play with regards to management of digital objects; do people treat decisions to retain or delete *qua* moral judgements?
- Is it possible to quantify the degree to which people deviate between statements of *ought* and actual practice?
- Will this effect impact the validity of this study?

### Research Questions for Section 2.8 on page 30

- Which features are most commonly used as selection criteria for decisions to delete? To retain?
- Which features are present as part of the object, and which are present within the mind of the subjects and only brought to mind by some aspect of the object?
- Can we witness instances of “extremeness aversion” or “asymmetric dominance” at work in decisions to delete or retain digital objects?

### Research Questions for Section 2.9 on page 33

- Are single digital objects more likely to be selected based upon visual attributes, and groups of objects by verbal or conceptual attributes?

- Are decisions to delete different than decisions to retain, in terms of how selection is performed? Do people select single items for deletion more readily than multiple items? For retention?
- If people select differently for deletion than for retention, does this selection process bias towards retention or deletion solely based upon how the object is presented?
- Is there a difference in how subjects engage—on the level of duty or responsibility—when selecting single objects as opposed to groups of objects?
- If there is a difference in engagement, can we confirm *that* the difference is due to the type of attribute (visual or conceptual) used in making the selection?

### Research Questions for Section 2.10 on page 34

- Do people retain certain types of digital object for their diagnosticity more than others?
- Is data which has been retained because of its possible diagnosticity privileged, in terms of its not being considered for deletion?
- Are people more likely to consider duty or responsibility when deciding to delete or preserve series of data than when making similar decisions with regards to singleton digital objects?
- If people are more likely to consider others from one context to the other, is this brought about due to the perceived value of one type of data as compared to the other? The diagnosticity of one type of data vs. the other?

### Research Questions for Section 2.11.1 on page 36

- What sort of decision process is involved in deciding to delete or preserve, in terms of the level of conscious vs. unconscious decision?
- What do agents consider, in examining a record, in order to tell them whether they must engage with that record on a more conscious level?
- If their relationship to the digital object is mostly conducted on an unconscious basis, what are the “triggers” which cause processing to switch into the conscious realm?

**Research Questions for Section 2.11.4 on page 45**

- Can digital objects be clearly categorized in terms of their exemplification of *an object about which I have a duty or responsibility*?
- Is there a prototypical *object about which I have a duty or responsibility*?
- Is there more than one prototypical *object about which I have a duty or responsibility*?
- Do people attach duty and responsibility to prototypes of objects (e.g., “personnel files”) or to individual objects?
- If there are multiple prototypes, and duty / responsibility attaches to the prototype and not to the instance except via the prototype, does the most salient prototype determine the level of duty or responsibility felt towards the object? Or is the individual aware, simultaneously, that the object exemplifies more than one prototype?
- Do people consider different properties when evaluating files within different folders?
- More generally, does a change in *effective context* cause any significant change in properties considered?
- If changing *effective context* alters selection / rejection criteria, is the reverse true as well: does citing something in a single *effective context* cause the user to consider *the same* properties (e.g., when different domains are all cited within a web browser)?
- In considering *valueless features*, it would seem that this phenomenon would be more likely to occur during group settings. As such, is it even within the domain of this research?
- Can we imagine any relevant *compromise effects*, to be characterized and witnessed?
- Can we imagine any relevant *contrast effects*, to be characterized and witnessed?
- Can we imagine any relevant instances of the use of *prior probabilities*, such that we might encounter instances where subjects ignore *prior probabilities* in favour of descriptive information?
- Are different types of digital object treated differently, with reference to the phenomenon of *anchoring*?
- Do people form multiple anchors, associated with different types of data?

**Research Questions for Section 2.12.2 on page 53**

- Does *framing* affect conceptions of duty or responsibility?
- Do subjects engage different types of thinking about duty and responsibility when the situation is framed as a possible loss versus a possible gain?
- Do subjects who perceive themselves to be *experts* engage in any greater or lesser degree of engagement with concepts of duty and responsibility, with reference to their deletion / retention decisions?
- Do *experts* make better decisions with regards to deletion / retention decisions?
- If, as section 2.12.1 states, experts engage with their task less from a scientific perspective, and more from a ‘gut feel’ perspective, are they also engaging with the deletion / retention task more or less so than non-experts with reference to duty / responsibility?

**Research Questions for Section 2.13.2 on page 56**

- With reference to the disjunction effect, at what level are decisions to delete or retain are being made: do agents engage in logical thinking about objects under their control, or do they perceive there to be ambiguity in the situation, and therefore act upon some other means of assessing the situation (i.e., “their gut feel”)?
- Are there particular ambiguities which are inherent in different types of digital object? For example, does relational data hold a greater amount of ambiguity simply because it *is* relational data?

**Research Questions for Section 2.14 on page 57**

- If we can predict the ways in which people will mistakenly decide to preserve or delete a digital objects, how can we modify the technology in such a manner as to eliminate this effect? Is it *desirable* for us to eliminate this effect?
- How can we address the situation of illogical decision-making so as to bring about the desired outcome? And what is that desired outcome?
- If deciding to delete or retain an object *not* based upon some logical chain of evaluation, are users *more likely* to consider duty and responsibility to be salient to their decision?

**Research Questions for Section 2.15 on page 58**

- Does this treatment—of records as means of society building—hold true with regards behaviour within organisations? Not just on the web?
- Does behaviour with regards to the *born digital* translate into standards of behaviour within organisations?
- Is the expectation that this standard—the Web 2.0 standard—will become the norm?
- Where do conceptions of duty and responsibility fit, when interaction surrounding this data is so sporadic?
- Does this disconnection within the user community mean that this type of data is destined to remain, lingering about, until someone decides to “clean it up?”

## 3.2 Study Aims

This study seeks answers to the following questions<sup>4</sup>:

- What do people believe their motivating beliefs to be, regarding digital object which they delete or preserve?
- Do they take others into consideration when carrying out their deletion or preservation act?
- How do people feel about the possibility of being given a more advanced information management system (without specifying what that would entail)?
- How do people explain their actions with regards to ethical grey areas (such as keeping copies of work-owned digital objects on their home computer, maintaining copies on their work computer outside of the approved records management system, taking work-owned digital objects when leaving a job, keeping personal objects on their work computer)?
- Is there a relationship between computer hygiene<sup>5</sup> and people's beliefs and practices?
- Does "home culture" influence retention/destruction decisions, and in what way?
- Do a variety of different factors (e.g. demography, employment industry, employment responsibilities) correlate with particular beliefs, behaviours, and attitudes?
- Do "hygiene behaviours" correlate with any of these other factors?
- What sorts of formal, records management training have the participants received, when, and of what quality; can this training be seen to have an influence upon beliefs, behaviours?

In focusing upon questions of belief and intention, this study may provide insight into how change may be brought about in behaviour; change to address unintentional deletion or preservation will not be considered. These questions are all attempting to examine the connection between intention, belief, and behaviour; this represents a narrowing of the study focus to a series of research questions which are closely related, answers to which were perceived as being valuable, which questions would make logical sense to participants when presented in a single participation, and about which there has been little study.

---

<sup>4</sup>These study aims are related to the research questions as stated in section 1.2 on page 6 but have been refined based upon the literature examined and also so as to present a set of questions more closely related to one another, for the purposes of the study. The study aims may be regarded as a refinement of the research questions initially set forward to shape the literature review portion of this research.

<sup>5</sup>"Hygiene behaviours" encompass such activities as regularly filing, organising, preserving, deleting digital objects.

These research questions do not take into account cases in which behaviour is not purposeful—cases in which decisions are made out of neglect, laziness, or in which accidental deletion or preservation take place<sup>6</sup>. This is by design, as it was felt that the study could not accommodate such questions without extending the size and complexity of the study, besides which such questions do not concern intentionality, at least not directly<sup>7</sup>.

### 3.3 Different Means of Fulfilling the Study Aims

#### 3.3.1 Quantitative Study

In general, quantitative methods are well suited to answer questions whose objects are *who*, *what*, *when*, or *where*, but not very well suited to answering questions which concern *how* or *why*. This is because quantitative methods require that behaviour be quantified or quantifiable in some manner. For the aims of this study, this would require that all behaviour be expressed in such a manner as to facilitate quantification, or would require that responses be codified in such a way as to allow quantification of the coded responses. The survey was constructed to allow application of either qualitative or quantitative methods.

Quantitative methods are deductive in nature: an hypothesis is deduced after some period of study within a particular domain, which hypothesis is then tested via empirical means [Bry08, p. 9]. Quantitative methods require a strong hypothesis, expressed in terms of measures which would indicate some conclusion regarding the validity of the hypothesis; the hypothesis must form the basis of the study, meaning that the behaviour under examination must be known *a priori*, and that something about the behaviour will be proven via the study. This hypothesis is tested against a (usually) random sample of the group of participants being studied, with preference given to large sample groups in order that statistical confidence may be achieved.

Quantitative methods may result in a definite conclusion as to whether there were a relationship between the variables being studied<sup>8</sup>; however, the results would not necessarily be explanatory in terms of causation: they would demonstrate merely *that a relationship exists* not *what that relationship consists in* nor *which factors are causal*, the direction of the causation, nor the nature of that causation; this is one of the more difficult to address concerns in quantitative studies.

---

<sup>6</sup>Although deviations in hygiene behaviours as they correlate to object type were considered.

<sup>7</sup>It could be argued that poor practices contribute to accidental deletion or preservation; pursuing this was felt to be secondary to the aims of the study.

<sup>8</sup>Please note that this statement should not be interpreted to imply any lack of such a finding upon the part of qualitative methods.

### 3.3.2 Qualitative Study

In general, qualitative methods are well suited to answer questions whose objects are *how* or *why*. This is at least in part because qualitative methods tend to elicit more detailed, individual responses, emphasising the value of the participants' voices and opinions. Qualitative methods expect that participants have knowledge about their own beliefs and motivations<sup>9</sup>, can (and will) share these beliefs and motivations with researchers, and that knowledge of those beliefs and motivations is an important domain of study. This stance regarding the value of the individual participant places the researcher in a different relationship in terms of power and authority than that of quantitative study, which takes a more remote, removed, and possibly objectifying stance towards study participants.

Qualitative methods approach the study of behaviour not with hypotheses to be tested, but with the aim of generating theory about behaviour, or the aim of gaining a rich understanding of the population being studied. They are inductive methods, resulting in the generation of theory from observation of a phenomenon [Bry08, p. 11]. Qualitative methods tend towards smaller sampling groups, and allow for more casual sampling methods, e.g. such as that of "snowball sampling" wherein "the researcher makes initial contact with a small group of people who are relevant to the research topic and then uses these to establish contacts with others" [Bry08, p. 184]. The richness of description and depth of understanding are seen as important reasons for choosing qualitative methods over quantitative methods.

Qualitative methods, though, are subject to criticism. Because the researcher may be quite closely involved with study participants, there is risk of the introduction of bias on the part of the researcher. Because of small sample sizes and lack of random sampling, qualitative methods may only really describe the sample population, and any theory generated may not be generally applicable to any other group. Qualitative methods generate results which are difficult to prove valid even to the sample participants, much less to a larger group; this is true in particular because qualitative study tends to ask about beliefs, thoughts, and opinions, rather than studying actions, but also because such study lacks some of the objectivity and rigour present in quantitative studies.

### 3.3.3 Mixed Methods Study

There are a variety of ways in which to take advantage of the strengths of both quantitative and qualitative methods, while attempting to address some of the criticisms and weaknesses of each. Generally, this type of study consists of multiple, smaller studies, with at least one utilising a

---

<sup>9</sup>This paper will not attempt to defend that premiss.

qualitative method of study and one utilising a quantitative method of study [Bry08, p. 607]. Usually, mixed methods are deployed serially, with one method being the primary method, and the other brought in to support the conclusions of the first study [Bry08, pp. 607,611]. This may be a study wherein the quantitative study is the primary study, which is then supported with description and causal narrative obtained by qualitative study. It may also look be the reverse: a qualitative study used to generate theory which is then refined into hypotheses which are subsequently tested via a quantitative study. The reasons given as justification for pursuing mixed methods may be along the lines of one “proving,” “triangulating,” or “validating” the other (in the case where the qualitative method is the initial method), or in order to add richness and description to an otherwise coldly analytic study (in the case where a qualitative tool is deployed in support of the quantitative tool) [Bry08, pp. 607,609,622]. “The most common form [...] is that the inferences that are derived from a qualitative study are then subsequently tested with quantitative research [Bry08, p. 622].”

There is a philosophical argument against mixed methods research, which basically says that each method subscribes to different epistemological and ontological views, which views are incompatible [Bry08, pp. 604-606]. In addition to arguments regarding the benefits of qualitative vs. quantitative data collection techniques, there is also the concern that mixed methods studies are more difficult to conduct, requiring the application of at least two different study methods; also, this type of study requires significantly more time and effort on the part of the researcher deploying the studies, if only because at least two discrete studies must be performed [Bry08, pp. 606-607,622]. There is also the concern that mixed methods may divide and dilute the research focus from either purely qualitative or purely quantitative aims, resulting in a weakness in both portions of the study [Bry08, p. 606].

### 3.3.4 Methodological Selection

Home culture may have some influence upon retention and deletion decisions; specifically, people who live in cultures regarded as having a strong tradition of records management may make decisions in this area differently than those who live in cultures which are believed not to have such a tradition. If such traditions influence retention and deletion decisions—or if significant variation between such cultures were to be found—it would be a significant finding, as such a finding would indicate that different strategies would be required depending upon the cultural context<sup>10</sup>.

---

<sup>10</sup>Note that this study will not argue as to whether any one culture actually possesses a stronger tradition of records management—it will explore only the differences between the cultures involved in the study, making no claim with regards to whether cultural differences may be ascribed to cultural traditions of records-keeping.

Conducting a study across such a variety of cultures, however, broadens the scope of this study immensely, making it logistically difficult to conduct a randomised, quantitative study with adequate rigour unless some other limiting factor were introduced (e.g., studying only a certain industry, or only certain companies within a particular industry located in different countries). The breadth of the study would seem to argue in favour of the use of qualitative methods, simply because large enough sample sizes could not be achieved using purely quantitative methods—at least, not with the time and financial constraints of the research project.

It was initially decided to make use of qualitative study methods, because they seemed to best fit with the study aims, and because qualitative methods seemed the likeliest to generate sufficient data for analysis. In particular, “grounded theory” was chosen as the foundational methodology, primarily because its goal is the generation of theory about human behaviour. Because this study seeks to generate theory which is explanatory of human behaviour, and because this is one of the stated goals of Grounded theory [Bry08, p. 544], this methodology seemed the best fit for the intended purpose of this research.

Initially, however, surveys were constructed with a large number of closed questions<sup>11</sup> interspersed with open questions to allow the participants to explain why they made their selection in the closed question<sup>12</sup>. They were constructed in this manner in the hopes of achieving a sufficient number of responses as to be able to analyse the data quantitatively *as well as* qualitatively, thus allowing both the definite conclusions of quantitative methods and the richness of response of qualitative methods.

Because the survey instrument was constructed to allow for mixed methods research, and because theoretical sampling<sup>13</sup> may be applied to either quantitative or qualitative research, the survey data was such that it could be analysed using either means of study. Data analysis was finally conducted using quantitative analysis, with open questions coded according to grounded theory; open questions were also analysed individually, to provide insight into the reasoning of participants.

Whether this methodology constitutes qualitative study or mixed methods study, however, is unclear. Such a mixing of methodologies may be subject to criticism by adherents of all methods described, for a variety of reasons.

---

<sup>11</sup>‘Closed questions’ are questions requiring a discrete answer, as opposed to ‘open questions’, which allow for an explanatory answer rather than a selection from a closed set of responses.

<sup>12</sup>This will be discussed in some detail in section 3.6.1 on page 81

<sup>13</sup>See section 3.3.5 on page 74.

### 3.3.5 Grounded Theory, Described

Grounded Theory provides an iterative process wherein the researcher engages with the participants through the use of some study instrument (e.g., a survey or interviews). This engagement is performed using *theoretical sampling* techniques until *theoretical saturation* has been achieved throughout all categories. Theoretical sampling involves obtaining new samples only where it is expected that a new sample of a particular type or with particular attributes would prove useful to the generation of ideas; it is *not* concerned with sample size; it is purposive, as opposed to random sampling [Bry08, p. 415]. Theoretical saturation of a category is achieved when results for a category have become predictable, or where there are very few new or different responses within a given category [Bry08, p. 416]. At the point of saturation, new responses are no longer sought, yet the researcher continues to engage with the responses via different forms of coding and analysis.

The central activity of grounded theory analysis is the coding of responses, which takes place in three distinct phases, and with different purposes to each type of coding. During the first phase of coding (open coding), the researcher examines each response for the purpose of breaking down, examining, comparing, and conceptualising responses into categories. The second phase of coding (axial coding) involves putting the category data back together, making connections between categories, to yield causes. The third phase of coding (selective coding) involves systematically relating a core category to other categories, validating relationships, further refining and developing categories. Hypotheses<sup>14</sup> about the data are generated during the second and third phases of coding, with the hypotheses generated at each coding step tested against the data, and further sampling performed if necessary [Bry08, p. 545].

There are a range of both practical and philosophical criticisms of grounded theory. Practical considerations centre about the researcher maintaining objectivity, the length of time needed to conduct qualitative study (although this seems primarily to be a concern regarding interview transcription, where interviews are the study instrument), and that “most grounded theories are substantive in character [...] they pertain to the specific social phenomenon being researched and not to a broader range of phenomena [...] [Bry08, p. 549].” Philosophically, grounded theory suffers criticisms because its epistemology and ontology are based in constructivism and interpretivism. Epistemologically, grounded theory involves gathering knowledge directly from the statements of study subjects, rather than by observing them or testing them through some deductive means; this presents problems for the justification and truth of such statements.

---

<sup>14</sup>It must be stressed that the hypotheses generated are not to be tested by engaging with study subjects, but with assuring that the data do, indeed, support the hypotheses. In that way, the hypotheses being tested represent a deductive process; however, because they are not tested against a random group of study subjects, this is merely the introduction of a rigorous process within a larger, inductive process.

Ontologically, criticisms are mainly levelled against interpretivism in general, rather than at how that interpretivism affects research based upon grounded theory practices. Further problems in epistemology are applied to the researcher, in that the researcher is making claims about the world which are founded poorly—the researcher can have no justification in believing the statements, yet is making claims based upon them, calling into doubt the epistemological soundness of the researcher’s statements.

### 3.4 Ethical Considerations

There are several ethical concerns with regards to people’s participation in the study: considerations of privacy; of the fact that individuals would essentially be donating some period of their time to the study, which might be construed as “harm” to the participant; that the study instrument might privilege certain voices over others, in particular that some voices might be given more credence were those participants better able to express themselves as compared to those who are not so able or as compared to those which are reluctant to express themselves for some reason.

Privacy concerns were addressed in the following manner:

- All identifying data about participants (name, email address) were maintained separately from response data.
- Personal information was considered when examining responses—no one examining response data could examine personal information while examining responses.
- The analysis and publication of the response data is entirely anonymous, consisting largely of aggregate information rather than individual responses.
- Personal information (name, email address) shall be destroyed at the conclusion of Ph.D. study.
- The anonymous response data was held on secure network storage and has been retained to facilitate further research.

Because participants were not directly compensated for participating in the study, it was of the utmost importance that the study instrument be both pilot-tested and validated, minimising the possibility that responses would be spoiled. Validation consisted of sitting with a test subject through a thorough examination of the study instrument, listening to their feedback as to what they understood each question to mean, and checking their understanding against

what the researcher *intended* to ask in each question; this feedback was used to further refine any problematically-worded questions. Pilot-testing of the study instrument was done prior to validation, to determine whether the experience of using the study instrument was perceived to be burdensome, and to provide some objective measure of the time required to participate in the study. This feedback was then provided to participants at the beginning of their participation so that they could elect not to participate or could delay participation in favour of a more convenient time or a more comfortable locale.

Participants were provided with aggregate statistics at the end of their participation and with the option to print their responses in hopes that they would find the exercise immediately useful. In addition to finding the exercise of use, participants were asked to help establish the credibility of findings [Bry08, p. 377]; participating in credibility testing consisted in the solicitation of feedback via email, which responses were incorporated into the final data for analysis. This incorporation of feedback about the study instrument itself contributed to the “authenticity” of the study. Bryman [Bry08, p. 379] suggests that research has a political impact, and that that it should be judged with regards to its authenticity:

- “*Ontological authenticity*. Does the research help members to arrive at a better understanding of their social milieu?
- “*Educative authenticity*. Does the research help members to appreciate better the perspectives of other members of their social setting?
- “*Catalytic authenticity*. Has the research acted as an impetus to members to engage in action to change their circumstances?”

Lastly, participants in the survey were offered both a version of the thesis in summary form and *in toto*, so that they might have a chance to benefit from this research directly.

Of minor ethical concern was the possibility that some of the questions would be troubling to the participant. In particular, because the study was concerned with reasons for information retention and destruction, some of the questions involved asking the participant to reflect upon any emotions involved in retention or destruction decisions. This invocation of emotion was minimised via the use of closed questions where possible, so that the emotional content was somewhat more contained than would be the case if participants were asked to provide an extensive narrative concerning their behaviour, beliefs, and feelings. The use of few open questions (which would require participants to express themselves in their own words) was also preferred in order to address concerns regarding equality of access and expression, and to avoid privileging those participants with greater verbal acuity or fluidity.

Participants were fully informed as to the domain of study (i.e., their home and work computer use) and the intent of the study (i.e., to learn about their beliefs and motivations with regards to a variety of digital objects). They were provided with an estimate of the time they might expect to spend participating in the study, and with the means to contact the researcher and the researcher's supervisors should the participant encounter any problems or have any complaints.

Ethical approval was sought from the university prior to soliciting participation in the study.

## 3.5 Alternate Methods Considered

Several different methods were considered for the execution of this study. Some methods available were obviously unsuitable<sup>15</sup> and will not be discussed here. The methods discussed below were each evaluated in terms of their strengths and weaknesses with respect to the study aims, with respect to the amount of bias possibly introduced by the researcher, as to their feasibility, in consideration of how much impact they would have upon participants, and with regard to how they would affect the study results. The decision upon study method was made primarily based upon how well the method would satisfy the study aims. The final study instrument selected was that of mixed-method surveys, to be discussed in section 3.6 on page 80.

### 3.5.1 Participant Observation

In order to address the aims of this study via participant observation, the researcher would need to combine participant observation with an interview process. The researcher would need to sit, observing a study subject as they went about their work, and to take notes upon the things saved by the subject (and where they were saved) and upon the things deleted by the subject; after this period of observation, the participant would be asked to explain why they made each decision. Alternatively, a study subject would need to be asked to perform a series of exercises in which they were asked to make retention / deletion decisions, which session would be recorded for later discussion and explanation.

Participant observation has the advantage of inquiring about real actions which the participant has clearly performed and about which must have been some reasoning. This is also a disadvantage, however, because asking someone to justify their decisions may be perceived as a criticism of their actions, and may lead to disingenuous answers simply because the reasoning may have been ephemeral or performed within the subconscious. Participant observation—

---

<sup>15</sup>E.g., those methods which rely upon extant documents as their primary objects of study (content analysis, semiotics, discourse analysis, hermeneutics).

unless of the scripted exercise type—also has the disadvantage that not all types of digital object will be considered, nor would the different contexts of home and work be easily studied using participant observation. The necessity of observing a study subject both at home and at work would require that each participant not only donate several hours of their time towards this study, but would require that they obtain permission for observation to take place at their place of work, and would need to open their home to the researcher; this level of commitment seems unreasonable in light of other methods available.

### 3.5.2 Ethnography

Ethnography was considered as a means of addressing the research questions. Ethnography generates truly rich insights into a particular culture or organisation, simply because the observer is present and participating as a member of that culture or organisation for a prolonged period of time; over time, the researcher gains trust and is granted access to privileged information, which information might not otherwise have been shared. The researcher in an ethnographic study takes part in the society as a member, and learns *what it is* to be a member.

For the purposes of this study, ethnography would have needed to be supplemented with interviews, or with some other study instrument, simply because of the physical impossibility of being simultaneously present within multiple geographies. The ethnographic portion, then, would likely perform a secondary role in the study, rather than being the primary study method: ethnography could possibly provide a personal account of what it is like to be present within each of several different organisations, but it is difficult to see how those experiences would address the study aims simply *by* being present—the researcher would need to gain access to the thoughts, reasons, motivations, and beliefs of the members of each organisation, which is only possible by querying them directly in some manner. In conducting an ethnographic study, the researcher would need to interview members of those organisations in just the same manner as if interviews alone were the primary study tool, or would need to conduct surveys of the members similarly.

It was felt that ethnography would not contribute significantly to the aims of this study. Also, it was difficult to see how disrupting a series of organisations could be ethically justified in light of other study techniques available. While ethnography would provide a richness in terms of understanding particular organisations, and could have told an important story about the interrelation between those organisations, their members, and their retention / deletion decisions, ethnography would be very limited in the populations considered and would be an unjustifiable burden to participants.

Further, it was difficult to perceive how ethnography would have been possible in terms of gaining access to the “home” context: while possible to gain access to a variety of organisations, gaining intimate access to a variety of families would not have provided anything like a natural interaction with those families. Organisations are accustomed to new members being added to the work context; families are nowhere near as like to do so. Thus, ethnography might have possibly provided insight into work behaviour, but would have been quite awkward in gaining access to the “home” context.

### 3.5.3 Interviews

Both structured and semi-structured interviews were considered as a means of addressing this research. Structured interviews have something of an advantage in terms of keeping the study within the bounds of the study area, enforcing a certain uniformity of responses. Semi-structured interviews allow for follow-up questions, for the researcher to pursue an interesting topic further, and to generate a greater depth of individual response. Structured interviews are less likely to introduce bias, simply because the interviewer is not allowed to interact in as dynamic a manner with the interviewee, yet semi-structured interviews may allow for more candour on the part of the interviewee simply *because* the interview is not scripted.

Interviews have the benefit of generating a greater amount of descriptive data, but this benefit is also a detriment, in that the data would need to be transcribed, analysed, and coded; the transcription process is estimated to take five to six hours of transcription time for every hour of interview time [Bry08, p. 452], in addition to issues of coordination and travel, which would place a limit as to the number of interviews conducted by the individual researcher. With such a limit upon the number of responses as imposed by transcription time, a corresponding limit would be necessary in the scope of the study in terms of either population or in terms of research aims. Transcription time could be reduced if interviews were conducted via email; conducting interviews via email, however, was viewed as something which would privilege certain interviewees over others (i.e., those with greater keyboard skills over those with poor keyboard skills); email interviews are also prey to a certain amount of risk in terms of their dragging out over time, with multiple exchanges taking place over several months (if the interview proceeds one question at a time), or of generating less than desirable responses (if the interview questions are provided in advance).

Given a working theory as to the behaviour’s origins, interviews would be quite beneficial, in particular because other methods may fail to tell the whole story of what is taking place when people make decisions across different contexts; interviews would help to provide exemplars of

individual behaviour, but would generate a depth of data which is not needed absent a working theory as to the phenomena under study, and which may not yield such a theory in and of themselves. Additionally, because this study is concerned with beliefs and values (sensitive subject areas), coupled with questions about computer usage and expertise, the presence of an interviewer might bias the responses given.

### 3.5.4 Focus groups

Focus groups may be viewed as an interview with a group of people. They present some challenges of their own, namely that they are difficult to transcribe, should generally be conducted by multiple researchers (one to act as scribe, one to observe participants and note behaviour, and yet another to mediate and guide the discussion), tend to be dominated by one or two participants, and tend to generate consensus rather than bringing out individual differences. Focus groups are similar to interviews, in terms of the amount and quality of content generated, yet are generally used when consensus is valued, such as in marketing research.

The choice of focus groups would have required a change in study focus away from gathering individual input and towards a series of study aims which were centred about the group or organisation. For example, rather than considering how the *individual* made deletion and retention decisions, the study could have focused upon how various *groups* arrived at such decisions. This would have proven valuable and interesting, perhaps; however, such a change in study aims was not desired.

## 3.6 Research Method: Surveys

Surveys were conducted via an electronic survey<sup>16</sup>. The use of electronic surveys is appropriate due to the fact that the subjects of this study are those who use a computer both at home and in the workplace; access to computing equipment is a precondition to taking the survey by definition of the study parameters; further, it has been found that “there is little evidence that the mode of administration makes a significant difference to the findings [Bry08, p. 647].” The software used to generate and conduct the survey was the open-source software LimeSurvey.

The survey instrument was constructed using closed questions where possible, in order that it might later be deployed in a quantitative manner<sup>17</sup>. Closed questions were preferred, also, in order to minimise the possibility of questions being perceived as burdensome or troubling,

---

<sup>16</sup>With allowances made so that the subjects could save it online and return to complete it later, or could complete the survey off-line and return it via postal mail.

<sup>17</sup>Triangulation might be achieved by a later study. “Triangulation [...] refers to the use of quantitative research to corroborate qualitative research findings or vice versa [Bry08, pp. 379,607].”

and to enable participants to answer questions more rapidly than otherwise; additionally, closed questions were regarded as a means of eliminating any bias in favour of those with better language and expression skills, or in favour of those who are better typists. Several open questions were interspersed throughout the survey in order to gather opinions surrounding questions of possible ethical conflict, such as in the case where the participant has stated that they intend to retain digital objects owned by their employer, or to gather opinions about how they would feel if presented with a better means of organising their digital objects. These open questions allowed for the researcher to characterise survey participants' relationship to their employer, their stance regarding records-management efforts, and their degree of reliance upon desktop-search tools.

Surveys were initially deployed via purposive, theoretical sampling<sup>18</sup>. Participants were encouraged to share awareness of the survey, so as to achieve a greater number of participants via the use of *snowball sampling* [Bry08, p. 459]. The use of snowball sampling was necessary in order to achieve *theoretical saturation*<sup>19</sup>. Respondents were made aware that the study was about “differences in people’s computer usage between when they’re at home and when they’re at work.” This was viewed as providing enough information to be ethically correct while not prompting the participants to consider any particular aspect of their behaviour more heavily.

### 3.6.1 Survey Composition

The design of the survey sought to strike a number of compromises: between covering all research questions completely vs. covering them adequately; between obtaining exact answers to such questions of industry and profession vs. providing the participant with something sensible to them; between asking multiple, indirect “check-questions” and asking more substantive questions.

To thoroughly cover the complete set of possible options—including such things as check questions, asked in a variety of different ways—would have resulted in several hundred questions other than those questions concerning demographics, industry, experience, education, etc. The survey design sought to balance the number of questions between the contexts of home and work, with a single series of questions concerning digital objects which travel between these contexts. It also sought to, within the contexts of home and work, examine deletion and retention practises.

---

<sup>18</sup>See Bryman [Bry08, pp. 184, 185, 415], Brink, Walt, and Rensburg [BWR06, p. 134], McMillan and Schumacher [MS97, p. 397], and Somekh and Lewin [SL05, p. 184] for more detail on purposive sampling and theoretical saturation; and see Patton and Patton [PP02, p. 234] for information on criterion sampling.

<sup>19</sup>‘Theoretical saturation’, here, may be taken to mean the state in which responses in a given category, or across a given axis of measurement, have become fairly predictable. [Bry08, p. 416] Because of the wide amount of variance due to the subject of the study, the fact that three large populations were under study, and because of time and resource constraints, the use of snowball sampling was viewed as the best means of achieving theoretical saturation.

Within *those* contexts, questions were asked regarding beliefs and values surrounding different types of digital object, and about different locations of digital object.

Subjecting the participant to multiple versions of the same question within a particular context was not reasonable; nor was subjecting the participant to several hundred very similar questions. While phrasing questions differently in order to have multiple indicators of a particular concept might have been beneficial<sup>20</sup>, this strategy was abandoned in favour of covering rather more ground than less. In addition, the inclusion of check questions implies that the study participant is untrustworthy either by intention or by inattention; this was thought to be at least slightly unethical, in particular as the set of questions was already so repetitive in asking essentially the same questions concerning the different contexts of home and work; to demand even more of the participants' time and effort could not be justified.

A list of possible survey questions were constructed with reference to the literature review and research questions (section 1.2 on page 6), with each subsection of the literature review generally contributing between five and twenty possible areas of interest<sup>21</sup>. These areas of interest were grouped into subjective categories based upon the similarity of area to be studied. From this master list of questions, the decision was made that only a subset of them could be reasonably answered by direct interaction with study subjects through surveys. Several categories of question were omitted by narrowing the study to questions of belief, motivation, and considerations of others in behaviour. The list was further narrowed to those questions which concerned the deletion or retention of digital objects and which were particularly likely to be performed in both the context of a home computing environment and a workplace environment.

The final survey instrument contained fifty-seven questions. Via the use of filter questions, at most fifty questions were asked of any single participant; for example, if a participant did not have greater than 1 computer in the home, they were not asked whether they shared files between computers at home, or if the number of people in their home was only 1, they were not asked whether they shared their computer with other family members; other filter questions involved branching, wherein a participant was asked 'why' or 'why not' depending upon how they had responded to a particular question (e.g., after having asked whether the participant believed it was right to keep work intellectual property, they were either prompted to answer "why" or "why not" depending upon how they had answered the previous question).

Non-personal demographic questions, such as home culture, were asked in the introductory

---

<sup>20</sup>“Sets of attitudes always need to be measured by batteries of indirect indicators. So too do many forms of behaviour [Bry08, p. 145].”

<sup>21</sup>See section 3.1 on page 62 for a listing of these questions

section. This was viewed as desirable in order to accustom the participant with the use of the instrument. Personal demographic questions were asked at the conclusion of the survey, as research has shown that participants are much more likely to give such information after they have completed the survey than if asked this information at the beginning of the survey [Bry08, pp. 204,208]

In the introductory section participants were also asked to evaluate their own computer skills on a seven-point scale ranging from “expert” to “beginner.” These measures were obtained because “[s]elf-efficacy perceptions have been found to influence decisions about what behaviors to undertake [...] , the effort exerted and persistence in attempting those behaviors [...] , the emotional responses (including stress and anxiety) of the individual performing the behaviors [...] , and the actual performance attainments of the individual with respect to the behavior [...] [CH95, p. 189].” Thus, in asking the participant to evaluate their skill level, two things were accomplished: first, the participant would provide useful information as to their own perception of skill; second, the participant would be subtly influenced to consider each question more carefully. While their self-evaluation of skill level did prove interesting (most participants viewed their skills as above average), the behavioural results of asking this question were the true motivation in asking this question. In later sections they were asked to evaluate their computer skills relative to other members of their household, and relative to others within their immediate work environment.

Survey questions were presented in groups [Bry08, p. 204], alternating questions about the computing environment with questions regarding attitudes, beliefs, and behaviours towards deletion and retention decisions. Questions were organised into the following sections<sup>22</sup>:

**1/12: Getting Started** *This section will ask some general questions to determine whether you are eligible to participate in this survey. (See appendix A.1 on page 158.)*

**2/12: Your Home Computer Use** *This section will examine how you use your computer at home. (See appendix A.1 on page 160.)*

**3/12: Things you keep on your Home Computer** *This section is all about what sorts of things you keep or save when using your home computer. (See appendix A.1 on page 162.)*

**4/12: Information about your home** *This section asks some general questions about your home. These questions will help determine whether attitudes and practices change based upon the composition of the home. (See appendix A.1 on page 164.)*

---

<sup>22</sup>See Appendix A on page 156 for the complete text of the survey.

- 5/12: Things you delete on your Home Computer** *This section is all about what sorts of things you delete or erase when using your home computer. (See appendix A.1 on page 165.)*
- 6/12: Sharing between Work and Personal Computers** *This section is about the things you might share between your personal computer and your work computer. For example, you make take work files home, or have personal files you also keep at work. (See appendix A.1 on page 166.)*
- 7/12: Your Work Computer Use** *This section will examine how you use your computer at work. (See appendix A.1 on page 168.)*
- 8/12: Things you keep on your Work Computer** *This section is all about what sorts of things you keep or save when using your work computer. (See appendix A.1 on page 170.)*
- 9/12: Information about your workplace** *This section asks some general questions about your workplace. These questions will help determine whether attitudes and practices change based upon the composition of the workplace. (See appendix A.1 on page 173.)*
- 10/12: Things you delete on your Work Computer** *This section is all about what sorts of things you delete or erase when using your work computer. (See appendix A.1 on page 175.)*
- 11/12: Information about your Work Position** *These are some general questions about your workplace and about the work you perform there. (See appendix A.1 on page 177.)*
- 12/12: Finishing Up** *Just a few more questions and you're done! (See appendix A.1 on page 178.)*

Based upon feedback from testing the survey instrument against a trial group of participants, visual cues (a graphic of a home, and a graphic of a factory) were provided at the top of each section, as appropriate, to help the participant in recalling the context of the questions being asked in that particular group of questions.

The list of industry sectors used consisted of a portion of the *North American Industry Classification System, 2007* provided by the United States Census Bureau [Uni06], and only used the second level in the hierarchy (99 sub-sectors). This list was selected after reviewing many different lists, all of which were compiled for different reasons, and were not felt to be as all-encompassing, or were too specific to manufacturing concerns (e.g. the *Global Industry*

*Classification Standard* [MSC08]). It was felt that using only the sub-sections, consisting of 99 industry classifications, was an adequate compromise between the needs of the survey and the needs of those taking the survey: to present the survey-taker with a structure consisting of the full list of 1,175 items would be unnecessary and difficult, and would take time away from the rest of the survey.

The list of occupations provided by the U.K. Office for National Statistics [U.K00] was non-hierarchical and, as such (and because it consists of 26,000 job titles), it could not be simplified for use within this survey. The list of occupations provided by the United Nations Statistics Division was also found unsuitable for the needs of this survey, as it blends “work activities engaged in by individuals into two major groups in relation to the institutional unit that produces the output [Uni09].” This was undesirable for the purposes of the survey, because of the desire to compile statistics *across* industry sectors—i.e., to analyse trends across a particular workplace activity, regardless of the service or good delivered by the primary organisation.

As a compromise, the list provided by the *Standard Occupational Classification, 2000* [U.S09] was selected as providing enough variety so as to cover most possible occupations, was hierarchical (allowing a multi-stage decision process by the survey participant), and was also somewhat exhaustive, consisting of all of the occupational classifications for civil-service employees within the U.S. Federal Government. Similar structures were considered from other, large employers, but were found to be too industry-specific, or not exhaustive enough. In testing the survey instrument, however, participants expressed concerns that they were overly constrained in stating their occupation; an open question was asked, instead, and results coded according to the U.S. Office of Personnel Management [U.S09] list.

Several questions were asked concerning the records management practices within each subject’s workplace, and the records management education provided by their employer. These questions were constructed according to principles outlined by JISC infoNet [JIS09]; for example, participants were asked whether their organisation or their department provided records management training, and upon what schedule (as part of the hiring induction of new employees, annually, etc.). It was hoped that these questions would provide a more direct axis of measurement than what could be inferred from descriptions of the subject’s industry or work description, and that this evaluation would be interesting as compared to their individual records management practices. By providing some means of evaluating the records management practices of their workplace, responses could be considered with this as an axis of measurement (and coding), to determine whether such records management education influences retention / deletion decisions in the workplace or at home.

## 3.7 Potential Weaknesses of Surveys as Study Method

Surveys are unable to provide insight into differences between verbal and pictorial presentation of digital objects<sup>23</sup>.

### 3.7.1 Context

Context assumed the presence of solely two contexts: that of ‘the home computer’ and that of ‘the work computer’. Judging by the results, in which participants seemed to behave with something approximating the same behaviour in either context, we must conclude that there is solely one context: that of ‘using a computer’, at least with regards computer usage in relatively private, controlled circumstances, i.e. at home and at work.

This study failed to consider computer usage outside of these two contexts; it failed to consider that participants might view these contexts as roughly equivalent (although certainly that is an interesting finding). How would people behave differently within such contexts as, e.g., a computer at a public library, or at a friend’s house; how do expectations of privacy influence behaviour? These factors were overlooked (or excluded) from the study.

### 3.7.2 Randomness of Subject Selection

Subjects were not selected randomly. This may be questioned: both quota sampling and purposive sampling are subject to the criticism [Bry08, p. 186] that they are not representative because they may under-represent certain groups. Also, a true random selection of participants would possibly allow for more generalisation to the populations under study [Bry08, p. 187].

In addition, subjects tended to be unequally distributed between the three home regions under study; this may be of concern, because generalisations made about all three populations are possibly less applicable to the smallest sample population (i.e., the home region of Canada). Any generalisations made about, e.g., archivists are likewise subject to the same criticism: the samples taken from a particular geography may be over-represented as compared to another geography.

### 3.7.3 Terms Used for ‘Employment’

The study needs to be considered in the light of the fact that certain home cultures may use different terms for the same employment, e.g. ‘archivist’ or ‘records manager’ vs. ‘information manager’. If there can be demonstrated that such terms are used primarily in certain home cultures as opposed to others, it may not be fair to compare those cultures on the same basis.

---

<sup>23</sup>see section 2.9 on page 33.

It may be that some umbrella term should be considered in these cases if any strong statements are to be made about, e.g., “what archivists believe.” If, however, “information managers” do, indeed, respond differently than “archivists”, some explanation may need to mention as to whether there exists a correlation between employment terms and home culture, and to consider home culture along with those employment terms so as to demonstrate that that particular employment term is significantly different to the overall attitude of that home culture.

### 3.7.4 Transformation of Digital Objects

Unlike formal records (and unlike physical objects), digital objects would seem to be more transitory (because of being digital and easily duplicated) and also to be likely to be adapted to other purposes after which adaptation they might be regarded as a different *thing*. While physical objects may likewise be adapted (e.g., a theatre ticket may serve to get one into the theatre, but may also be included in a scrapbook because it serves as an *aide-mémoires*), it is perhaps less common for physical objects to be “remixed” in the same manner as digital objects; digital objects, because of their ease of duplication and modification, may be *transformed* to meet some other purpose than was originally intended. Transformation adds yet another difficulty for this study, because transformation is possibly both destructive and generative: the original digital object may not have any further social significance, but the product of transformation may indeed. Thus, in social terms, the conversation between members of the community continues<sup>24</sup> although the original content may have ceased to exist. For the purposes of this study, however, are we to consider transformation an act of deletion? If the intent in transforming a digital object is to generate new objects, that would not be a fair assessment.

This study did not consider transformation or re-purposing of digital objects; this omission may or may not be an oversight, as the act of transformation, re-purposing, or revising digital objects involves both aspects of deletion and preservation. When considering physical objects—e.g., a draft of a manuscript—it is possible to consider individual drafts of the object; with digital objects, this is not necessarily the case. With physical objects, the decision to delete or preserve is a discrete action whereas with digital objects both actions may take place simultaneously. For the purposes of this study, we may have enquired as to the modification of digital objects, but would then have needed to understand what such modification should be considered: deletion or preservation. Certainly, content is destroyed in modification of a digital object; however, this study is interested in the reasons given for deletion, yet the modification process would seem to fall outside of the realm of a decision to delete because it is likely to be considered

---

<sup>24</sup>See 2.15 on page 58.

generative rather than destructive. Thus, to include the modification of digital objects in this study would have broadened the study and possibly confused the results. While this omission was an oversight in that such modification was not considered during the formulation of the study, it is likely that such modification would have been excluded as not within the scope of the study in any event.

### 3.7.5 Frequency of Reasons Used in Decisions

While participants were asked which reasons they have given for deletion or preservation activities, this study did not ask with what *frequency* such reasons were used as they were applied to a particular decision. Such frequency information (if obtainable and reliable) could have proven valuable if only because frequency of behaviour would help to provide a clearer picture of possible differences between work and home behaviour: the mere fact that participants may have accepted a personal reason as valid for both home and work is not adequate to determine whether there is a difference in behaviour except at the gross level—if participants accept, e.g., ‘I am emotionally attached to it’ as a reason for preserving something in both contexts, yet that reason is used with different frequency between the contexts, we may be inclined to infer that there is a difference in behaviour.

Alternatively, the collection of such frequency information may not have been useful due to the fact that participants may not be able to recollect frequency very accurately, on one hand, and that they should be expected to be relating to different digital objects between the contexts, on the other. Thus, some alternate means for collecting this information may have been more desirable than simply asking for frequency information via a survey; for example, some limited number of participants might have been asked to perform a detailed analysis of their behaviour during the course of a day or week, with said analysis being conducted over some extended period of time. In addition to this more reliable method of determining frequency, those participants should be asked whether they deal with certain types of information, and perhaps to explain a bit about why certain reasons are *not* used within a given context: it may be the case that the participant simply does not relate to information about which they feel passionately attached at work, or it may be that they are operating within a regulated environment, etc.

With accurate frequency information, and with detailed descriptions as to the reasons both given and not given, the analysis presented in this study may have been better able to determine whether there is a difference in decision-making between contexts. As this study stands, however, it is only able to state definitively *that* a certain reason is given for decisions within a particular

context; while such statements are revealing as to the broader decision-making process, they do not examine that process in such a manner as to be able to quantitatively differentiate between decisions made in different contexts.

### 3.7.6 Reasons Used for Different Object Types

Considering that the contexts do not appear to have a significant effect upon behaviour<sup>25</sup>, it might be that deeper consideration needs to be given to the sociological aspects of the problem<sup>26</sup>. If participants view certain activities performed on the computer as social interaction more so than as business or personal activities, then we may have been able to explain the response information if the study had asked which reasons were given for various classes of digital object rather than asking which reasons were given within each of the larger contexts of “home” and “work.” If different types of digital object elicit different decision-making processes because they are related to in some cases for the purpose of relating with the other members of a community and in other cases because they are either pertinent to job function or to some personal activity, then the survey should have considered digital objects not by file type or format but by function or purpose.

Examining objects by the function or purpose they serve within a given context, however, is insufficiently explanatory of the decision-making process: reasons given for decisions made with regards to the function of a given object may be affected by the organisational culture and by questions of whether such decisions benefit the individual or others<sup>27</sup>. In such cases wherein the decision to delete or preserve involves some benefit, and wherein a the outcome of a decision is ambiguous or “elastic [Wil11, p. 159],” questioning the participant about such decisions may result in false responses: an alternate explanation may be presented rather than an admission as to the full panoply of reasons, as participants may elect to view themselves as having behaved ethically in their decision-making. Asking about such decisions with regard to object function may also elicit idealised responses rather than revelations as to the particular action performed.<sup>28</sup>.

While the study focused on object type rather than object function, it is unclear whether focusing on object function would have been possible or whether such a shift in focus would have revealed more with regards to decisions to delete or preserve. An examination of decisions may require a different method of study, considering that such decisions would seem to engage the participant in a greater degree of speculation as to their reasoning rather than asking them

---

<sup>25</sup>See section 6.4 on page 137.

<sup>26</sup>See section 2.15 on page 58.

<sup>27</sup>See section 2.6.5 on page 26.

<sup>28</sup>See section 2.7 on page 29.

to consider whether they have used a particular reason for deletion or preservation within a given context, *simpliciter*.

Additionally, object function would likely engage the participant in considerations of their particular expertise, which would in turn affect results by asking participants to consider objects over which they have made decisions in consideration of their perceived competence<sup>29</sup>. Decisions made within and without areas of perceived competence are likely to result in different decisions being made and, thus, different reasons for having made those decisions. The results of a broad examination of “reasons for deletion or preservation of an object which serves XX purpose,” then, would seem to necessarily vary between participants whose particular expertise involves “XX purpose.” Such an examination might not reveal reasons for deletion or preservation—particularly across individuals with different perceived competencies—but would merely reveal the underlying psychological errors invoked by perceived competence.

### 3.7.7 Object Presentation

This study does not take into account that the decision-making process differs when people make judgements with regards to an object which is presented pictorially or verbally<sup>30</sup>, nor does it consider that decisions may be made unconsciously<sup>31</sup>. The former would necessitate the construction of an exercise or series of exercises through which to place participants in which digital objects of equal content were presented in different ways; the latter would be difficult to test except via a similar exercise [...] except that it would be difficult to evaluate *what*, exactly, was being tested.

In the case of verbal vs. pictorial presentation, the study could have been constructed so as to have the participant make the decision and then, after the fact, to enquire as to the reasoning behind the decision. In this way, we would be able to examine the outcome of a series of decisions, and could determine whether such outcomes differed between the two presentation methods. We would also be able to examine the difference in reasons given, which would give us some insight as to how differences in presentation evoke differences in response in accordance with the findings in Gati and Tversky [GT04, p. 127] in which different hemispheres of the brain were invoked by presentation differences because one hemisphere is better specialised for “difference detection” and one for “sameness detection.” Such a finding would permit argument that certain types of decision are better made based upon certain methods of presentation. However, such a finding would not have considered many of the research questions<sup>32</sup>, necessitating that the

---

<sup>29</sup>See section 2.12.1 on page 52.

<sup>30</sup>See section 2.9 on page 33.

<sup>31</sup>See section 2.11.1 on page 36.

<sup>32</sup>See the formalised list of research questions on page 6.

study be supplemented with the kind of information as gathered in the survey.

The next chapter details the implementation of the study. Included in the next chapter are the data analysis methodology and details of how data was consolidated for statistical purposes.

## Chapter 4

---

# Implementation of Study

---

One of the aims of this study was to understand whether there were differences in computer usage which might be attributable to such factors as age / generation, education level, home culture, industry / profession, etc.; because of this aim, surveys were selected as the instrument of study. The survey was constructed in such a manner as to allow for survey results to be analysed either quantitatively or qualitatively—a “mixed methods” approach; the same survey instrument may be re-deployed to a randomised group, in order to prove relationships between variables, should such study be desirable. The survey consisted largely of closed questions, but with a few dozen open questions which could be either coded for quantitative analysis or coded using qualitative methods.

Data was initially analysed in accordance with grounded theory [Bry08, pp. 542,545,549,550]. As surveys were completed, any open questions were coded; responses were then categorised using brief, descriptive terms. As each record was examined, a memo was kept, detailing the researcher’s thoughts about the particular response and how that response seemed compared to other responses. In addition, after completion of the survey, participants were sent a follow-up email thanking them for their participation in the study and asking them to provide any feedback as to the process, their thoughts about the subject matter, etc.; any emailed response was stripped of identifying information and included as part of the response data. Periodically, an analysis of the data (both indicators and categories<sup>1</sup> was performed, seeking relationships across axes of measurement, and to determine areas in which more sampling should be performed; this practise is roughly in line with both quota sampling and purposive sampling [Bry08, pp. 185,415,697].

---

<sup>1</sup>The terms ‘indicator’ and ‘variable’ describe data resulting from closed questions; ‘category’ and ‘code’ describe data generated as a result of coding of open questions by the researcher. ‘Indicator’ and ‘variable’ may be regarded as synonymous, as may ‘category’ and ‘code’.

The survey was closed when theoretical saturation [Bry08, p. 417] was achieved in the areas of age, profession, and home culture. At this point a full analysis was done in order to formulate theories as to the relationships between variables. These theories were tested against the data in an attempt to understand any outliers and to encapsulate all survey responses. From these tested theories a substantive theory was generated as to the nature and causes of the behaviour being studied<sup>2</sup>.

Surveys were deployed from January 2010. The survey was active until September, 2010. Data analysis and hypothesis generation was completed as of November, 2010. At this point, participants were asked whether they would be willing to participate in a brief discussion of findings, in order to test hypotheses for credibility. “The establishment of the credibility of findings entails [...] submitting research findings to the members of the social world who were studied for confirmation that the investigator has correctly understood that social world. This latter technique is often referred to as *respondent validation* or *member validation* [Bry08, p. 377].”. Credibility testing was performed with approximately 5% of the study participants (17 individuals) and consisted in a brief presentation of findings and a discussion with them as to whether they believed it to be a credible representation of their beliefs and actions; their opinions were overall that the study represented their beliefs and actions<sup>3</sup>.

## 4.1 Data Analysis Methodology

Initial coding of responses was performed using QSR NVivo software. However, NVivo did not easily allow multiple imports from the survey software; because of this, data was imported into Microsoft SQL Server 2000 and coded using a Microsoft Access interface to the Microsoft SQL Server database. Data was then exported from SQL Server to plain text format for analysis using PSPP<sup>4</sup>. During the export process, no differentiation was made between variables entered by the survey participants and codes applied by the researcher, although the researcher needed to be *aware* of the difference when analysing the data: when considering whether a particular relationship between a variable and/or code was significant, the researcher needed to pay particular attention to cases wherein one or both items was a code applied *by* the researcher, in particular because these codes were formulated from a close examination of open questions,

---

<sup>2</sup>See section 5 on page 104.

<sup>3</sup>Indeed, several participants wondered why the findings should be remarkable—that *of course* people are individual in how they use their computers, that they keep copies of things, that they ignore policies when it helps them do their jobs, that they use their home and work computers in the same way, etc.

<sup>4</sup>PSPP is an open-source implementation of the statistical analysis package SPSS. It was selected over SPSS not only because it is free and open-source, but because its implementation of certain statistical calculations are supposedly superior in performance to those as implemented by SPSS [Psp], although such behaviour could be overridden if so desired.

rather than the variables which were the product of closed questions. Because the structure of the survey consisted largely of closed questions, codes were in some cases much less likely to have been applied as compared to response-variables. Therefore, in a case wherein a code was only applied sparsely, careful thought was given as to whether any relationship between that code and another code or variable was merely the product of not having sufficient information, rather than indicating a particular trend or relationship; in such cases wherein coding was particularly sparse and likely to have been sparse due to a lack of information, relationships between codes and indicators was regarded as dubious and ignored.

The survey was implemented using the open-source package LimeSurvey. LimeSurvey exports to Microsoft Excel, a variety of text files, the R statistical language, and to SPSS command files. It was found that exporting to SPSS command files did not facilitate coding of the data, nor did it allow for the segregation of free-text answers from more structured answers. Exporting to plain text files caused some issues when importing the data in SQL Server, as the plain text file format deviates from that expected by SQL Server.

The most practical (and repeatable) way for this data to be accessible to both SQL Server and to PSPP was found to be to:

1. export the data to Excel from LimeSurvey,
2. perform a few, basic manipulations of the data using an Excel Macro<sup>5</sup>,
3. export the data to a comma-separated-values (CSV) file,
4. import the CSV file into SQL Server<sup>6</sup>,
5. code new responses in Access,
6. denormalise codes into multiple columns for export<sup>7</sup>
7. export all valid responses from SQL Server to a second CSV file<sup>8</sup>,
8. import the second CSV file into PSPP.

Initial statistical analysis was performed using the open-source software package PSPP. Further statistical analysis was performed using custom-written SQL Scripts<sup>9</sup> Although LimeSurvey provided basic statistics, these were found to be of little use, in particular because these statistics did not allow for the consolidation of multiple variables into higher-level categories, nor did LimeSurvey consider relationships between variables.

---

<sup>5</sup>See appendix B on page 179 written in Microsoft Visual Basic for Applications for Microsoft Excel 2000.

<sup>6</sup>See appendix B on page 181, and appendix B on page 186.

<sup>7</sup>See appendix B on page 191.

<sup>8</sup>A 'CSV' file is a plain-text file containing tabular data separated by commas. See appendix B on page 209 for data export code.

<sup>9</sup>See appendix B on page 179.

## 4.2 Notes on Data Consolidation

Initial data analysis was conducted upon individual questions; this, however, proved to be both too granular and insufficiently explanatory: individual responses varied wildly across participants, with no clear correlation between responses: individual variables considered in isolation (with the exception being that respondents tended to utilise instant messaging both at home and at work) showed no statistically significant positive or negative correlation. After this analysis was examined, the decision was made to consolidate responses into a series of categorised scores; in so doing, individual questions were made somewhat less significant and generalised behaviour was made more apparent (e.g., a response which indicated that 4 of 6 filing activities were practised in the home could be considered, rather than whether individual activities were practised). Higher-level aggregations of responses eliminated a certain degree of ‘noise’ within responses, and allowed the analysis to more easily identify areas of interest. These categories were determined by the content of the questions

Data was consolidated into the following, higher-level categories:

**‘Clean up Home’** With a theoretical maximum of 36, is comprised of responses to questions related to the frequency (on a scale of 0 to 6) of the following “clean up activities” taking place at the home:

1. Email (variable name HomeCleanupEmail)
2. Contacts (variable name HomeCleanupContacts)
3. Digital Objects resident upon the computer file system (variable name HomeCleanupDocuments)
4. Social Network Status (variable name HomeCleanupSocialNetworkStatus)
5. Text Messages (variable name HomeCleanupTextMessages)
6. Mobile Phone Contacts (variable name HomeCleanupMobileContacts)

**‘Deletion Affordances Home’** With a theoretical maximum of 9, consists of responses to 9 yes/no questions concerning whether particular metadata of digital objects was used in making decisions to delete, at home:

1. Whether the digital object was contained within a particular folder (variable HomeHelpDeleteFolder)
2. The file type (variable HomeHelpDeleteFileType)
3. The file name (variable HomeHelpDeleteFileName)
4. The object’s creation date (variable HomeHelpDeleteCreationDate)

5. The object's last modified date (variable HomeHelpDeleteModifiedDate)
6. The sender of email (variable HomeHelpDeleteEmailSenders)
7. Subject line of email (variable HomeHelpDeleteEmailSubject)
8. Whether an email was flagged (variable HomeHelpDeleteEmailFlagged)
9. Whether an email had attachments (variable HomeHelpDeleteEmailAttachments)

**'Deletion Affordances Home 2'** With a theoretical maximum of 10, consists of responses to 10 yes/no questions concerning whether particular *data or metadata* of digital objects was used in making decisions to delete, at home:

1. Whether the digital object was contained within a particular folder (variable HomeHelpDeleteFolder)
2. The file type (variable HomeHelpDeleteFileType)
3. The file name (variable HomeHelpDeleteFileName)
4. The object's creation date (variable HomeHelpDeleteCreationDate)
5. The object's last modified date (variable HomeHelpDeleteModifiedDate)
6. The sender of email (variable HomeHelpDeleteEmailSenders)
7. Subject line of email (variable HomeHelpDeleteEmailSubject)
8. Whether an email was flagged (variable HomeHelpDeleteEmailFlagged)
9. Whether an email had attachments (variable HomeHelpDeleteEmailAttachments)
10. Whether the item's *content* was used in decisions to delete (variable HomeHelpDelete-Content)

**'Deletion Affordances Work'** With a theoretical maximum of 9, consists of responses to 9 yes/no questions concerning whether particular metadata of digital objects was used in making decisions to delete, at work:

1. Whether the digital object was contained within a particular folder (variable WorkHelpDeleteFolder)
2. The file type (variable WorkHelpDeleteFileType)
3. The file name (variable WorkHelpDeleteFileName)
4. The object's creation date (variable WorkHelpDeleteCreationDate)
5. The object's last modified date (variable WorkHelpDeleteModifiedDate)
6. The sender of email (variable WorkHelpDeleteEmailSenders)
7. Subject line of email (variable WorkHelpDeleteEmailSubject)
8. Whether an email was flagged (variable WorkHelpDeleteEmailFlagged)

9. Whether an email had attachments (variable WorkHelpDeleteEmailAttachments)

**‘Deletion Affordances Work 2’** With a theoretical maximum of 10, consists of responses to 10 yes/no questions concerning whether particular *data or metadata* of digital objects was used in making decisions to delete, at Work:

1. Whether the digital object was contained within a particular folder (variable WorkHelpDeleteFolder)
2. The file type (variable WorkHelpDeleteFileType)
3. The file name (variable WorkHelpDeleteFileName)
4. The object’s creation date (variable WorkHelpDeleteCreationDate)
5. The object’s last modified date (variable WorkHelpDeleteModifiedDate)
6. The sender of email (variable WorkHelpDeleteEmailSenders)
7. Subject line of email (variable WorkHelpDeleteEmailSubject)
8. Whether an email was flagged (variable WorkHelpDeleteEmailFlagged)
9. Whether an email had attachments (variable WorkHelpDeleteEmailAttachments)
10. Whether the item’s *content* was used in decisions to delete (variable WorkHelpDelete-Content)

**‘Filing Activities Home’** With a theoretical maximum of 6, consists of responses to 6 yes/no questions concerning whether certain types of “filing activities” take place, at home. Filing of:

1. Email (variable HomeFileEmail)
2. Music (variable HomeFileMusic)
3. Pictures (variable HomeFilePictures)
4. Videos (variable HomeFileVideos)
5. “Records,” in the formal sense of the term (variable HomeFileRecords)
6. Other Digital Objects (variable HomeFileDocuments)

**‘Functional Reasons Home’** With a theoretical maximum of 6, consolidates responses to 6 yes/no questions as to whether the following reasons are reasons the participant would give for deleting a digital object, at home:

1. It gets in the way (variable HomeDeleteGetsInWay)
2. It’s old (variable HomeDeleteItsOld)
3. Someone else has a copy (variable HomeDeleteSomebodyHasCopy)

4. Too much trouble to keep (variable HomeDeleteTooMuchTroubleToKeep)
5. No way to keep (variable HomeDeleteNoWayToSave)
6. Takes up space (variable HomeDeleteTakesUpSpace)

**‘Functional Reasons Work’** With a theoretical maximum of 6, consolidates responses to 6 yes/no questions as to whether the following reasons are reasons the participant would give for deleting a digital object, at work:

1. It gets in the way (variable WorkDeleteGetsInWay)
2. It’s old (variable WorkDeleteItsOld)
3. Someone else has a copy (variable WorkDeleteSomebodyHasCopy)
4. Too much trouble to keep (variable WorkDeleteTooMuchTroubleToKeep)
5. No way to keep (variable WorkDeleteNoWayToSave)
6. Takes up space (variable WorkDeleteTakesUpSpace)

**‘Personal Reasons Home’** With a theoretical maximum of 15, consists of responses to 15 yes/no questions as to whether the following reasons are reasons the participant would give for keeping *or destroying* a digital object, at home:

1. I might need it (variable HomeKeepMightNeed)
2. To remember something important to me (variable HomeKeepImportantMemory)
3. Want to work on it (variable HomeKeepWantToWorkOn)
4. I spent time on it (variable HomeKeepSpentTimeOnIt)
5. Someone else spent time on it (variable HomeKeepSomebodySpentTimeOnIt)
6. Legal reasons (variable HomeKeepLegalReasons)
7. Emotionally attached to it (variable HomeKeepEmotionallyAttached)
8. It’s interesting (variable HomeKeepInteresting)
9. It’s my creation (variable HomeKeepMyCreation)
10. Evidentiary reasons (variable HomeKeepProveActions)
11. No longer need it (variable HomeDeleteDontNeed)
12. It would be wrong to keep it (variable HomeDeleteWrongToKeep)
13. It’s emotionally troublesome (variable HomeDeleteEmotionallyTroublesome)
14. It’s not important to me (variable HomeDeleteNotImportantToMe)
15. Reasons of privacy (variable HomeDeletePrivacy)

**‘Personal Reasons Work’** With a theoretical maximum of 15, consists of responses to 15 yes/no questions as to whether the following reasons are reasons the participant would

give for keeping *or destroying* a digital object, at work:

1. I might need it (variable WorkKeepMightNeed)
2. To remember something important to me (variable WorkKeepImportantMemory)
3. Want to work on it (variable WorkKeepWantToWorkOn)
4. I spent time on it (variable WorkKeepSpentTimeOnIt)
5. Someone else spent time on it (variable WorkKeepSomebodySpentTimeOnIt)
6. Legal reasons (variable WorkKeepLegalReasons)
7. Emotionally attached to it (variable WorkKeepEmotionallyAttached)
8. It's interesting (variable WorkKeepInteresting)
9. It's my creation (variable WorkKeepMyCreation)
10. Evidentiary reasons (variable WorkKeepProveActions)
11. No longer need it (variable WorkDeleteDontNeed)
12. It would be wrong to keep it (variable WorkDeleteWrongToKeep)
13. It's emotionally troublesome (variable WorkDeleteEmotionallyTroublesome)
14. It's not important to me (variable WorkDeleteNotImportantToMe)
15. Reasons of privacy (variable WorkDeletePrivacy)

**'Reasons Involving Others Home'** With a theoretical maximum of 4, this aggregate consists of the responses to 4 yes/no questions about whether the following reasons are reasons which the participant would give for keeping or destroying a digital object, at home. These reasons were felt to consider others in the decision, by definition, and were analysed as a separate aggregate group:

1. It's important to the household (variable HomeKeepImportantToHousehold)
2. Someone else might need it (variable HomeKeepSomebodyMightNeed)
3. No one else needs it (variable HomeDeleteNobodyElseNeeds)
4. It's confidential (variable HomeDeleteConfidential)

**'Reasons Involving Others Work'** With a theoretical maximum of 4, this aggregate consists of the responses to 4 yes/no questions about whether the following reasons are reasons which the participant would give for keeping or destroying a digital object, at work. These reasons were felt to consider others in the decision, by definition, and were analysed as a separate aggregate group:

1. It's important to the business (variable WorkKeepImportantToBusiness)
2. Someone else might need it (variable WorkKeepSomebodyMightNeed)

3. No one else needs it (variable WorkDeleteNobodyElseNeeds)
4. It's confidential (variable WorkDeleteConfidential)

**'Record Generating Activity Home'** With a theoretical maximum of 300, participants were asked to complete a percentage of time spent on various activities, at home, with the acknowledgement that certain activities might overlap:

1. Interacting with email (variable HomeEmail)
2. Documenting personal finances (variable HomePersonalFinance)
3. Documenting other finances (variable HomeOtherFinance)

**'Record Generating Activity Work'** With a theoretical maximum of 300, participants were asked to complete a percentage of time spent on various activities, at work, with the acknowledgement that certain activities might overlap:

1. Interacting with email (variable WorkEmail)
2. Documenting personal finances (variable WorkPersonalFinance)
3. Documenting other finances (variable WorkOtherFinance)

**'Use Default Stores Home'** With a theoretical maximum of 6 (yes/no question responses), this series of questions asked whether the respondent used the following, default storage locations (or the operating-system equivalent for such) for housing digital objects, at home:

1. "My Documents" (variable HomeUseDefaultDocuments)
2. "My Downloads" (variable HomeUseDefaultDownloads)
3. "My Ebooks" (variable HomeUseDefaultLibrary)
4. "My Movies" (variable HomeUseDefaultMovies)
5. "My Music" (variable HomeUseDefaultMusic)
6. "My Pictures" (variable HomeUseDefaultPictures)

**'Use Default Stores Work'** With a theoretical maximum of 6 (yes/no question responses), this series of questions asked whether the respondent used the following, default storage locations (or the operating-system equivalent for such) for housing digital objects, at work:

1. "My Documents" (variable WorkUseDefaultDocuments)
2. "My Downloads" (variable WorkUseDefaultDownloads)
3. "My Ebooks" (variable WorkUseDefaultLibrary)
4. "My Movies" (variable WorkUseDefaultMovies)
5. "My Music" (variable WorkUseDefaultMusic)

6. “My Pictures” (variable `WorkUseDefaultPictures`)

**‘Work RM System’** With a theoretical maximum of 7 (yes/no question responses), this series of questions asked about the Work Records-Management environment:

1. Workplace has an RM system (variable `WorkHasRM`)
2. Workplace has RM policies (variable `WorkHasRMPolicies`)
3. The current employer has instructed you on RM policies (variable `WorkInstructedOnRM`)
4. All employees are instructed as to RM policies (variable `WorkAllEmpsInstructedOnRM`)
5. Email is moved to a records-management system (variable `WorkEmailMoveToOtherSystem`)
6. Documents (separate to email) are kept in an RM system (variable `WorkCentrallyManagedDocuments`)
7. The workplace audits documents kept on individual systems (variable `WorkAuditsDocs`)

**‘RM Violation’** With a theoretical maximum of 3 and comprised of variables:

1. Keep work objects at home, despite RM policy (variable `HomeKeepWorkObjects`)
2. Keep workplace digital objects after leaving the company (variable `HomeKeepWorkObjectsAfterLeavingCompany`)
3. At the workplace, despite any RM policy, keep local copies of centrally-managed documents (variable `WorkCentrallyManagedDocumentsKeepCopy`)

**‘Training Organisation’** With a theoretical maximum of 5 (yes/no question responses), this group of questions asked where the employee obtained their job skills:

1. Formally acquired, during work-time (variable `JobSkillAcquiredFormallyDuringWorkTime`)
2. Formally acquired, outside of work (variable `JobSkillAcquiredFormallyOutsideWork`)
3. Through seminars or conferences (variable `JobSkillAcquiredSeminarsConferences`)
4. The workplace provided instruction on RM procedures (variable `WorkInstructedOnRM`)
5. All employees are provided workplace instruction (variable `WorkAllEmpsInstructedOnRM`)

**‘Training Self’** With a theoretical maximum of 3 (yes/no question responses), is meant to serve as comparison to workplace-provided training and is comprised of the following self-training areas:

1. “On the job” (training provided by the organisation or a coworker) (variable `JobSkillAcquiredOJT`)
2. On their own, whilst at work (variable `JobSkillAcquiredOnOwnAtWork`)
3. On their own, outside of work (variable `JobSkillAcquiredOnOwnOutsideWork`)

**‘Home Work Mixing’** With a theoretical maximum of 6 (yes/no question responses) are a series of questions used to determine the degree to which respondents “mixed” their home and work contexts:

1. Keep digital objects from work at home (variable HomeKeepWorkObjects)
2. Keep digital objects from work after leaving the organisation (variable HomeKeep-WorkObjectsAfterLeavingCompany)
3. Keep digital objects from home on work systems (variable WorkKeepHomeObjects)
4. Perform personal finance activities whilst at work (variable WorkPersonalFinance)
5. Keep a variety of personal digital objects at work (variable WorkPersonal)
6. Keep other, personal, digital objects at work (variable WorkOtherPersonal)

‘Yes/No’ Answers were not always simply a choice between only these two options; sometimes a question was not applicable, and sometimes the participant was offered the choice of ‘Uncertain’. Because summary scores were searching for positive results (e.g., ‘What is the “Filing Activities Home” Score’?), ‘Yes’ was scored as having a value of 1, and all other responses were considered to have a value of 0. This may prove controversial because a lack of response cannot be truly considered to be a negative response, nor can the response ‘not applicable’ be considered as such. In order to alleviate any such concerns, however, conclusory statements regarding responses were framed in terms of positive responses, rather than addressing the responses which may have included definitively negative responses mixed with ‘uncertain’ or ‘not applicable’ responses. Additionally, analysis code has been retained to re-perform the same series of statistical operations with the ‘N/A’ and ‘Uncertain’ coded as desired.

Participants in the survey were asked to place their age within a 5-year range (e.g., ‘25-29 Years’). Similarly to grouping questions into higher-level categories, in performing high-level data analysis these 5-year bands were aggregated into 2 different variables: ‘Age2’ included ‘15-29 Years’, ‘30-49 Years’, and ‘50-74 Years’; ‘Age3’ included ‘15-34 Years’, ‘35-54 Years’, and ‘55-74 Years’. The generation of two different, high-level variables was done because while it was suspected that while there might be a difference between generations, it was not immediately clear as to exactly where such generational breaks might lie; additionally, it was unclear as to whether generational differences might appear at different age levels between different cultural groups. Thus, analysis of ‘Age2’ and ‘Age3’ and their relationships with other categories were conducted independently, in the hopes that any such generational affects would become apparent.

It is dubious as to whether the variables ‘HomeHelpDeleteContent’ and ‘WorkHelpDeleteContent’ should be included in ‘Deletion Affordances Home’ and ‘Deletion Affordances Work’,

respectively, as the content should not really be considered an ‘affordance’, per se. Therefore, the calculations were performed both with and without this variable included, to be determined later whether it is desirable to settle upon one or the other method of calculation. The ‘2’ version of the summary calculations includes the ‘Content’ as well.

‘HomeKeepWorkObjects’ and ‘HomeKeepWorkObjectsAfterLeavingCompany’ fall into two calculations: ‘Home Work Mixing’ and ‘RM Violation’.

‘WorkPersonalFinance’ is included in 2 categories: ‘Record Generating Activity Work’ and ‘Home Work Mixing’. Because this study is searching for context effects in behaviour, it was felt that, while the records generated are of a personal nature and not of concern to the organisation, the context in which they were generated might have some effect upon this behaviour. Also, for the purposes of comparison between the two summary calculations, ‘WorkPersonalFinance’ was included in this category. Further examination of the individual variables ‘HomePersonalFinance’ and ‘WorkPersonalFinance’ may prove of value here.

‘WorkInstructedOnRM’ and ‘WorkAllEmpsInstructedOnRM’ are included both in ‘Work RM System’ and in ‘Training Organisation’.

‘WorkDeleteBusinessDoesntNeed’ has no direct equivalent within the home context. Therefore, calculations were performed including and excluding this variable; in cases wherein comparisons are being made between home and work context, ‘Reasons Involving Others Work’ will be the basis of comparison. Otherwise, ‘Reasons Involving Others Work 2’ will be considered.

‘WorkEmailMoveToOtherSystem’ is included in both ‘Work RM System’ and ‘Email Management Activity Work’. These two categories were not compared to one another, so there was no risk of a particular ‘WorkEmailMoveToOtherSystem’ being given more weight than it ought in the calculations.

‘Record Generating Activity Home’ and ‘Record Generating Activity Work’ were already comprised of solely numerical variables, all of which were on a sliding scale representing the percentage of time spent in each location on various activities.

## Chapter 5

---

# Interpretation of Results

---

### 5.1 Analysis of Consolidated Data

Higher-level analysis of survey data can be considered to fall into three broad categories: what sorts of activities are the participants engaged in with their computers, what built-in features of their computer's operating system help them to perform deletion and preservation activities, and what sorts of reasons do they give as salient to their decision-making processes. Additional areas of interest are whether participants have received formal, records-management training; to what degree they are self- or formally-educated; and to what degree do they tend to engage in activities which might be viewed as bad practice, in terms of records-management.

#### 5.1.1 Demographic Findings

There were no statistically significant differences between different age ranges, the three cultures studied, nor between gender. From this we may conclude that computer skills and practices are fairly uniform across these axes and perhaps that the computer operating systems encourage consistent behaviour in some way. That the operating systems encourage certain behaviour is not unreasonable at all, as computers and/or operating systems may be regarded as actors which modify the social context within which decisions are made<sup>1</sup>.

Because this study was not limited to certain professions, the quantities within each profession were too small to allow for the formulation of statistically significant comparisons between professions. Comparisons between those professions with significant numbers of participants (i.e., 'Educator', 'Records Manager', 'Archivist', 'Technologist') did not result in any significant differences across the other axes of measurement.

---

<sup>1</sup>See section 2.4 on page 15.

Profession	Participants	Profession	Participants
Educator	41	Secretarial	5
Records Manager	41	Lawyer	4
Archivist	36	Journalist	3
Technologist	26	Social Work	3
Medical Professional	13	Other	2
Librarian	9	Curator	2
Civil Servant	7	Business Person	2
Administrative	7	Ecologist	1
Student	7	Marketing	1
Project Manager	5	Recruiter	1

Table 5.1: Breakdown of ‘Profession’ of study participants

Within the population studied, although there was a larger preponderance of American participants than any other group, the age ranges within each demographic were fairly consistent (see table 5.2). Participants, although unequally distributed between the three study areas of the United States, Canada, and the British Isles, were rather equally distributed by age range within each geographic area. Thus, it is not reasonable to expect any variation in behaviour found between age ranges and other variables would be a hidden artefact of geographic variation instead; rather, age and geographic location should be considered independent variables.

Age Range	United States	British Isles	Canada	Percent of Total
15-19 Years	1%	0%	0%	0%
20-24 Years	2%	2%	0%	1%
25-29 Years	10%	20%	19%	14%
30-34 Years	14%	25%	29%	19%
35-39 Years	15%	17%	19%	16%
40-44 Years	15%	11%	10%	13%
45-49 Years	13%	12%	10%	13%
50-54 Years	12%	9%	0%	10%
55-59 Years	10%	2%	10%	7%
60-64 Years	9%	2%	5%	6%
65-69 Years	1%	2%	0%	1%
<i>Responses</i>	130 (60%)	65 (30%)	21 (10%)	216 (100%)

Table 5.2: Age Range as Percent of Respondents within ‘Home Country’

### 5.1.2 Filing Activities Home

Survey participants were asked about whether they engaged in such activities as filing their home email; organising their personal documents, music, photographs, videos, or what they might consider to be formal records (e.g., income tax documents or bank statements). They were not asked at what frequency they engaged in such activities, but merely whether they

performed such tasks. The six questions which comprise this consolidated view allowed for the elimination of some of the outliers (for example, participants who did not store a particular type of media on their home computer), while still gaining some insight into the presence of such activities.

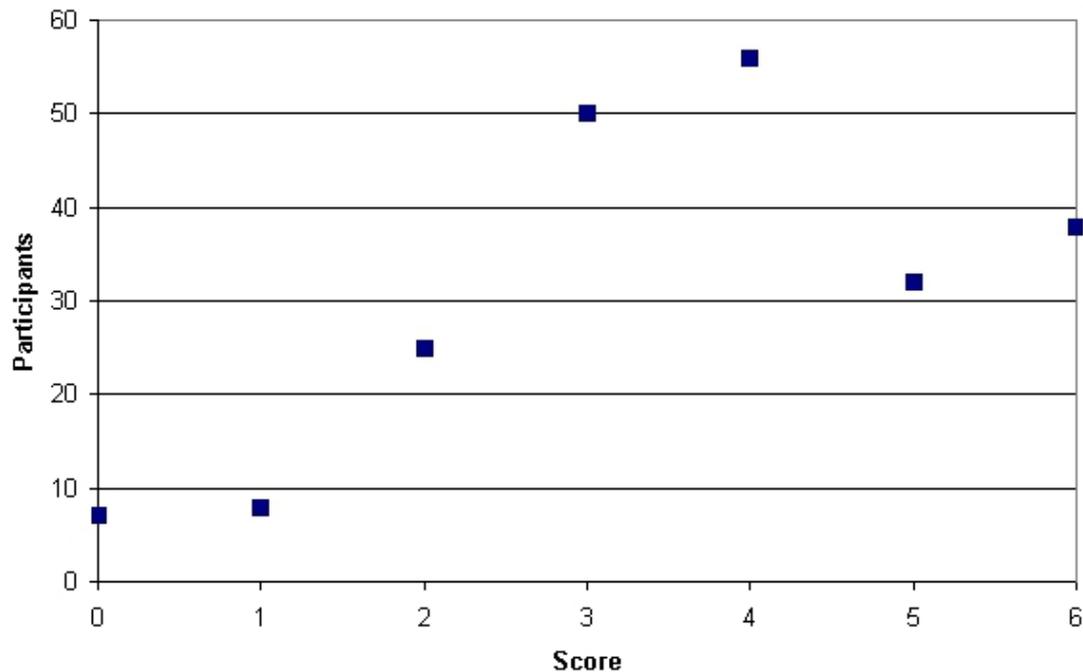


Figure 5.1: Filing Activities Home

Only 7 of 216 participants performed no filing activities whatsoever, with the majority of participants performing 4 or more such filing activities (see figure 5.1). This indicates that, of the sample group, people are generally invested in maintaining some sort of an organisational system for their own use.

There were no statistically significant differences between different age groups in the number of home filing activities performed (see figure 5.2). This indicates that, while reasons given may differ, participants from all age ranges are relatively similar in their desires to maintain some sort of organisational system. Had there been a generational effect, this should have been shown in a difference in the number of filing activities performed by the participants<sup>2</sup>. As none was found, we may conclude that age does not help determine the degree to which participants are interested in organising their documents at home. We may also conclude that filing behaviour is relatively similar between generations of computer user, thus indicating that the age at which computer skills were acquired may not play a significant role in how computers are used: if older computer users had acquired their computer skills on older computer systems and these

<sup>2</sup>Indeed, in analysing the individual variables which comprised the aggregate variable, no statistically-significant correlation was determined to be present.

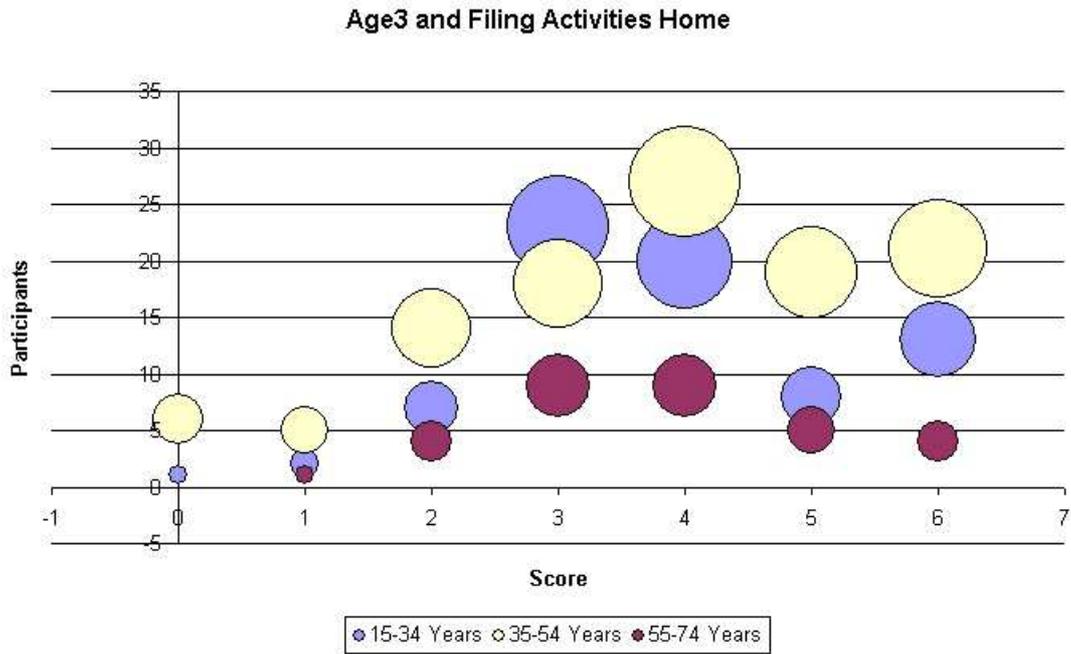


Figure 5.2: Age3 and Filing Activities Home

skills carried forward into their current usage, we might have expected to find a difference in filing behaviour.

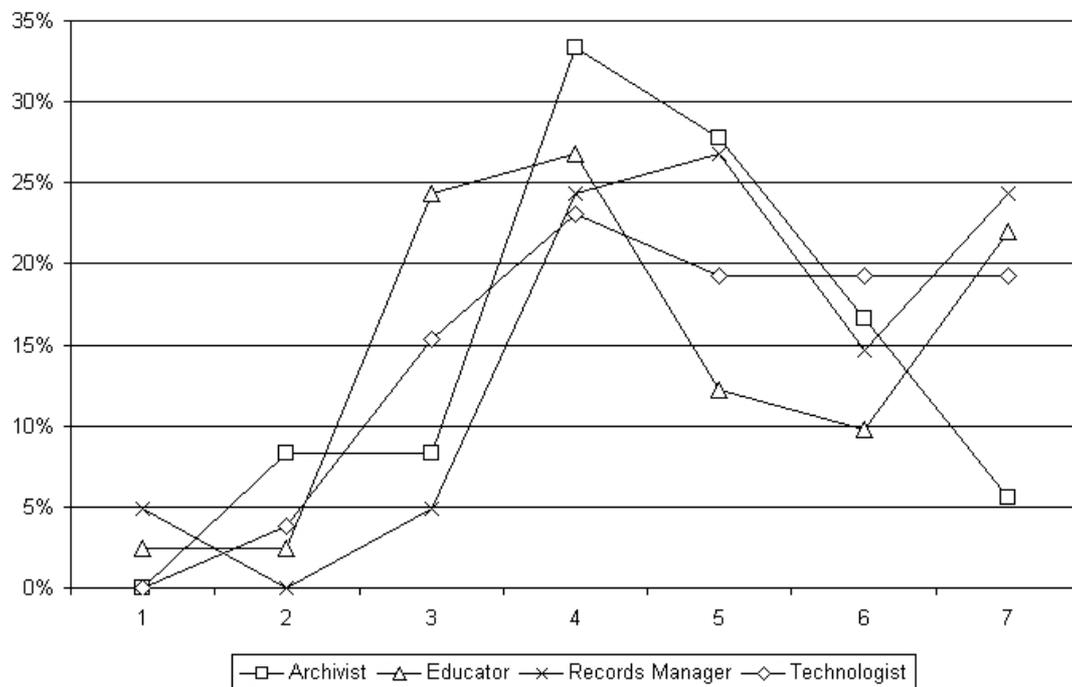


Figure 5.3: Profession and Filing Activities Home

Likewise, there were no major differences between profession in terms of their home filing activities (see figure 5.3). Archivists and Records Managers were more likely to perform these

activities as compared to Educators and Technologists<sup>3</sup>, but not by any large percentage. This is particularly true given that there were only 36 archivists and 41 records managers; 3 or 4 individual responses could account for a difference of 10%.

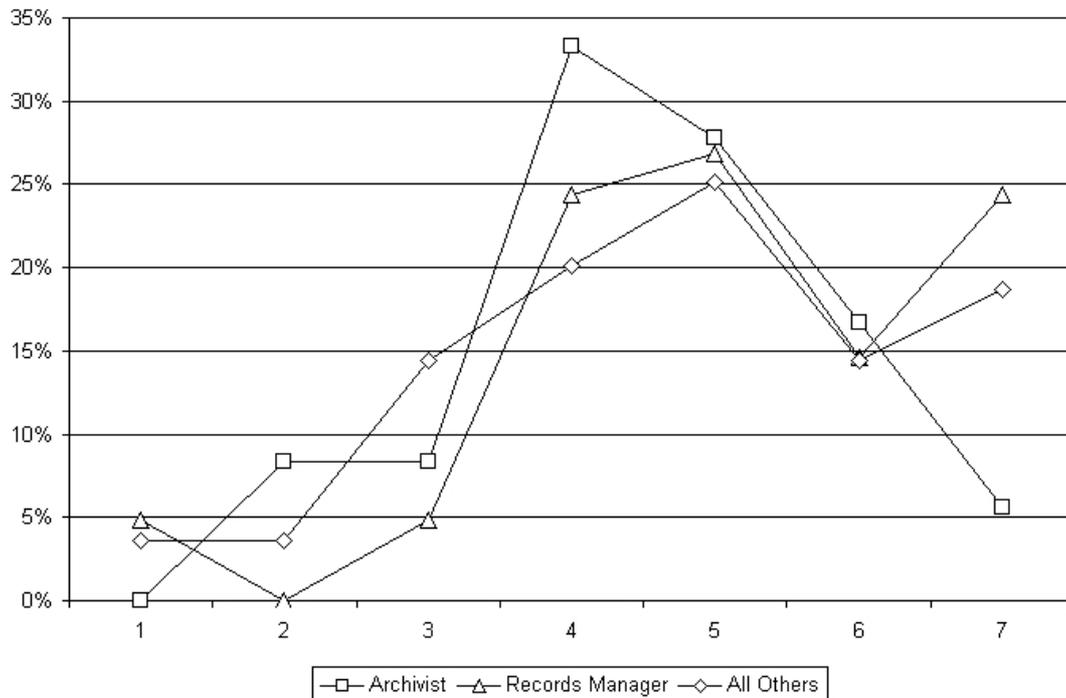


Figure 5.4: Profession and Filing Activities Home 2

Compared to the rest of the survey participants, the picture was much the same: archivists and records managers were only slightly more likely to report filing activities at home (see figure 5.4). This may indicate that the home behaviour of archivists and records managers has been informed by their profession; on the other hand, this may indicate that people who are more likely to perform home filing activities are more likely to choose to become archivists or records managers. In any event, the difference is so small as to be of doubtful significance.

### 5.1.3 Use Default Stores Home

All but 33 of the participants utilised at least 1 of the default storage locations provided by their computer's operating systems (e.g., 'My Documents'), as compared to 28 participants who utilised all of the default storage locations (see figure 5.5). This statistic may be a bit misleading, as there are certainly users who do not store a particular type of digital object<sup>4</sup> (indeed, this was one of the most common comments in response to this question) or do not store *any* media upon their computer system (a comment made by several participants). We

<sup>3</sup>'Technologists' includes such professions as 'computer programmer', 'software tester', and the like.

<sup>4</sup>E.g., "pictures," "video," "email," etc..

may conclude, however, that the majority of computer users do not go out of their way to modify the built-in behaviour of their systems when it comes to organising their files<sup>5</sup>.

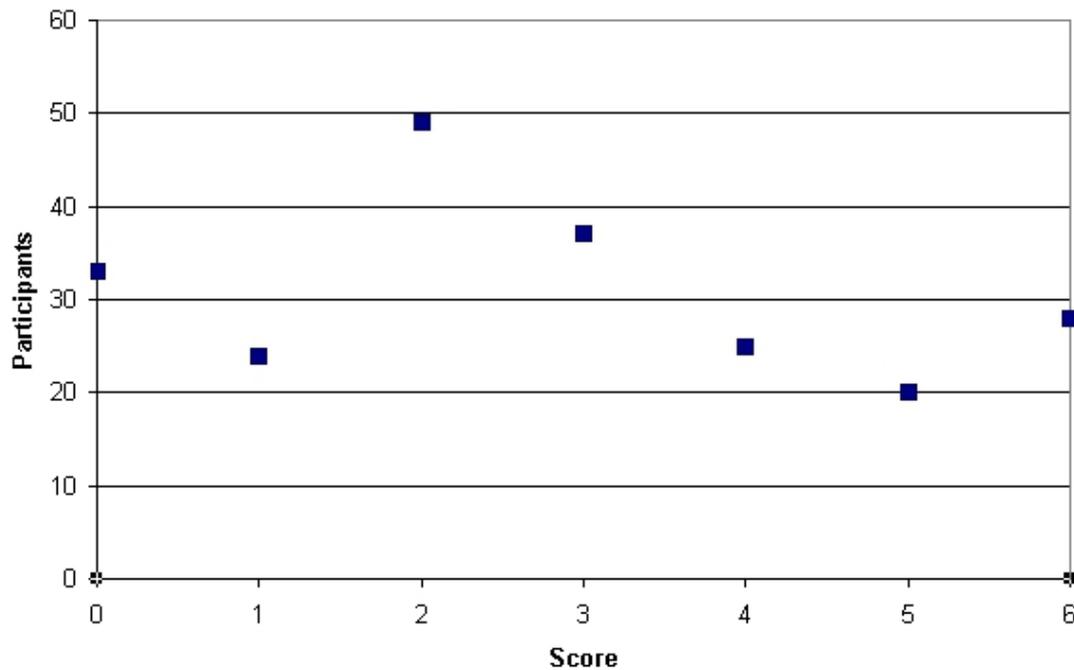


Figure 5.5: Use Default Stores Home

Considering the usage of default filing locations in light of the previous question concerning whether users participate in filing activities upon their computer systems<sup>6</sup>, we may have cause to wonder what, exactly, the users consider to be valuable features of organisational system; do they simply accept what is given to them (most of the time) and work within those confines? Or are users simply accepting the limitations of their computer systems, then periodically taking the time to reorganise the information kept in their default stores into their preferred organisational structure?

Profession does not appear to have a significant correlation with the use of ‘default stores’ on the home computer (see table 5.3). Again, while the percentages within each profession may seem to indicate a difference, the numbers present within the ‘Archivist’ and ‘Records Manager’ profession are small enough that the differences are not statistically significant: 3 or 4 different responses would be enough to alter the percentages enough to cause those two professions to come into alignment with the rest of the population. In addition, the fact that not all users store every type of content upon their home computers must be taken into account: it is certainly

<sup>5</sup>There exist for the various operating systems a number of tools for manipulating the default behaviour of the computer file system; e.g., the tool TweakUI is provided by Microsoft for the purpose of manipulating such behaviour.

<sup>6</sup>See section 5.1.2 on page 105.

Score	All Others	Archivist	Records Manager
0	17%	6%	20%
1	9%	19%	12%
2	22%	19%	27%
3	15%	17%	24%
4	16%	06%	2%
5	7%	19%	7%
6	14%	14%	7%

Table 5.3: ‘Profession’ and ‘Use Default Stores Home’

possible that the apparent variation may be attributed to differences in content stored upon home computers rather than attributing these differences to the profession of the participants.

#### 5.1.4 Clean-up Home

Participants reported that, for the most part, they “cleaned up” their various data stores (email, contacts, documents, social network statuses, text messages, and mobile phone contacts) several times per year (See figure 5.6). Depending upon the data store, this was found to be more or less frequent, with email being “cleaned up” most frequently (30% performing this task daily, another 25% performing this task weekly).

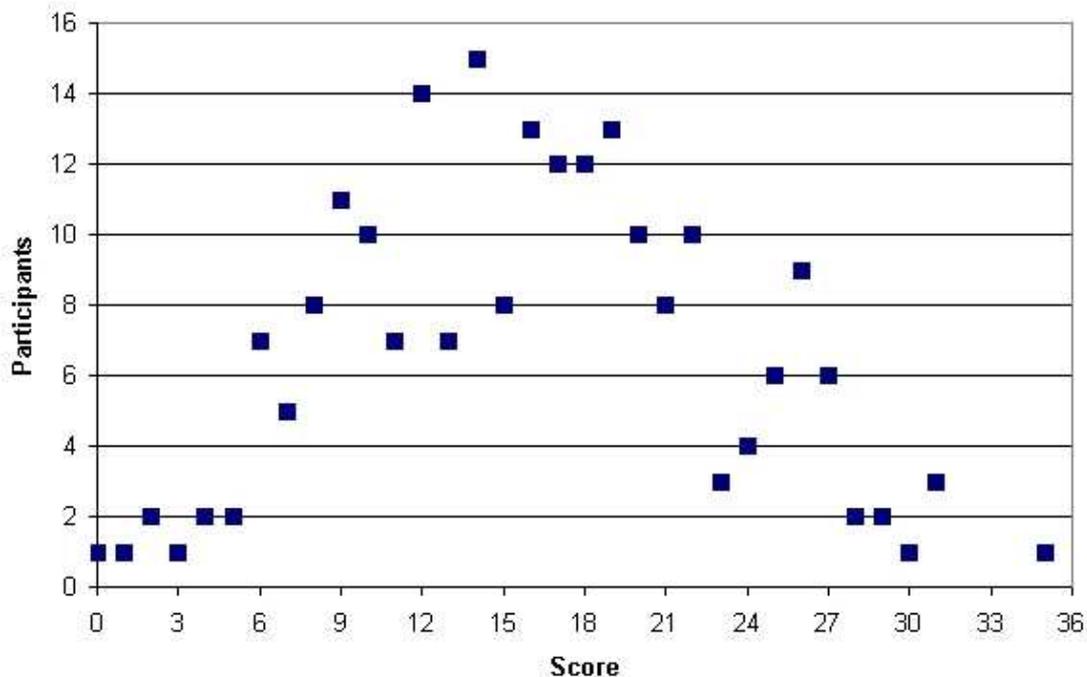


Figure 5.6: Clean-up Home

In contrast to email, documents were found to be very infrequently revisited for the purposes of organisation, retention, or deletion (see table 5.4).

Frequency	Participants	Percent of Total
Never (or Not Applicable)	8	3.70
Hardly ever	35	16.20
Every year	31	14.35
Several times a year	85	39.35
Once a month	33	15.28
Once a week	15	6.94
Every day	9	4.17

Table 5.4: Frequency of “clean-up” of user’s documents

Age, Gender, Home Country, Profession, and other demographic variables were found to make little difference in behaviour with regards to the frequency of cleaning up documents at home. This would seem to indicate that the behaviour is fairly consistent amongst the participants: cleaning up documents takes place periodically for all involved in this study, with email being most frequently cleaned up and documents kept in folders being relatively infrequently addressed.

### 5.1.5 Deletion Affordances Home

The study enquired as to which of 9 properties presented by operating systems were found to be helpful in selecting files for deletion<sup>7</sup>. The majority of respondents made use of only 2 or 3 of these features in their deletion decisions (see figure 5.7). This is potentially worrisome because such a finding might indicate that respondents were unaware of some of these features, unaware of how to access them, or were perhaps making extensive use of the item’s content rather than considering the metadata presented to them by their computer systems. It is unclear how to interpret this result, because of the structure of the study—analysing the data as individual variables does not (and would not) provide any greater insight as to whether individual respondents were aware of the metadata presented to them by their operating systems.

When we include the item’s content in the evaluation of the affordances utilised in a deletion decision, it becomes clear that certain users make almost exclusive use of the item’s content to the exclusion of all other properties of a digital object (see figure 5.8). 176 of 216, or 81% report that they use item’s content in their deletion decision. More revealing, however, is that 28 participants (13%) claim to make use of no other feature of the object in their decision to delete.

In the case of some participants, perhaps the file name serves to inform them as to the content, as 105 of 216 (49%) attest to using the file name in their decisions to delete; or perhaps

<sup>7</sup>‘File Content’ has not been considered in this category.

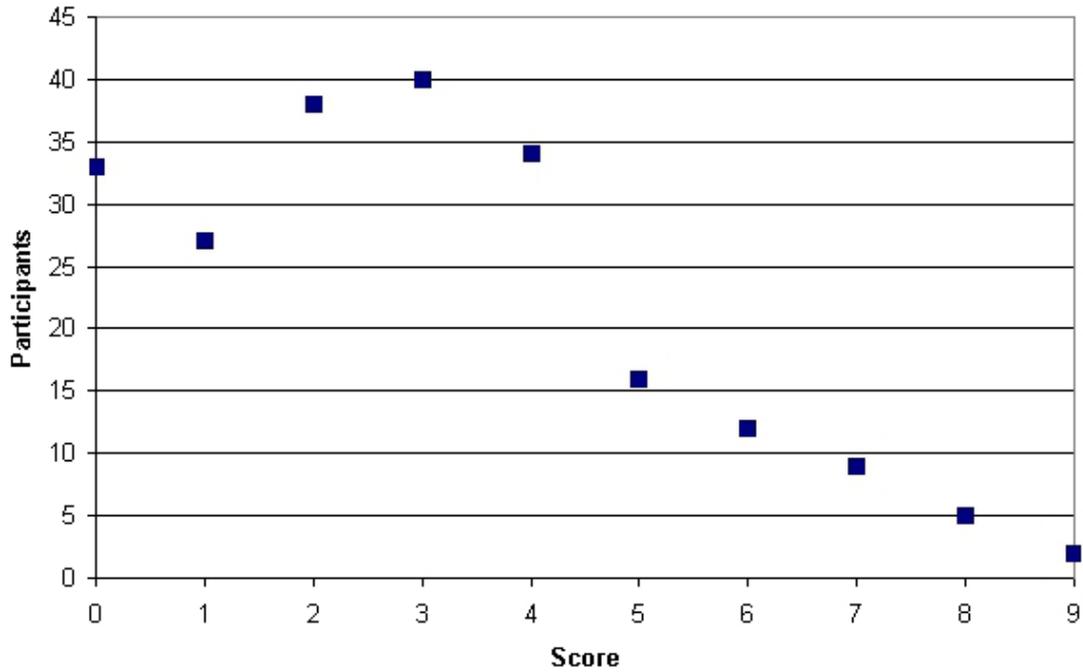


Figure 5.7: Deletion Affordances Home

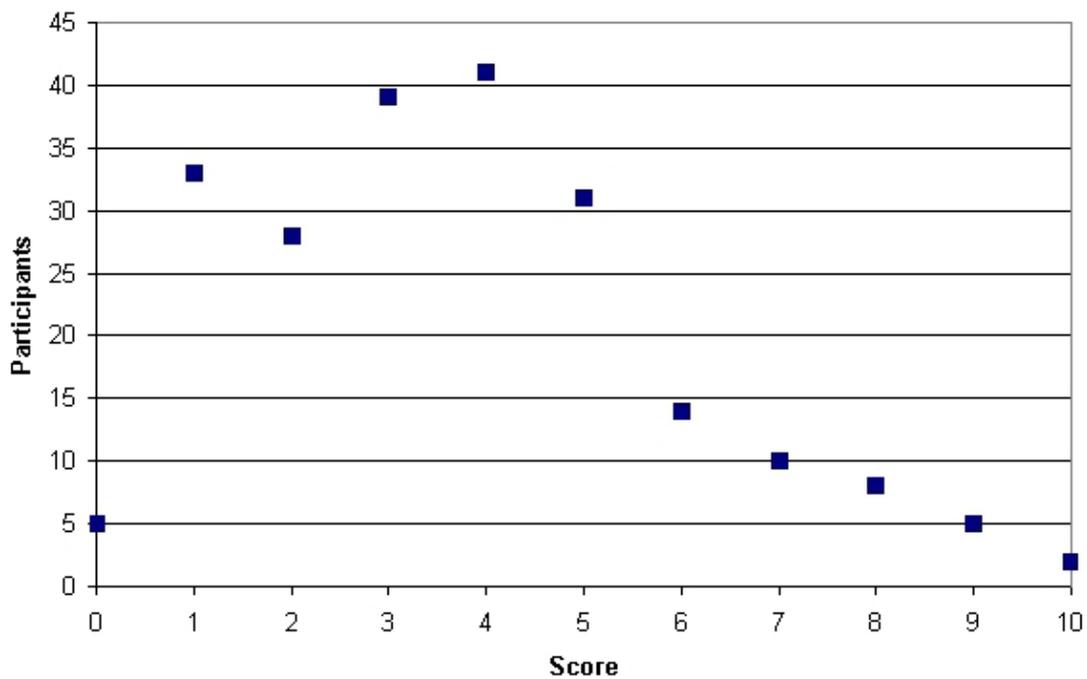


Figure 5.8: Deletion Affordances Home 2

the file location performs this service (as it is used by 73 of 216, or 34%). If we interpret the responses in a narrow manner, then the survey participants are stating that they have direct access to the content and that guides their decision. In a wide interpretation, the participants would seem to be using the metadata presented by the operating system to inform their decision,

in that the metadata acts as *aide-mémoires* as to the content.

Of the responses, 40 (19%) stated that they did not use the file content as the basis for their decision to delete. This would seem to indicate that, for nearly 1 in 5 participants, the content of the file was not used in their decision-making process. From this, we may wish to conclude that the narrow interpretation, above, is best: when participants state that they use the content of the object as the basis for their decision, they are (at least with regards to some decisions) indicating that they are *not* utilising any metadata in addition to that file's content.

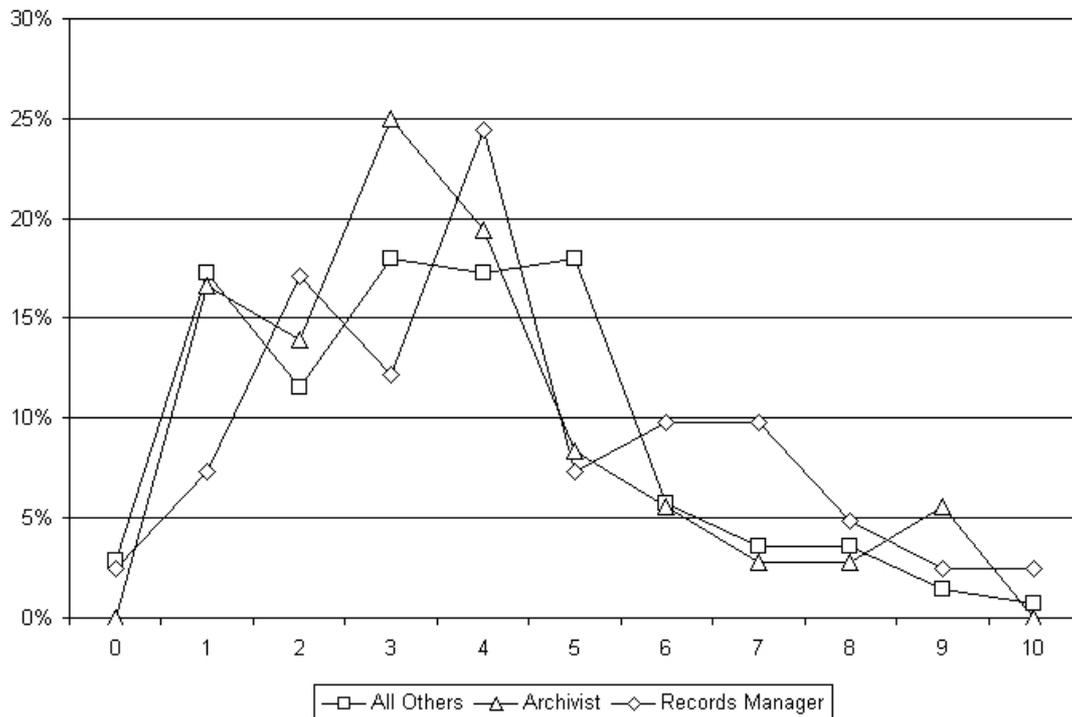


Figure 5.9: Profession and Deletion Affordances Home 2

When examining the deletion affordances used by various professions, Archivists and Records Managers stand out slightly in that they may actually be less likely to use available information in making deletion decisions (see figure 5.9). This may be due to these two populations having stricter naming and filing conventions in place and, thus, do not *need* to make use of other properties of documents in their decision. It is unclear, however, whether this is the case. The difference is only slight and the two professions appear to follow the same trend as the rest of the study participants, with the majority using between 2 and 5 properties of files to help them determine whether to delete.

### 5.1.6 Deletion Affordances Work

This study also sought to be about to consider possible differences between decision-making processes within the different contexts of home and work. In order to do so, the same question was asked with reference to deletion decisions.

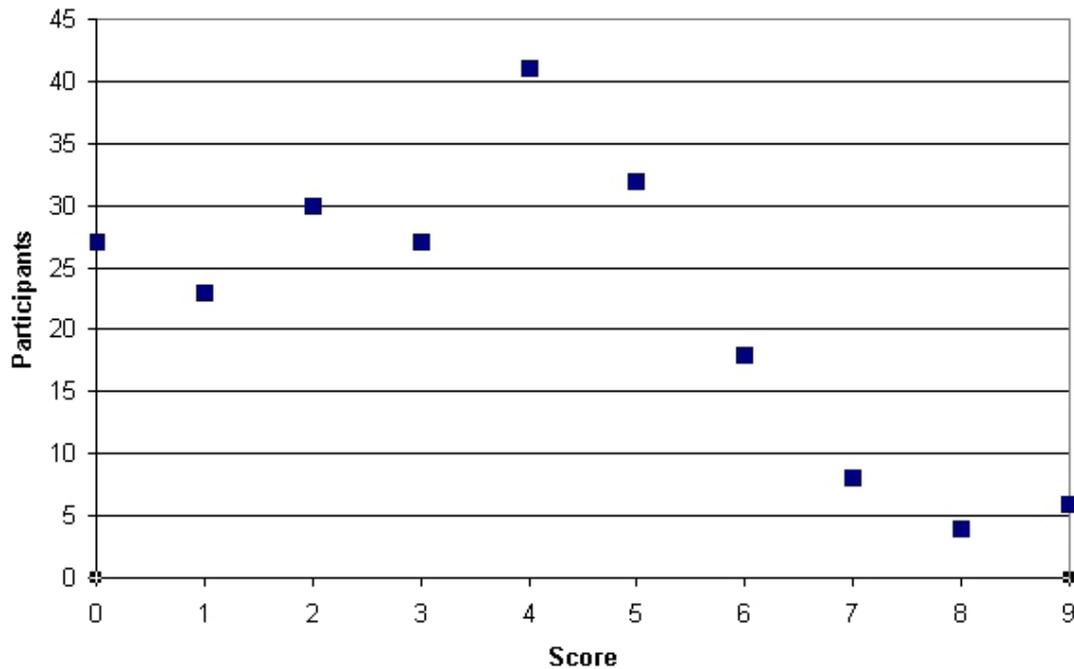


Figure 5.10: Deletion Affordances Work

Graphically, the results appear very similar, which would seem to indicate that there is not much difference between the two contexts—that the participants mostly utilise the same affordances at work as they use at home (see figure 5.10). There is some small measure of difference, however, in that at least a few the respondents claim to make more use of these affordances at work than they do at home.

Affordances	Work	Home
0	27	33
1	23	27
2	30	38
3	27	40
4	41	34
5	32	16
6	18	12
7	8	9
8	4	5
9	6	2

Table 5.5: Comparison of ‘Deletion Affordances Work’ and ‘Deletion Affordances Home’

Some of the difference between home and work affordances may be attributable to a difference in computer usage between the two contexts; indeed, this seems likely, in particular because some respondents claimed that their home computer usage was solely for the purposes of entertainment and the like and, therefore, these participants would not engage with the same type of content in the different contexts. Because the variation between the two contexts was quite small (see table 5.5), we may infer that participants tend to relate to their digital objects in much the same way between the two contexts. That is not to say that their *reasons* for making decisions differ between contexts, but that their manner of relating to their computing systems does not vary much between the two contexts.

In general, participants claimed to use a greater variety of metadata in their work deletion decisions as compared to their home deletion decisions. This could indicate that deletion decisions are more carefully considered within the work environment. It could also indicate that people are, in general, more aware of the content of their home computers and, therefore, do not need to consider as many aspects of files about which they are making this sort of decision.

### 5.1.7 Computer Skills Training

This study asked the participants about what computer skills training was provided formally, either within the organisation as part of a training course, or outside of work as part of an organised (not self-led) study course. 203 (94%) of the respondents indicated that they had received some sort of formal training with regards to their computer skills (including records management training provided by the organisation), and half (107 of 216) had received 3 or more types of formal training (see figure 5.11).

In addition to formal training, the study asked participants about any computer skills training they may have pursued on their own, in addition to that provided formally. 214 (99%) stated that they had pursued their own computer skills study over and above that provided by their organisation or provided in a classroom setting.

From these statistics, we may surmise that the participants in the study are fairly well-informed with regards to the types of metadata available to them from computing systems. Yet, in examining their behaviour with regards to how they make use of that metadata in their deletion decisions<sup>8</sup> we are confronted with the fact that at least some participants do not make use of this information. We must ask, then, whether it is not the skills of the participants which are lacking, but the quality or accessibility of the metadata itself which is not adequate to the

---

<sup>8</sup>See section 5.1.6 on page 114.

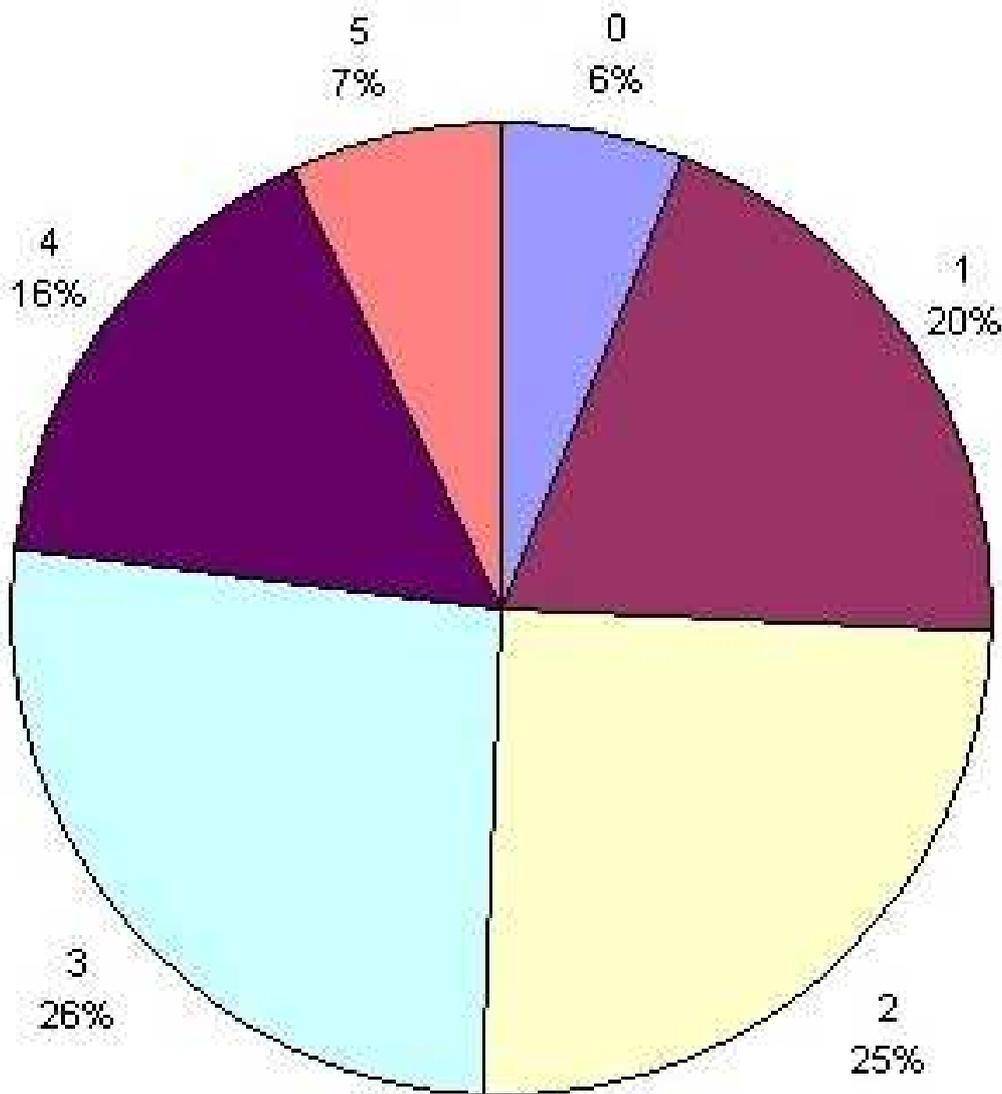


Figure 5.11: Training Organisation Pie

needs of the computer user.

### 5.1.8 Record Generating Activity

In an effort to understand the difference in computer usage between the contexts of home and work, participants were asked to provide a percentage of time spent on various activities. They were allowed to exceed 100%, with the recognition that certain activities might overlap (e.g., ‘composing email’ might also involve ‘documenting personal finances’).

These questions were selected for consideration at this point because they allow for the comparison of three discrete activities between the contexts of home and work. About half of the participants (114 of 216, or 53%) believed that 25% or less of their home computing time

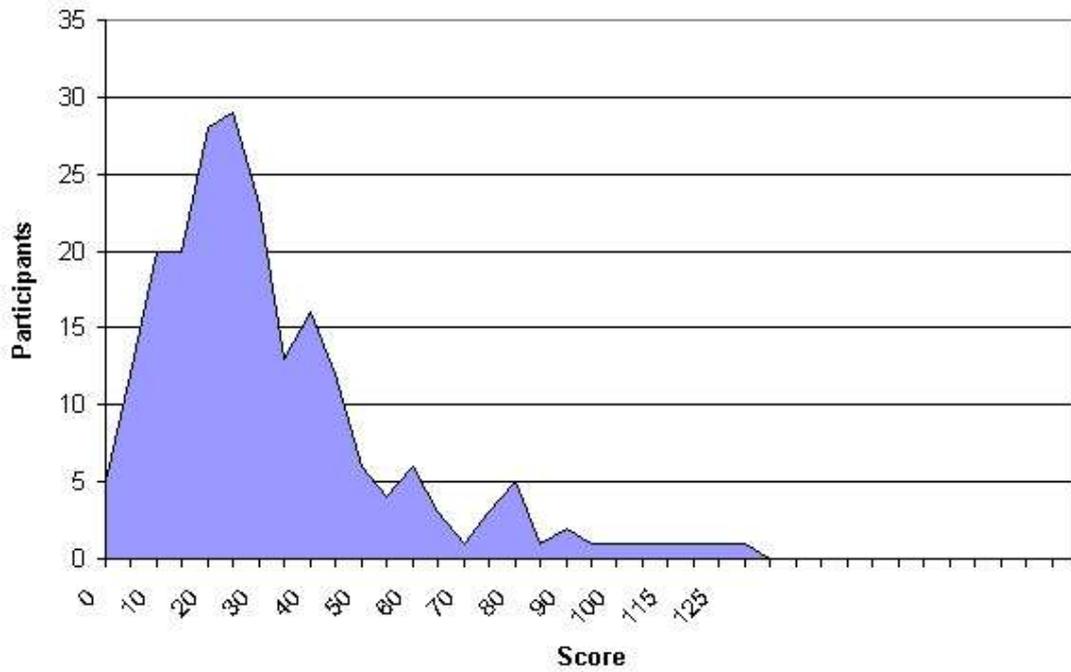


Figure 5.12: Record Generating Activity Home

was spent on activities which might generate content of this type (see figure 5.12).

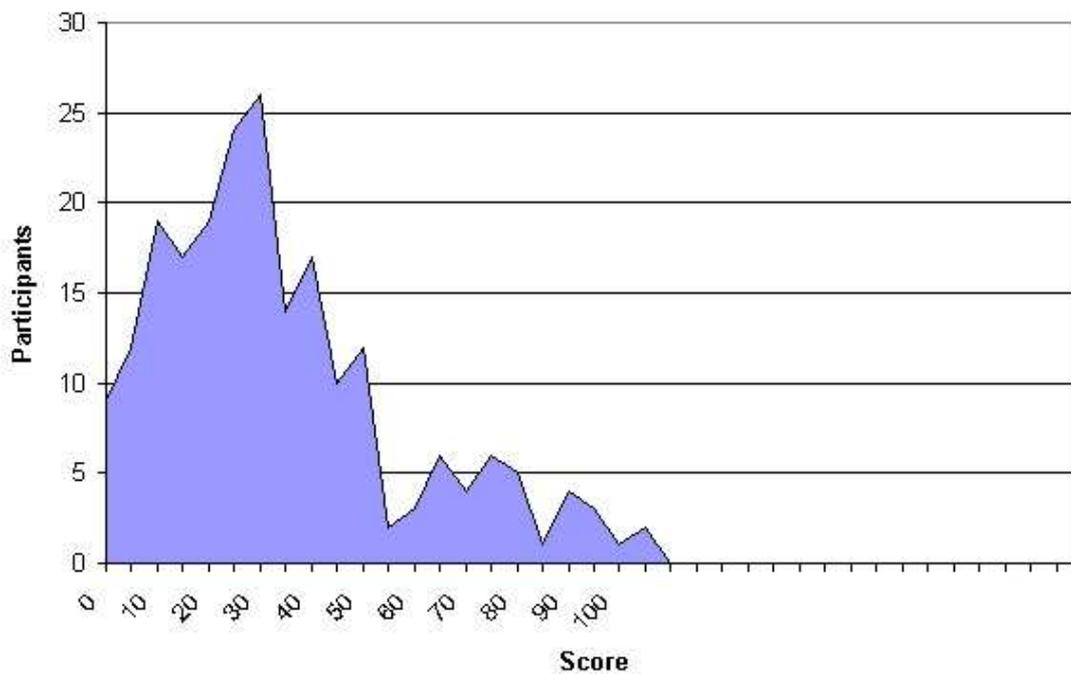


Figure 5.13: Record Generating Activity Work Area

Within the context of work, these measures yielded nearly the same percentage of responses (100 of 216, or 46%) as within the context of home (see figure 5.13). This provides an indicator that the activities performed between the two contexts are relatively similar, albeit for different

purposes at least in some degree. That the measures overlapped somewhat is also significant: part of this measure was comprised of time spent engaged in documenting ‘personal finance’.

That there was little change in the measures is significant, as 97 participants (45%) spent at least 5% of their work computing time in recording personal financial information. This would seem to indicate that not only did participants use their home and work computing systems in very similar ways, but that they use them to accomplish the same *types* of task, if not the same *exact* tasks. This blending of home and work contexts is possibly significant in terms of explaining the similarities between behaviours, thoughts, and feelings: if participants are relating to the same information in both contexts, they may be less likely to relate to their computers differently in each context.

### 5.1.9 Reasons Given for Deletion and Preservation

Up until this point we have examined the ways in which participants stated that they use their computer systems. We have not considered, however, the *reasons* behind their behaviour. These reasons are important not only because they may differ between the contexts of home and work, but because they may reveal some interesting features about the ways in which people approach their decisions to delete or preserve digital objects. Also, one of the goals of this study was to consider whether and how people consider others in their decisions to delete or preserve digital objects.

#### Functional Reasons

One category of question asked in this study was regarding functional reasons for deleting digital objects. These included such reasons as ‘it gets in the way’, or ‘it takes up space’. Most participants (120 of 216, or 56%) accepted one or two of these as reasons they might give for deleting something at home, with 32 (15%) accepting none of these as reasons they would give for deleting something at home (see figure 5.14).

These statistics change very little when considering the same respondents within the context of work: 121 of 216 (56%) accepted one or two functional reasons for deletion at work, and 39 (18%) accepted none of these as reasons for deleting something at work (see figure 5.15).

This indicates that, at least for functional reasons given for deleting digital objects, participants tend to rationalise their decisions in much the same way, regardless of context. This is important as it would seem to indicate that not only do computer users relate to their digital objects in the same ways between the two contexts, but that they have very similar mental processes when considering personal and work objects.

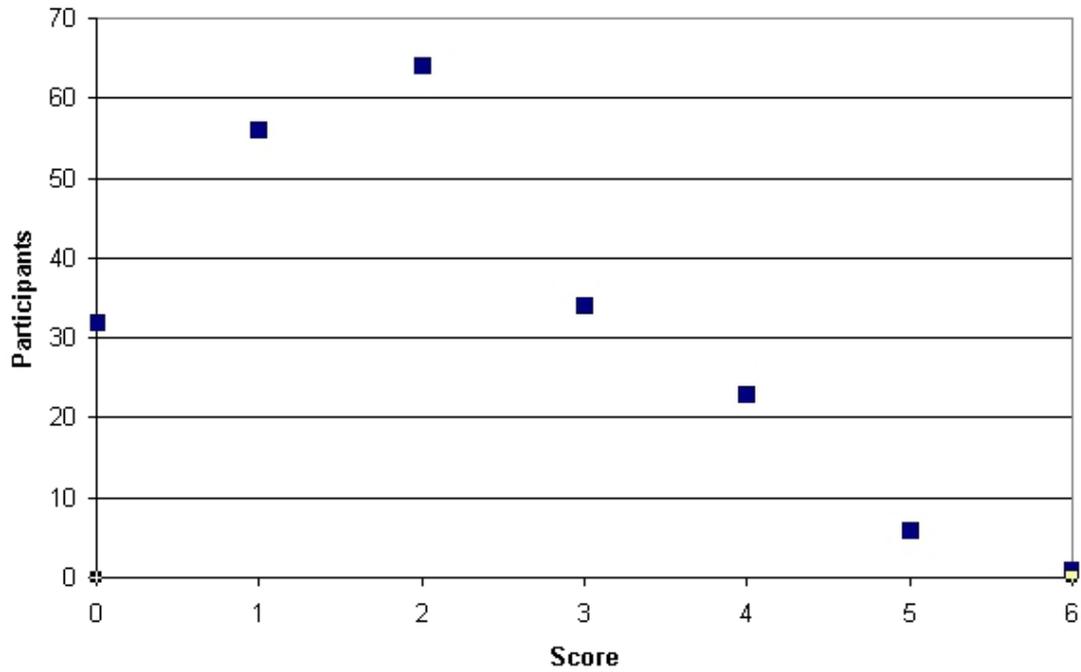


Figure 5.14: Functional Reasons Home

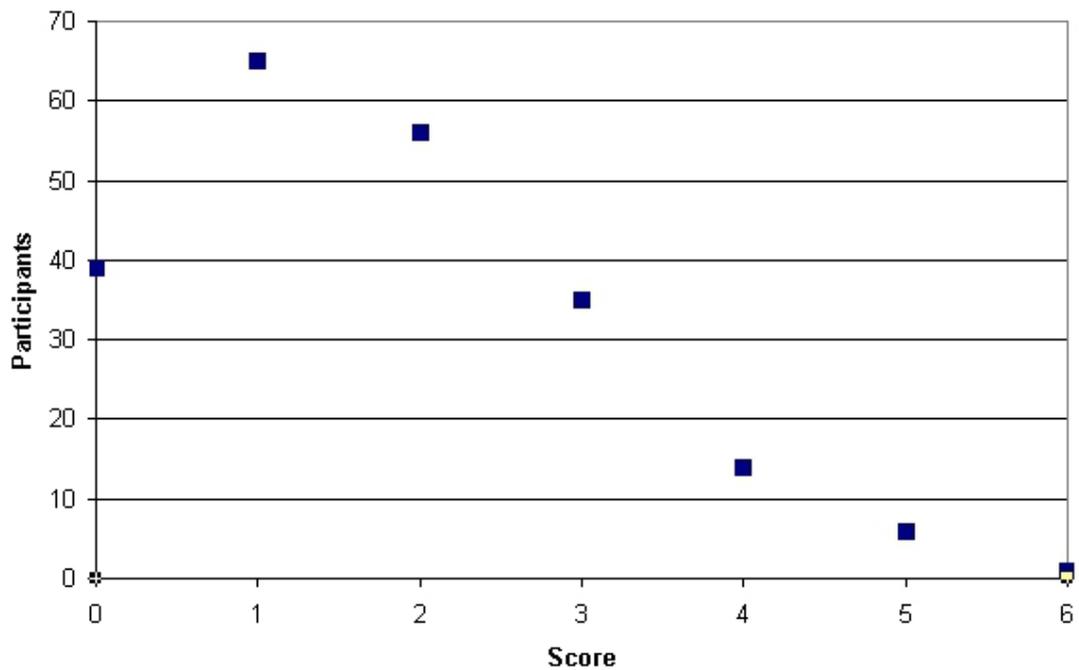


Figure 5.15: Functional Reasons Work

A very small percentage (7 participants, 3%) accept no functional reasons for deletion when at work, yet do so at home. This supports the conclusion that computer users do not approach the two contexts differently in terms of functional reasons to destroy digital objects.

### Personal Reasons

Another area to be considered is whether participants cited emotional reasons either for deletion or preservation. This includes the reason ‘somebody spent time on it’ because that feeling could not rightly be considered to be ‘considering others’, as opposed to saving a digital object for someone else (e.g., ‘somebody else might need it’) which could. It is, indeed, a fine distinction, but one which may be valid, as ‘somebody spent time on it’ indicates that the participant *should* value it for some reason other than the intrinsic value of the digital object; ‘somebody spent time on it’ does not indicate that the participant is keeping the digital object *for* the other person.

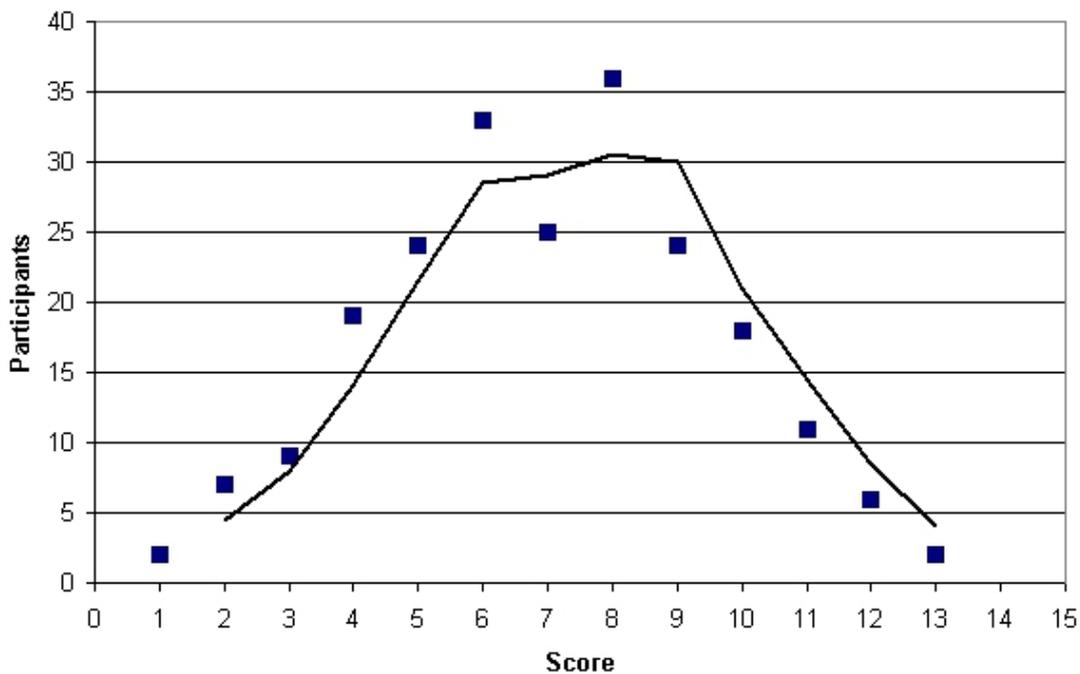


Figure 5.16: Personal Reasons Home Moving Average

As is to be expected, participants largely accepted personal reasons as reasons they would give for keeping or destroying digital objects at home. Surprisingly, however, participants also accepted personal reasons as valid reasons within the work context, with only 1 participant not accepting any personal reason as valid within the work context. 123 participants (57%) accepted 6 or more personal reasons as valid within the work context (see figure 5.16).

Shown with moving averages, figure 5.16 and figure 5.17 appear nearly identical, suggesting that participants make very little distinction between the contexts when it comes to personal reasons for deletion or preservation. This is surprising because the work context would seem to involve more decisions which would have an impact upon others and, therefore, would seem to argue against personal reasons being perceived as valid within that context; it would seem,

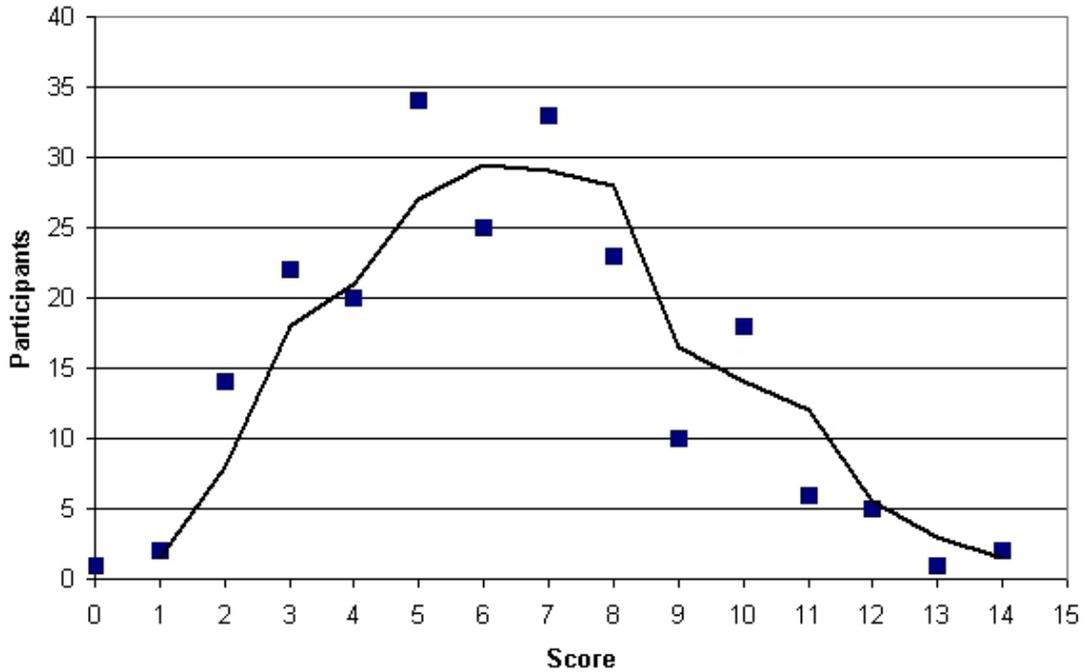


Figure 5.17: Personal Reasons Work Moving Average

then, that fewer personal reasons would be cited, or at least the frequency with which those reasons are cited as valid would seem to decrease. Unfortunately for this study, participants were not asked to evaluate the validity or frequency with which these reasons were applied<sup>9</sup>, nor were there any extant studies investigating these types of decisions between the two contexts. This is certainly an area requiring further study.

### Reasons Involving Others

Considering that some participants were single (40, or 19%), others may have only had young children, and still others may not have shared computing resources with the other members of their household (103, or 48% did share computers with others in their household), comparing ‘reasons involving others’ between the two domains may not be a fair comparison. However, one of the aims of this study was to consider *whether* computer users considered others in their deletion and preservation decisions; with this aim in mind, we can understand that it is significant that the majority of participants (171, or 79%) considered others in their deletion and preservation decisions at home (see figure 5.18).

The picture is largely the same when considering the context of work (see figure 5.19), although the percentages have certainly shifted. This is probably due to the already mentioned

<sup>9</sup>See section 3.7.5 on page 88.

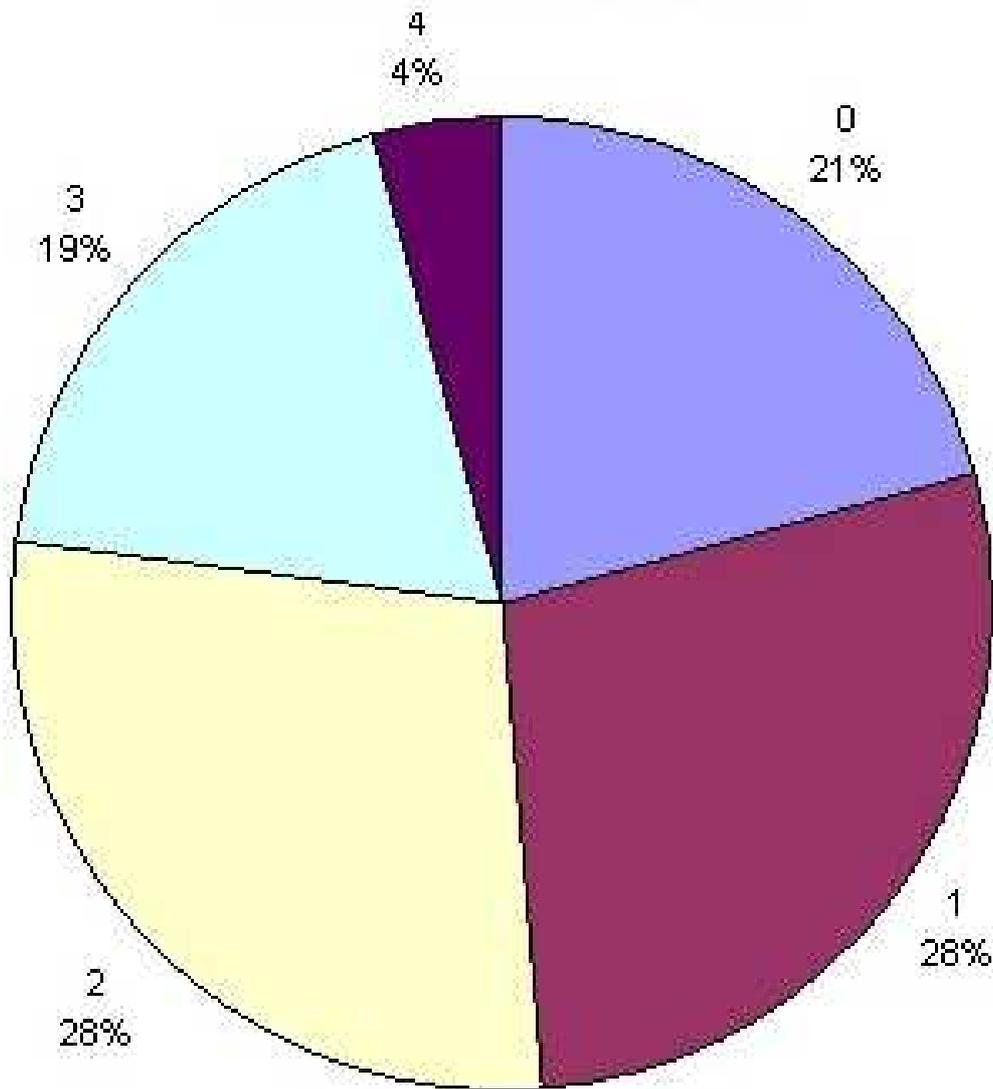


Figure 5.18: Reasons Involving Others Home Pie

fact that participants do not uniformly belong to family groups with multiple computer users—or, indeed, with multiple members.

It is an important finding that participants do, indeed, consider others when making their deletion and retention decisions, as this finding serves to balance the perspective a bit; while the participants in this study may perform the same types of record-generating activities<sup>10</sup>, and may make decisions for similar functional and personal reasons, they are also aware of the broader social context within which they are making these decisions. That is not to say that participants would agree that they were making a *moral judgement* in these cases (that would certainly be too strong an assertion), but that they take into consideration that others might have some interest in the digital objects under consideration for deletion or retention.

<sup>10</sup>See section 5.1.8 on page 116.

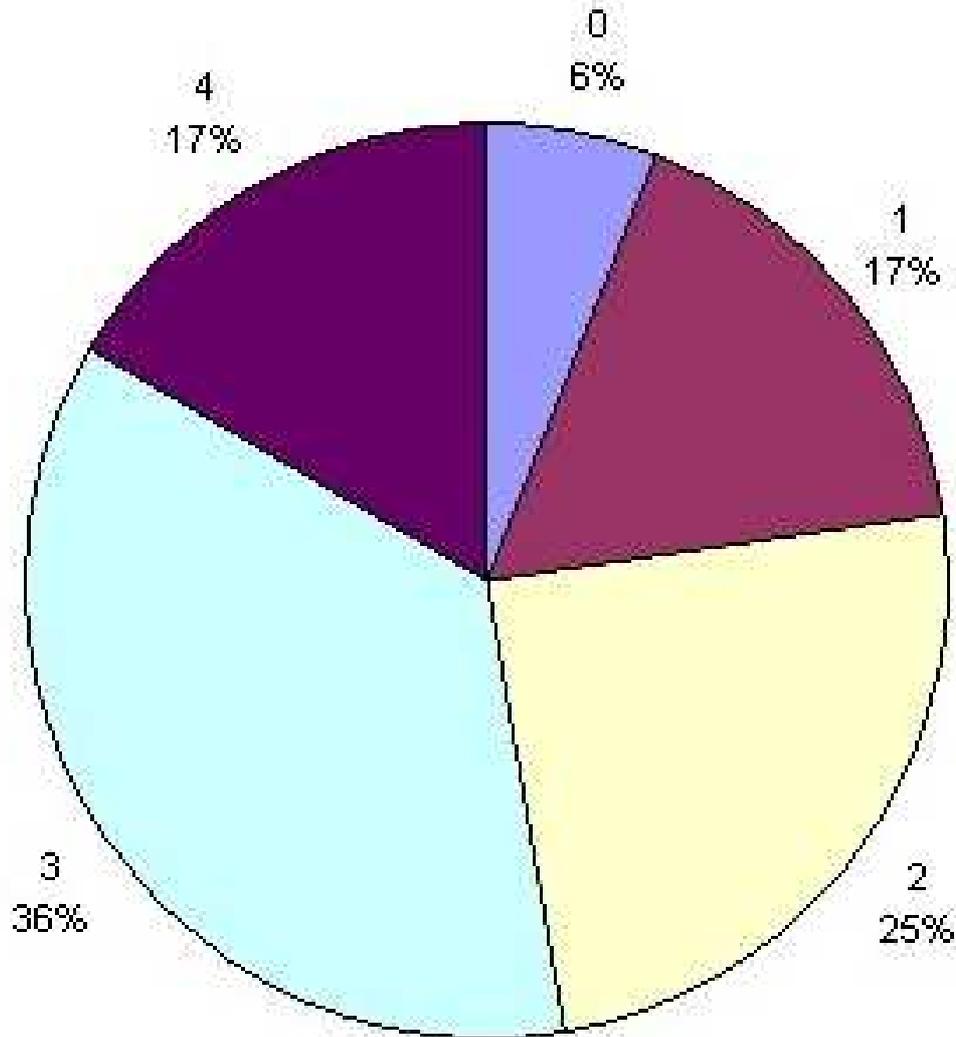


Figure 5.19: Reasons Involving Others Work Pie

### 5.1.10 Home / Work Mixing

It was thought that the degree to which the contexts of home and work have become intermixed might play some role in the behaviour of participants, in particular because the exchange of digital objects between contexts might affect how participants perceived each context. Not unexpectedly, the vast majority of participants engaged in some of these types of activity, whether it was keeping personal files at work, keeping work files at home, or engaging in work or personal tasks when in the other context.

However, 52 participants (24%) did not attest to any of this type of activity (see figure 5.20). It is unclear whether those 52 participants would have done so, given the opportunity, or whether they were opposed to such. What is clear, however, is that the other participants viewed the separation between home and work as porous, freely performing personal tasks at work and vice versa. It is impossible to know the effect that this has had upon the behaviours and perceptions

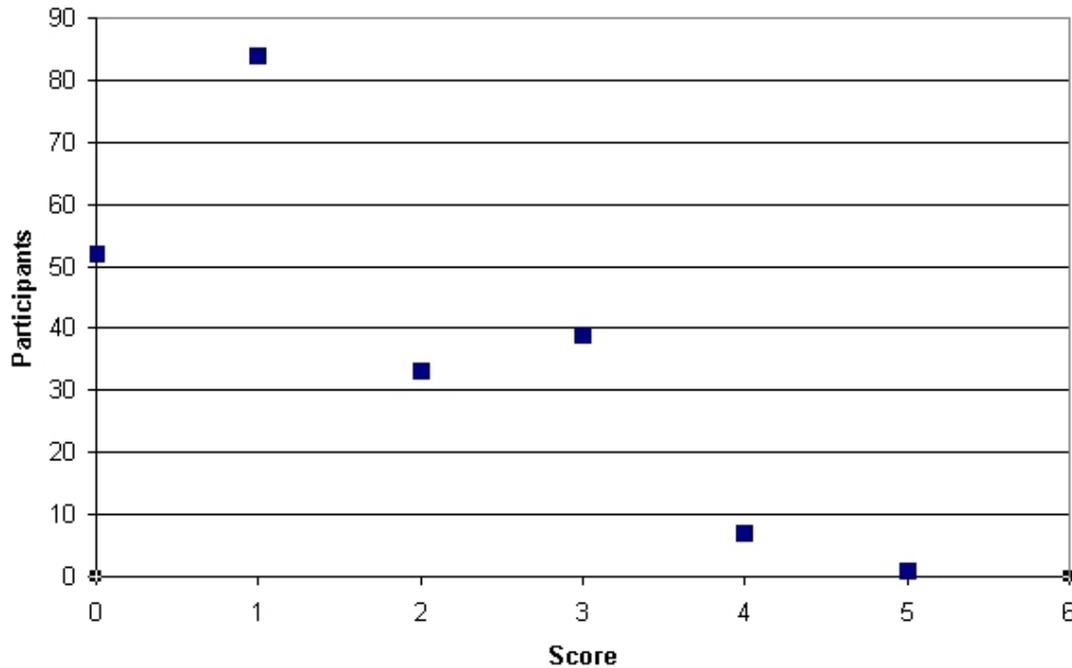


Figure 5.20: Home Work Mixing

of the participants, but it is not unreasonable to suspect that this free intermingling of the two contexts has contributed to the consistency in responses between the two contexts. The participants were asked about whether they kept digital objects on their work systems that they feel they own, personally, and how they would feel if their workplace were to examine these objects<sup>11</sup>. Their responses were fairly consistent in being opposed to the idea of their workplace computer being examined:

1. “I have some interests which my work doesn’t need to know about me such as political or religious views or random ramblings which might be considered racial or sexist by others and are none of their business but might be damaging to my career. I will not let anyone else examine anything on my computer, without a court order.”
2. “Even though I know it is on a work computer and that I do not own the computer, I am the only person to use the pc. Therefore, I feel no one else should have the right to use the pc unless I leave the company.”
3. “It feel that it is the content and not location of an item that determines ownership. Even if something is stored on a work computer, I don’t think it means I surrender it to the workplace—although I am sure management may disagree. I bet that would be their position if I were to place some of their items in my briefcase or laptop though.”
4. “I don’t draw a hard line between work and home.”

---

<sup>11</sup>See page 167.

5. “All personal files I access at work come from an encrypted file (I only access these files quickly and then close them). I feel my behavior limits my exposure to ‘big brother’ looking into my personal info.”
6. “I don’t keep anything on my work computer that I feel could be damaging to me. However, just as it would be inappropriate for my workplace to search through my desk without first asking me permission to do so, I also feel that it would be inappropriate for them to examine my computer. Though the computer (and desk) may belong to my workplace, I would feel that to examine these things, without my approval, would be disrespectful.”
7. “I use the university email address as a prime reliable address and so receive invoices from Amazon, for example. The employer owns the equipment and has the right to examine, but I would feel I worked in an environment of distrust if it were to happen, and that would upset me.”
8. “I keep my work separate from my home computer for privacy.”

In general, the majority of comments could be said to fall either into *a*) feeling that their workplace computer should be free from what they regarded as unwarranted surveillance, *b*) acknowledging that their organisation had the right to view their personal information but would feel violated if the organisation did so, or *c*) actively drawing a boundary between the two contexts, only keeping non-sensitive personal items on their work system. These comments indicate that although participants may engage in a degree of mingling of workplace digital objects on their personal computers, they are aware that personal files may be examined on the workplace computer and take measures to limit potentially embarrassing or damaging exposure. Intellectually at least, participants are aware that there are restrictions upon their computer use in the workplace, even though they may do give different reasons for their deletion or preservation decisions between the two contexts.

### 5.1.11 Work Records-Management Environment

Also thought to be a contributing factor towards the behaviour of individuals within organisations, the presence or absence of a formal records-management system was measured. Nearly all (193, or 89%) of respondents either worked with a records-management system or had received formal training in records management, as provided by their organisation (see figure 5.21). It is difficult to judge just what effect this has had upon behaviour, however, particularly in light of the fact that responses do not vary between contexts—it was expected that the presence of a records-management system and/or records-management training would affect a change in behaviour whilst participants were acting within the work context.

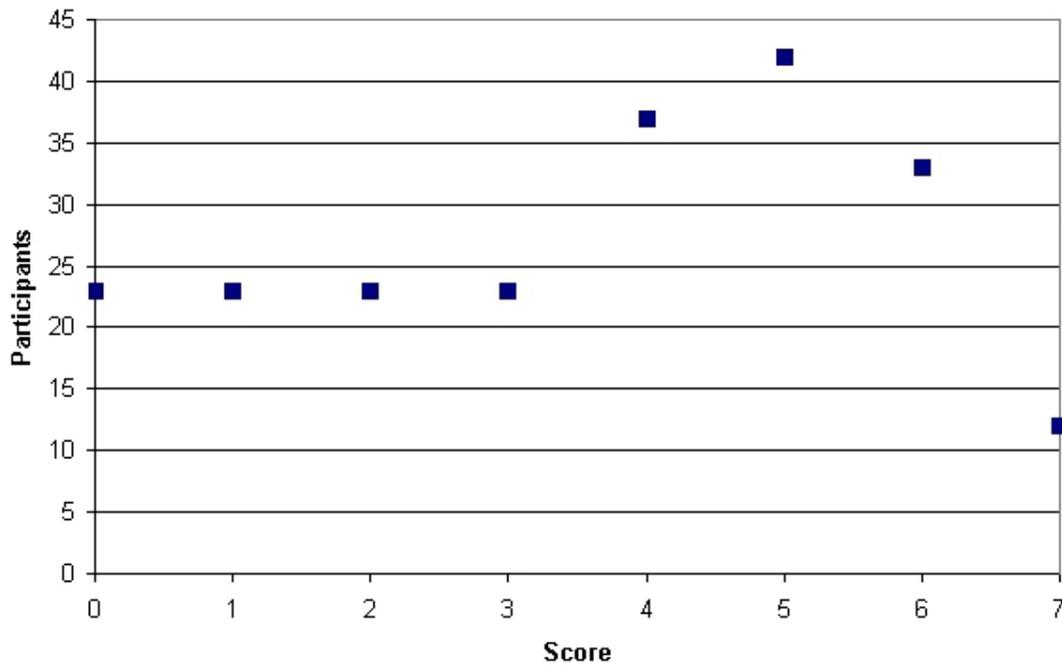


Figure 5.21: Work RM System

70% of participants (151) work in organisations with a formal records-management system, 13% (28) were uncertain as to whether their organisation had such, and 17% (37) did not work in such an organisation. This individual variable was the most widely-selected of the 7 variables comprising this measure; although it was commonly agreed that their organisation had a records-management system, however, only 8% (18 participants) agreed that their organisation managed documents centrally. This apparent contradiction could indicate that some organisations managed their own documents within their organisational unit, it could indicate that only certain types of content were maintained in their records-management system, or it could indicate that participants were uncertain about what was contained within their organisation's records-management system.

That the majority of participants had at least been exposed to formal records-management concepts would seem to indicate that participants *ought* to relate to formal records (or documents resembling records) in a particular way. Indeed, this was found to be true for 162 (75%) of the participants, who admitted to none of the three practises which comprise the 'RM Violation' measure (see figure 5.22).

The remaining 25% of participants gave a variety of reasons why violating their organisation's records-management policy wasn't a bad practise, with the most-cited reason being one of convenience or ease-of-access to documents also kept within the document management system. Some participants stated that they were concerned about the integrity of the document

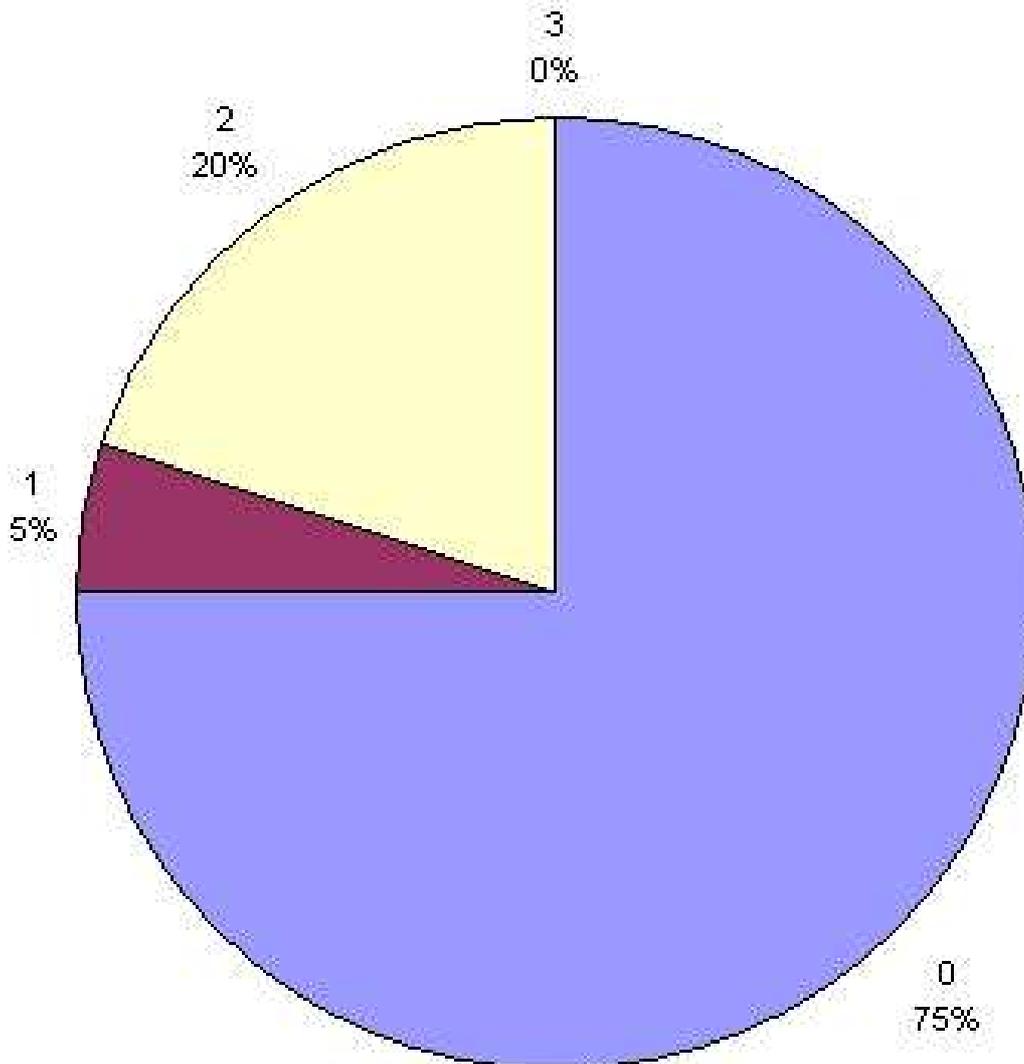


Figure 5.22: RM Violation

management system, so kept a back-up copy. Some participants also expressed concerns about proving their actions and, therefore, kept copies of documents which had been transferred to the document management system. Others kept copies of documents because they wished to memorialise a project, or because they felt emotionally attached to the content.

Generally speaking, participants who kept copies of documents which were also kept in a document management system were aware that their actions were in contravention of their organisation's policies. Their reasons for behaving in this manner were varied, but can be generally characterised as involving feelings of information ownership: they felt that the documents were important to them, and that the loss of those documents to a document management system would in some way affect them negatively.

Of the 162 participants, 121 (75%) provided commentary concerning these types of action; although only 25% admitted to these actions, the commentary provided indicated a high level

of concern regarding centrally managed documents, the keeping of copies of documents on a home computer, or keeping documents after leaving their present organisation. Additionally, the comments very often conflict with the response given. For example, one participant does not claim to keep copies of centrally managed documents, yet comments that, “It’s easier to access them—just for the ones I use every day.” The majority of commentary may be fairly represented by the following quotations:

1. “In case there are errors or something gets lost. It’s good to keep your back covered.”
2. “Proof of documentation of actions taken during a particular timeframe re: a particular topic.”
3. “Keep all sent mail for documentation to prove that others were notified—helps recreate what others may have forgotten.”
4. “Don’t trust repository or don’t trust other editors.”
5. “It is easier to edit and access them on my local computer and then to update the central files after.”
6. “It’s easier to access on my dedicated computer than searching for it in shared docs every time I need it.”
7. “Reminder of work created.”
8. “Record and aide-memoire of actions, instructions and responsibilities.”

Participants’ concerns may be best characterised as *a)* proof of action, *b)* lack of trust in the repository or members of the organisation, *c)* ease of access or use, *d)* memorialisation of work, or *e)* some combination of these factors. The choice between ‘yes’ or ‘no’ was not sufficient to explain the behaviour of the participants, who largely view their activities as not in violation of their organisation’s records management policy although they are keeping certain content on their work or home computer and even after they have left their organisation.

A related question reveals a wide range of feelings regarding the possibility of being given a particular way to keep things organized, on their work computer<sup>12</sup>:

1. “I would LOVE it!”
2. “Relieved, would make filing decisions easier.”
3. “Elated.”
4. “We have a bit of that already. If done well, a godsend.”

---

<sup>12</sup>See page 176.

5. "I would have no problem with this. It might help me to be more organized. I'd also like to have administration authorize me to delete old documents, instead of making the call myself."
6. "I think I would have a hard time adjusting to it at first, since I tend to have my own ways of organizing my files that makes sense to me."
7. "I might find it frustrating, as not everyone likes the same method for organising things. Consistency is important though, so I think it would be a good thing."
8. "I think everyone has their own organization method, but I think suggestions on how to organize electronic items is useful."
9. "I wish things were more standardized (including naming conventions, filing conventions, etc.). It would make things easier and less time consuming for everyone."
10. "I wouldn't like it as I prefer to organise my docs/files my own way."
11. "I would prefer my own organizational system."
12. "I would probably resent the intrusion, since I already have a quite rational organizational scheme."
13. "I'd feel that it was an intrusion on my autonomy."
14. "Annoyed."

Participants were fairly well distributed between *a*) strongly in favour of being given an organisational system or *b*) strongly opposed to the idea. Although there weren't strong correlations between receptivity to organisational systems and other individual variables, it seems that feelings are fairly strong on one side or the other, suggesting that there is at least a modest desire for some sort of externally-imposed organisational system but that any such system must not be too inflexible in its requirements if it is to be adopted by a majority of those within an organisation.

Regards Work IP As Their Property	Quantity	Keeps Work IP	Quantity	Sample Probability
No	184	No	168	0.91
Yes	32	No	10	0.31
No	184	Yes	16	0.09
Yes	32	Yes	22	0.69

Table 5.6: Given 'Regards Work IP As Their Property', analysis of 'Keeps Work IP'

Most participants (184, or 85%) did not regard work intellectual property as also belonging to them (see table 5.6). 69% of those which did regard intellectual property as their property (32 participants, or 15%) said that they would keep that intellectual property if they were to

leave their current employer. Ownership of intellectual property, then, can be said to be a key factor in whether individuals will keep content which also belongs to their employer. A very small percentage (16 participants, or 9%) said that they would keep digital objects even though they felt no ownership of these digital objects. The comments provided by this small percentage can be characterised as falling into two categories: “it would be too much trouble to figure out what to get rid of” and “the content will be useful in my next job.” These 16 participants, however, represent a small minority of those who said that they would keep digital objects if they were to leave their present employer. Feelings of ownership of information may be regarded as predictive of whether an individual will retain digital objects when leaving an employer.

The following chapter provides an overall statement of study findings. The following chapter also provides summary conclusions (section 6.6 on page 141) which shall be provided to study participants as per their expectation in participating in the study (See section 3.4 on page 75.).

## Chapter 6

---

# Findings

---

This chapter addresses those questions set forth in both section 1.2 on page 6 and in section 3.2 on page 69. Additionally, it provides hypotheses for further research, stated as hypotheses within each section.

### 6.1 Organisation of Digital Objects Desired

This study suggests that people are generally invested in maintaining some sort of organisational system for their digital objects, regardless of the type of digital object. In general, respondents engage in “filing activities” fairly regularly, although these activities vary depending upon the media type<sup>1</sup>. Aside from *the location* of such organisational systems<sup>2</sup>, this study finds no strong correlation with such factors as age, home culture, or profession. It was initially believed that there would be some correlation between these factors and the desire to organise digital objects; finding otherwise suggests that this desire may possibly be generalised to: 1) all computer users, 2) advanced computer users, where ‘advanced’ is taken to mean a computer user with some threshold of involvement with computers sufficient to require an organisational strategy, or at least 3) users who use multiple computers. It is unclear to which population this desire for organisation may be generalised; however, it is certainly a desire held by nearly every participant in this study and, as such, may be generalised at least to users of multiple computers.

The qualities of a person’s digital organisational system become significant when considering such comparison errors<sup>3</sup> as compromise effects<sup>4</sup> or contrast effects<sup>5</sup>. While these phenomena

---

<sup>1</sup>See section 5.1.2 on page 105 for analysis of filing activities.

<sup>2</sup>See 6.2 on page 133 for findings regarding storage location.

<sup>3</sup>See section 2.11.2 on page 36.

<sup>4</sup>See section 2.11.3 on page 41.

<sup>5</sup>See section 2.11.3 on page 42.

may be found in both the presence and absence of a strong organisational system, they may be somewhat mitigated by the presence of some sort of organisational system simply because these effects depend upon their being present within the items available either a clear loser (resulting in a contrast effect) or there being present a ‘medium-sized cola’ (resulting in a compromise effect).

In this study, the 7 participants without any organisational system may actually have a slight (though dubious) advantage in escaping such cognitive errors in the area of comparison errors<sup>6</sup>: because they do not choose from the pool of feature-rich and feature-poor digital objects and attempt to make sense of them, they do not end up misfiling their digital objects<sup>7</sup>. They are, however, subject to both compromise and contrast effects should they engage in deletion activities.

One factor influencing the regularity of performing organisational tasks (where ‘organisational’ may be taken to mean ‘filing *or* deleting’) was found to be the type of digital object under consideration<sup>8</sup>. This is perhaps unsurprising, as certain types of digital objects (e.g., email) are viewed as more transitory than other types of digital objects (e.g., music files). The determining factor, however, was not the type of digital object, but the means of access to those objects: objects accessed via the interface of the operating system’s file system interface were managed relatively similarly in terms of frequency, whereas those items which were presented to the user via an interface which encouraged management<sup>9</sup> were organised far more frequently. This finding suggests that, if changes in behaviour are desired with regards to the frequency at which these actions are performed, changes should be made to the means of access to these types of digital object; interfaces which allow ease of interaction with the content of the digital object and with which users must interact regularly (e.g., email systems) are far more likely to result in organisational activities.

**Hypothesis for Further Research 1.** Software interface design strongly affects the type and frequency of destruction and retention activities.

**Hypothesis for Further Research 2.** Software interfaces which provide immediately-accessible means of organising digital objects may result in improved decisions with regards to destruction and retention activities.

---

<sup>6</sup>See section 2.11.2 on page 36.

<sup>7</sup>See also section 2.8 on page 30.

<sup>8</sup>See section 5.1.4 on page 110 for analysis of “clean-up” activities.

<sup>9</sup>This is to assume that, for example, email systems encourage management by presenting the user with an interface which presents the item’s content at the same time as presenting folders: participants in this study reported that they relied on the item’s content very heavily when dispositioning digital objects, it would seem that email software encourages the user to engage in filing activities. Also, because email is generally regarded as demanding that the recipient take some action, computer users may be more likely to file or delete items from the inbox rather than leaving them where they present a demand for action.

## 6.2 Operating Systems' Storage Locations Accepted

Despite the desire for organisation of their digital objects, most participants accept the behaviour of their computer system when it comes to the location in which to place digital objects. This finding seems to indicate that the desires of participants are not well supported by their computer systems: they either routinely work around the “suggestions” of the operating system, or they periodically reorganise the files kept in the default locations to better support their own organisational system<sup>10</sup>. This finding is significant not solely because of the inefficiencies in either working around the operating system or in periodically reorganising the default stores, nor in the possible errors in either process (both procedural and cognitive). This finding is significant because it would seem that altering the operating systems to better support the organisational desires of the computer user would be a rather trivial change to ask of operating system manufacturers (or I.T. personnel) which would have a major impact upon the quality of the data both preserved and deleted upon computer systems. This finding does not take into account catastrophic loss of digital objects kept in those default stores, some of which are difficult to locate absent expert knowledge, yet it would seem intuitively correct to believe that, should the operating system itself facilitate better organisation of digital objects—and in a manner which were clear to the computer user—catastrophic loss would be much less likely, if only because the computer user would be aware of the exact location *of* those digital objects.

If users are periodically reorganising their filing systems—essentially, leaving their desk messy until something prompts them to clean up—then we are almost guaranteed to find some interesting cognitive errors<sup>11</sup>. We may expect to find comparison errors as users attempt to fit individual files into their existing filing structure and perhaps not finding an exact match. For example, consider a user attempting to classify a particular musical album when their filing system contains the categories ‘rock’ and ‘classical’. Should this user be presented with an album which does not fit either category exactly but could possibly be categorised as both, that user will feel forced into making a decision. This in itself is not a problem; the problem arises when the user, having proceeded along with the existing filing scheme, encounters an album which exactly straddles both categories. At that point, the user is forced into an undesirable decision: they must either reorganise all of their files somehow (after all, there may be any number of other albums which similarly straddle both categories although not to the same degree), the user must duplicate the album into both categories, or they must make an arbitrary decision to place the album into one or the other category. This is one type of comparison error. Given,

---

<sup>10</sup>See section 5.1.3 on page 108 for a discussion of the use of “default stores” and possible errors to be found during periodic reorganisation *of* those default stores’ contents.

<sup>11</sup>See section 5.1.3 on page 108.

this type of comparison error is not unknown in the analogue world; however, we must ask why this issue persists in the digital world, as technology would seem to be able to address this issue.

Digital objects were found to be very infrequently revisited for the purposes of organisation, retention, or deletion<sup>12</sup>. This infrequency of disposition would seem to exacerbate two special types of comparison error: compromise and contrast effects. Comparison errors should be expected, in general, as users attempt to fit their digital objects into an existing file structure (if, indeed, this is how they perform their “clean up”). One compromise effect to be expected when considering users’ periodic disposition of their digital objects is the compromise of keeping digital objects which are no longer needed—this is the less extreme option, therefore more likely to be chosen when considering digital objects within a large store of digital objects. However, contrast effects may be found when considering disposition of digital objects as they compare to other objects within the same store: certain objects will be selected because they offer a greater contrast in features from the other objects within the store and will, therefore, be selected either for retention or destruction.

We may also expect to find comparison errors during the (re)organisation process due to certain digital objects being feature-rich and others being feature-poor. Consider that certain digital objects are recognisable due to descriptive file names (one of the primary means of determining the content of a digital file) whereas others are only recognisable due to their context, or due to some other feature presented by the computer operating system. Some digital objects will present rich metadata to the user while others will present very little. This disparity between features presented would result in certain digital objects being selected more frequently for both a filing activity *and* a deletion activity; also, certain digital objects would likely be selected for filing in multiple locations, if that were an option perceived as available to the user. This comparison error differs from that of the contrast effect because the contrast effect is concerned with selecting based upon a comparison against a poor option; the generalised form of comparison error is based upon the accessibility of information, rather than the contrast between a poor option and a good option.

**Hypothesis for Further Research 3.** A change in operating-system presentation of digital objects may improve decisions made regarding their deletion or preservation.

**Hypothesis for Further Research 4.** When users understand how and where things are stored, they may be less subject to catastrophic data loss.

---

<sup>12</sup>See section 5.1.4 on page 110.

**Hypothesis for Further Research 5.** Catastrophic data loss may be more likely when users perform infrequent “clean up” operations than would be the case if they were encouraged to perform such tasks on a more regular basis.

**Hypothesis for Further Research 6.** File-system based organisational systems (e.g., when music albums are stored according to some attribute such as whether they are “classical” or “rock”) may present more opportunity for data loss than would be the case if taxonomical classification were a separate activity to organising digital objects within their storage location.

### 6.3 Poor Metadata Usage Despite Training

Participants reported using relatively few metadata elements to assist them in their deletion processes. This was consistent both in the home and work contexts<sup>13</sup>, largely relying upon the actual content of the digital object to inform their deletion decision, despite the fact that not only had the participants received some computer training from their organisations but had also invested significant amounts of their own time to obtain training in the use of computers<sup>14</sup>. The survey did not address any particular qualities of training received by the participants, however—it may be that training received was particular to a fairly narrow range of application functionality, rather than in how best to use a general-purpose computing device.

This finding is worrisome, because it indicates not a failure on the part of participants, but upon the part of the interfaces to digital objects: digital objects must be examined in detail in order to formulate a deletion decision, because metadata does not act as a sufficient surrogate to the content. While it is true that computer users may still wish to view the content of an object prior to deletion (e.g., due to worries that they may be losing something of value), the fact that participants with high levels of computer expertise report that they may not utilise metadata in their deletion decisions indicates that the operating systems do not facilitate decision-making based upon metadata. Thus, in order to perform many of the routine organisational tasks upon computer systems, computer users must examine the contents of the digital objects. If that is the case, then we may infer that those organisational tasks will not only be neglected, but that they will be carried out haphazardly, and with several varieties of cognitive error taking place during those organisational tasks<sup>15</sup>.

This finding is also troublesome because it indicates that metadata is not being provided, and that the metadata which *is* provided is not viewed as trustworthy or useful *even when*

---

<sup>13</sup>See section 5.1.5 on page 111 and section 5.1.6 on page 114.

<sup>14</sup>See section 5.1.7 on page 115 regarding computer training of participants.

<sup>15</sup>See section 2.11.2 on page 36 and section 2.11.3 on page 36

*generated by the same individual*. This results in a lack of easy manipulation of digital objects by their metadata, which in turn may result in a lack of disposition of digital objects: digital objects may linger far past when they are useful to anyone, only to be dealt with under extreme circumstances, or only disposed of or organised very rarely. In cases wherein there are legal concerns for the destruction of digital objects, this is undesirable because digital objects remaining beyond any mandated destruction date are potentially available to the legal discovery process; in cases wherein there is a desire to *locate* digital objects, this is undesirable, because “old” objects tend to obscure the presence of valuable objects; this is undesirable in cases wherein there is a desire for organisation because objects may only be assessed as individual objects rather than via the use of metadata.

There is some measure of risk, however, in thinking that the presentation of additional metadata to the user might result in a better decision: as shown by Shafir, Simonson, and Tversky [SST04, p. 956], “valueless features” are perceived as *negative* features<sup>16</sup>. Thus, if additional metadata were provided to computer users and it were perceived as being of no value, this additional metadata might actually provoke a less desirable decision than if the user were provided with less metadata: it may be expected that computer users would be influenced more heavily in favour of deletion *by* these valueless features simply by their being presented.

It is of some concern that participants use file content rather than metadata<sup>17</sup>. We must ask how, exactly, the study participants gain access to the content of these files about which they are making a decision to delete. In the case of email, this is perhaps trivial, as many if not all email clients provide easy access to such content via the use of a ‘preview pane’ (which may go some way towards explaining the frequency of email disposition). In the case of files accessed via an operating system, however, such is not the case: users must either open the document itself, or must remember the content in some manner. Considering the volume of digital objects under consideration—excluding text messages, blog posts, and social networking statuses because these digital objects are presented to the user via presenting their content—the fact that participants consider the content in the decision-making process is somewhat troublesome. This study did not enquire as to the volume of digital objects, but it is not unreasonable to believe that several thousand such digital objects are dispositioned each year by the participants. The participants allow that they do not “clean up” digital objects (aside from email) very frequently<sup>18</sup>; thus, in the process of making disposition decisions, participants are effectively stating that they periodically examine the content of many dozens or hundreds

---

<sup>16</sup>See section 2.11.3 on page 40.

<sup>17</sup>See section 5.1.5 on page 111.

<sup>18</sup>See section 5.1.4 on page 110.

of digital objects in order to determine whether to delete or preserve them. In the absence of metadata, such decisions would not only seem prone to cognitive errors, but would be subject to such human factors as fatigue and boredom: digital objects considered at the beginning of the process would be considered using different criteria (and different mental faculties) than those considered towards the end of the process.

**Hypothesis for Further Research 7.** If users were more knowledgeable about the types of metadata provided to them by their operating system or application software, they might be more likely to make use of such metadata rather than examining the contents of the digital object directly.

**Hypothesis for Further Research 8.** If metadata were more aggressively presented—e.g., displayed along with individual digital objects rather than displaying an icon—users might be more likely to make use of that metadata and, thus, be able to make dispositioning decisions more rapidly.

**Hypothesis for Further Research 9.** “Valueless features” presented to the user may provoke unexpected results; many metadata elements of digital objects may be perceived by users as being valueless, perhaps contributing to the lack of metadata usage reported by participants, but also possibly contributing to poor dispositioning decisions.

**Hypothesis for Further Research 10.** Because disposition decisions are frequently made based upon the content of individual digital objects, users may be less likely to make decisions about classes of digital object, resulting in less consistent application of retention or destruction goals.

## 6.4 Computer Usage Not Context-Dependent

A major finding of this study is that the reasons given for deletion and preservation decisions do not significantly vary whether the participant was using the computer at home or at work<sup>19</sup>. This was most unexpected because the reasons examined comprised of functional reasons<sup>20</sup>, reasons involving others<sup>21</sup>, and personal reasons<sup>22</sup>. These areas were selected for examination out of the expectation that they would differ significantly between the two contexts: that the reasons given for the home context would largely be concerned with the individual or their family,

---

<sup>19</sup>See section 5.1.9 on page 118

<sup>20</sup>See section 5.1.9 on page 118

<sup>21</sup>See section 5.1.9 on page 121

<sup>22</sup>See section 5.1.9 on page 120

with concerns for posterity, and would include far more personal reasons than that of the work context, whereas the work context would provide a different set of functional reasons and would include far fewer personal reasons given as reasons for deletion or preservation decisions. This was not the case.

That the reasons given for deletion and preservation decisions were largely identical between the two contexts may indicate that participants view the experience of interacting with a computer to be a single context, rather than considering the location of the computer itself. It could be that the cues provided by the operating systems in the two contexts (e.g., being some flavour of Windows in both places, both running the same web browser, and both running the same document manipulation packages) are so similar that participants did not distinguish between the experience of interacting with these systems, despite the fact that they may have been interacting with different digital objects<sup>23</sup>.

If it is the case, however, that the behaviour of participants is influenced by the cues presented by the computer systems, and that those cues are similar enough as to fail to indicate a difference in context, then this would seem to indicate that a change may need to take place so as to inform the computer user that they are in a different context, with different rules or different behaviours expected of them. It is not known whether such a change in computing environment would be sufficient to cause a change in behaviour, or if such a change would be desired, but it may be one means of achieving a change in behaviour between contexts. However, it is unknown and not determined by this study whether the computing environment is causative of the consistency in behaviour; while it may seem likely, given the consistency in the reasons given for decisions between both home and work contexts, it has not been proven.

This finding is also surprising in light of Habermas [Hab03]; Jones [Jon91]; Rapoport [Rap63]; Vaughan [Vau98]; Webster and Trevino [WT95]<sup>24</sup>. Those authors all considered that there would be some variation between different contexts—that people behave differently in different contexts. Given, the reasons may be the same between contexts while behaviour is in reality different. However, should normative engagement differ while reasons remain the same, there is something very important missing from the picture. Normative engagement *ought* to be different between contexts. Normative engagement *ought* to shape the decision-making process. If normative engagement is different and normative engagement shapes the decision-making process, then language used to describe that decision-making process *ought* to be different. Participants provided essentially the same reasons independent of context. Therefore, we should conclude that there is something missing from our understanding of normativity when it comes

---

<sup>23</sup>This will be considered further in section 6.4 on page 137

<sup>24</sup>See section 2.6.4 on page 23

to people making decisions facilitated by computers. Let us examine the argument logically<sup>25</sup>:

(P1) Home and work computers comprise different contexts.

(P2) Normative engagement is different between contexts.

(P3) Normative engagement shapes the decision-making process.

$(P1) \wedge (P2) \wedge (P3) \rightarrow (P4)$  Reasons given for decisions should vary between contexts.

(P5) Reasons given do not vary between contexts.

---

$\neg(P1) \vee \neg(P2) \vee \neg(P3) \vee \neg(P4)$

If we are to conclude  $\neg(P2)$ , then we are committed to disagreeing with the understandings of Habermas [Hab03]; Jones [Jon91]; Rapoport [Rap63]; Vaughan [Vau98]; Webster and Trevino [WT95], and many others. This does not seem a very likely conclusion in light of the extant research. Likewise if we were to conclude  $\neg(P3)$ —this just would not seem to make much sense given the nature of moral judgements<sup>26</sup>.  $\neg(P4)$ , on the other hand, might show some promise: it *could* be possible that people give the same reasons for their actions in different contexts, but perhaps to different degrees<sup>27</sup>. However, because some of the reasons given by study participants were personal reasons (e.g., they found something emotionally disturbing and therefore destroyed the content), this also does not seem likely. We are given (P5) by the study results, so are left with the conclusion that there is something flawed here, with the most likely conclusion being that this study did not capture any difference in context, thereby not fulfilling (P1). We may conclude, then, that “the computer” is a context in and of itself, regardless of where that computer is situated. We conclude  $\neg(P1)$ : home and work computers do not comprise different contexts, but comprise a single context with multiple instances.

Participants reported that the types of content that they generate, and the types of activities that they performed, were largely the same between the contexts of home and work<sup>28</sup>. This finding is interesting—not necessarily unexpected—because it indicates that people generally use their computer systems for the same types of activity regardless of context. There are, of course, a small number of activities conducted in one context and not in another (e.g., computer gaming), but the degree of overlap between the two contexts is significant because it indicates that there are few differences between the data systems used at home and those at work. This is an interesting finding because it demonstrates that the computer is a general-purpose device,

---

<sup>25</sup>The symbol ‘ $\wedge$ ’ denotes a logical conjunction. The symbol ‘ $\vee$ ’ denotes a logical disjunction. The symbol ‘ $\neg$ ’ denotes a logical negation. The symbol ‘ $\rightarrow$ ’ denotes a material implication.

<sup>26</sup>See section 2.6.4 on page 23

<sup>27</sup>See section 3.7.5 on page 88.

<sup>28</sup>See section 5.1.8 on page 116.

yet many of the same applications are installed on both the home and work computer or many of the applications used are within a web browser, further reducing context cues.

The similarity in content generated between the two contexts is also significant because this may alleviate some concerns regarding the possibility of different object types being considered between the two contexts<sup>29</sup> and, therefore, different reasons for deletion or preservation being reported. Because the content generated (and, subsequently, deleted or preserved) is fairly consistent between the two contexts, we may feel relatively confident that the similarity between reasons given for deletion and preservation is similar because the decision-making process is similar. If the content were dissimilar or if the reasons given were dissimilar then we might be concerned that a difference in content contributed to differences in reasons, or we might wish to conclude that the decisions made in the different contexts were affected *by* the context. Because the content between contexts is so uniform, however, we are able to be more confident that the context of ‘the computer’ really does comprise a single context.

Many participants reported that they interacted with the same information in both contexts (e.g., “recording personal financial information”), or that they kept personal digital objects at work and vice versa<sup>30</sup>. While not terribly surprising, this finding may indicate that participants give the same reasons for deletion and preservation decisions within both contexts—that participants, rather than reporting their responses as if such contexts were separate, were reporting that in some way they viewed the contexts as contiguous, without a hard boundary. If the contexts are identical not only in form, function, and usage, then they are effectively the same context: the idea that there is a home computing environment separate from the work computing environment only holds true insofar as there are distinct computer systems involved, and that those systems are in separate locations one of which is termed “home” and one “work.” Because of this degree of mixing, it is difficult to determine the cause for the similarity in reasons given for deletion and preservation decisions, i.e., are the reasons given so similar because participants view the context of “the computer” *in general* as being a single context, or do participants view the context of “the computer” as being a single context because they view their computer *instances* as comprising a single context?

**Hypothesis for Further Research 11.** It may be that radically altering the computing environment’s visual presentation in, e.g., the work context may force a computer user’s awareness of a difference in context and, thus, expectations regarding their decision-making processes.

---

<sup>29</sup>See section 3.7.6 on page 89 for a discussion of how different types of digital object may affect survey responses.

<sup>30</sup>See section 5.1.10 on page 123.

## 6.5 Records Management Systems Usage

Participants largely violate the records management policies of their workplace, yet provide some very interesting and perhaps valid reasons for these violations; participants do not feel that they are truly violating the policies, or that their behaviour should not be *considered as* a violation of these policies<sup>31</sup>. Reasons of efficiency were not the only ones provided as explanatory of why violating the records management policy was somehow a good thing, but were overwhelmingly in the majority of such explanatory comments. These responses result in the finding that, for at least half of the study's participants, records management policies either interfere with or detract from their work in some way. For these participants, there seems to be a tension between doing work and adhering to policy: they are aware that the policy exists, perhaps are even aware for the reason *for* the policy, but at the same time are expected to conduct their work and feel that their work would not be as efficient should they follow the records management policy to the letter.

**Hypothesis for Further Research 12.** It may be that users of RM systems feel they are not given enough flexibility to organise things to suit their own conception of how information *ought* to be organised—that the RM systems do not organise things as they would do and, therefore, users of RM systems work outside those systems for this reason.

**Hypothesis for Further Research 13.** It may be that users of RM systems are unable to “clean up” their working environment within the RM system such that they can easily locate digital objects, leading them to a less cluttered, local storage of those digital objects.

## 6.6 Summary Conclusions

### 6.6.1 Object Disposition

Participants in this study desire organisation of their digital objects, yet need to conduct periodic reorganisation of these objects. Participants generally accept the default storage locations proposed by their computer operating system when initially saving a digital object, then some months later move that digital object into a more appropriate storage location or delete it. In order to determine where to place their digital objects or whether to delete those digital objects, participants examine the content of each digital object rather than making use of metadata.

---

<sup>31</sup>Only 25% of participants agreed that they had violated their organisation's records management policy, yet 75% of participants provided explanatory comments as to why their behaviour was not in violation of said policy. See section 5.1.11 on page 125.

From this we may conclude that computer users are not supported by their computer operating systems at the level desired: that they are “working around” their computer systems, “making do” with the way things function today. We may also conclude that computer users are not making the decisions that they, themselves, would make under ideal conditions: if computer users are periodically reorganising their files, the operating systems could incorporate this reorganisation in such a way as to make use of machine learning combined with metadata, perhaps prompting the computer user through an organisation process, with subsequent digital object disposition actions incorporating that machine learning, resulting in a better “default” decision.

### **6.6.2 Context**

Participants perform the same types of activity on their computers regardless of context. Participants frequently conduct personal business at work and vice versa. Participants give the same reasons for their deletion and preservation actions regardless of context. If it is desired that the decision-making process differ from context to context, the operating systems need to incorporate some deeper means of encouraging decisions; perhaps this would be something so simple as the incorporation of context cues, or something more complex such as machine-learning technologies such that the decisions of the group could be pooled so as to allow greater consistency in the decisions being made. Such a change would not be so straightforward as implementing a set of rules, however, given the fact that work and home tend to intermingle; separating out the two contexts into discrete environments may prove difficult if not impossible, given the manner in which people tend to engage with work and personal information.

### **6.6.3 Records Management Systems**

Participants in this study felt the need to work around their organisations’ records management systems and policies in a manner similar to the way in which they worked around their computer operating systems’ proposed filing locations. In doing so, participants most frequently kept copies of digital objects in convenient locations so that they could have them for reference or because they needed them for a particular task or project. This behaviour would seem to indicate that participants were accustomed to relating to objects stored in familiar locations, be that on their local computer or in their email system. This should not be taken to mean that participants had not had adequate training with records management systems, but should indicate that participants were more comfortable with using their computer systems as they are now than they are comfortable with using records management systems. This area is one

in which changes to the underlying operating system could have the greatest impact: if records management systems and policies were implemented in a manner transparent to the computer user, then the user may be less likely to circumvent the system, simply because they would be more comfortable.

#### 6.6.4 Research Implications

The findings detailed by this study are perhaps most broadly applicable to computer interface design (HCI). There is a significant problem in the way in which general-purpose computing machines present information to users, in that deletion and preservation decisions are not structured so as to encourage the most optimal outcome: users approach such decisions in idiosyncratic and unique ways, based not upon some formalised decision-making process but based upon their own background, preferences, and indeed whims, not necessarily considering others in their decisions or considering that they themselves may ideally want to have made different decisions. This study demonstrates that the general-purpose user interfaces presented by computer operating systems, in presenting a huge range of choice with regards to information management, do not assist users to behave as they would wish of themselves—users make decisions not based upon some well-conceived plan for their needs, for their future selves, or out of a sense of duty to others, but make decisions which are affected by a wide range of social, psychological, and emotional. These issues clearly result in less than optimal circumstances.

These findings should be important to HCI designers not just of general-purpose operating systems but for designers of specific information systems, as the same issues have been found to affect such single-purpose systems as records-management systems. While generally speaking such systems may be imagined to be less susceptible to these issues, records-management systems were examined in particular because it was thought that they would provide a controlled arena against which to measure other behaviour; idiosyncratic behaviour was found in relating to records-management systems as in the context of other types of information-management systems, including decisions having been made from what can only be seen as emotional responses to information presented.

Further study should be made which examines the types of decisions users would ideally like to make within the information management arena, and such study should be extended into a formal design methodology including how best to facilitate such decisions. Such a decision process should incorporate the fact that users currently refer to the content of individual digital objects in order to disposition them; because users are currently required to consult the objects on multiple occasions, they are somewhat removed from context; if users were given the means

to disposition objects at the time of receipt, it is likely that disposition decisions would be of higher quality. Currently, such decisions regarding disposition of objects consist primarily of the options ‘delete,’ ‘file,’ or ‘do not decide right now.’ Such a limited set of choices could be expanded to incorporate options to delete at some future date, to preserve indefinitely but to hide from view after a certain date, to prompt the user for a decision at some future date—essentially, rather than being a general-purpose device, operating systems could actively engage with the user and assist them in making decisions about the future disposition of digital objects.

Future HCI design should include improvements in filtering information presented to the user, such that ad-hoc decisions are not made with regards to ‘all of the objects present in this context’ but only to those objects which are currently ‘active’, where ‘active’ may be taken to mean ‘those things about which a decision has not been made’, perhaps.

Further study should also be made with regards to incorporation of machine-learning technologies, in order that record- and archival-management decisions be aided by automation. For example, users should be able to instruct their machines that ‘all email from my estate agent should be moved to this folder and retained for seven years.’ This type of *ex ante* action might tend to incorporate a greater degree of conscious decision-making than is currently the case. Additionally, such actions, taken upon a large number of objects, would serve to limit the number of decisions required of the user in that they would not need to continually instruct their information to perform repetitive tasks; rather, such instructions would need to be made only in exceptional cases, and might tend to increase the quality of such decisions.

With regards to records-management systems, further research is required in order to determine some means of providing for customised views of records held in such systems. This is particularly true in light of those users whose sole reason for maintaining their own copies of certain records outside of their organisations’ records-management systems was because they feared not being able to easily locate those records again. Rather than providing a static view of such information, mediated via a technological implementation of a taxonomical structure, users would prefer to be able to organise such records into something they would find meaningful and easily negotiable. That is not to say that structures are not necessary, but that they are not necessary to the end user and that they tend to discourage the use of records-management systems (or, at least, to encourage the *abuse* of said systems). If users were presented with custom views of repositories, they might be encouraged to more closely adhere to institutional or regulatory standards with regards to copying.

Likewise, more research is required into the facilitation of group information management, with particular emphasis given to a technological implementation of retention and destruction decisions. Users currently view their computing practises as taking place within their own,

dedicated system; they do not regard the needs of others in their decisions to delete or preserve information. This is undesirable in a shared environment, nor is it desirable for each user to be required to maintain their own copies of any digital objects they believe may be important. Yet because users are accustomed to doing so, and because users do not agree on particular standards in terms of the organisation of digital objects, any technological solution must allow for each user to interact with digital objects as if those objects were present upon their own system while simultaneously preserving and deleting appropriately. Rather than manipulating the digital object itself (or a copy of said object), users could manipulate a symbolic link representing the digital object, moving it about, renaming it, ‘deleting’ it, while the digital object remained ensconced within the group’s management system. In this way, users would be allowed the flexibility granted by general-purpose computing devices, while the organisation would be granted some degree of security with regards to its content—security that the content would be maintained within the system, that the content would not be duplicated needlessly, and that the content would be appropriately subject to retention and destruction decisions.

---

# Bibliography

---

- [Are03] Hannah Arendt.  
*Responsibility and Judgment*.  
New York: Schocken Books, 2003,  
P. 295.
- [BS09] Mutlu Binark and Günseli Bayraktutan Sütçü.  
“Practicing Identity in the Digital Game World: The Turkish Tribes’ Community Practices in “Silkroad Online””.  
In: *Digital technologies of the self*.  
Ed. by Yasmine Abbas and Fred Dervin.  
Newcastle: Cambridge Scholars, 2009.  
Chap. 3, pp. 61–84.
- [BWR06] Hilla. Brink, C. Van der Walt, and Gisela. Van Rensburg.  
*Fundamentals of research methodology for health care professionals*.  
Cape Town: Juta, 2006.
- [Bai08] Steve Bailey.  
*Managing the crowd : rethinking records management for the Web 2.0 world*.  
London: Facet Publishing, 2008,  
P. 172.
- [Bal99] Anne Balsamo.  
“Reading Cyborgs, Writing Feminism”.  
In: *Cybersexualities: a reader on feminist theory cyborgs and cyberspace*.  
Ed. by Jenny Wolmark.  
Includes bibliographical references and index. Edited and with an introduction by Jenny Wolmark.  
Edinburgh University Press, 1999.  
Chap. 7, pp. 145–156.
- [Bov04] Mark Bovens.  
*The quest for responsibility : accountability and citizenship in complex organisations*.  
Cambridge: Cambridge University Press, 2004,  
P. 252.

- [Bry08] Alan Bryman.  
*Social research methods*.  
Oxford [u.a.: Oxford University Press, 2008,  
P. 748.
- [CH95] Deborah R. Compeau and Christopher A. Higgins.  
“Computer Self-Efficacy: Development of a Measure and Initial Test”.  
In: *MIS Quarterly* 19.2 (1995). <http://www.jstor.org/stable/249688>, pp. 189–211.
- [Coh07] Laura B. Cohen, ed.  
*Library 2.0 initiatives in academic libraries*.  
Chicago: Association of College and Research Libraries, 2007.
- [DR09] Fred Dervin and Tanja Riikonen.  
““Whatever I am, Wherever I am, How does it Matter?... Why does it Matter?”  
Egocasting in-between Identities”.  
In: *Digital technologies of the self*.  
Ed. by Yasmine Abbas and Fred Dervin.  
Newcastle: Cambridge Scholars, 2009.  
Chap. 6, pp. 125–156.
- [DX04] J. Dutra and Q. Xiao.  
*A NASA Technical White Paper: Implementing the NASA Taxonomy Through  
Service Oriented Architectures to Promote Knowledge Sharing and Increased  
Mission Success*.  
Tech. rep.  
<http://trs-new.jpl.nasa.gov/dspace/bitstream/2014/37996/1/04-0370.pdf>.  
National Aeronautics and Space Administration, 2004.  
(Visited on 10/18/2009).
- [EN98] S.L. Esler and M.L. Nelson.  
“Evolution of scientific and technical information distribution”.  
In: *Journal of the American Society for Information Science* 49.1 (1998).  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.60.481&rep=rep1&type=pdf>,  
pp. 82–91.  
(Visited on 10/18/2009).
- [Edm93] William A. Edmundson.  
“Review: Rethinking Exclusionary Reasons: A Second Edition of Joseph Raz’s  
“Practical Reason and Norms””.  
In: *Law and Philosophy* 12.3 (1993). <http://www.jstor.org/stable/3504852>,  
pp. 329–343.
- [FT04a] Craig R. Fox and Amos Tversky.  
“A Belief-Based Account of Decision under Uncertainty”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 32, pp. 795–823.

- [FT04b] Craig R. Fox and Amos Tversky.  
“Ambiguity Aversion and Comparative Ignorance”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 31, pp. 778–795.
- [Foo89] Philippa Foot.  
“Utilitarianism and the Virtues”.  
In: *Ethical theory : classical and contemporary readings*.  
Ed. by Louis P. Pojman.  
Belmont, Calif.: Wadsworth Pub. Co., 1989.  
Chap. 5.7, pp. 205–224.
- [Fox92] Robin Fox.  
“Prejudice and the Unfinished Mind: A New Look at an Old Failing”.  
In: *Psychological Inquiry* 3.2 (1992). <http://www.jstor.org/stable/1449204>,  
pp. 137–152.
- [GN82] Mary Ann Von Glinow and Jr. Novelli Luke.  
“Ethical Standards within Organizational Behavior”.  
In: *The Academy of Management Journal* 25.2 (1982).  
<http://www.jstor.org/stable/256001>, pp. 417–436.
- [GT04] Itamar Gati and Amos Tversky.  
“Weighting Common and Distinctive Features in Perceptual and Conceptual  
Judgments”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 4, pp. 97–128.
- [Gon99] Jennifer Gonzalez.  
“Envisioning Cyborg Bodies: Notes from Current Research”.  
In: *Cybersexualities: a reader on feminist theory cyborgs and cyberspace*.  
Ed. by Jenny Wolmark.  
Includes bibliographical references and index. Edited and with an introduction by  
Jenny Wolmark.  
Edinburgh University Press, 1999.  
Chap. 14, pp. 264–279.
- [HT04] Chip Heath and Amos Tversky.  
“Preference and Belief: Ambiguity and Competence in Choice under Uncertainty”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 26, pp. 645–673.
- [Hab03] Jurgen Habermas.  
*Truth and Justification*.  
Cambridge, Mass.: MIT Press, 2003,  
P. 327.

- [Hei93] Martin Heidegger.  
*Being and Time*.  
Translated by John Macquarrie and Edward Robinson.  
Oxford [u.a.]: Blackwell, 1993,  
P. 589.
- [Hos89] John Hospers.  
“Rule-Utilitarianism”.  
In: *Ethical theory : classical and contemporary readings*.  
Ed. by Louis P. Pojman.  
Belmont, Calif.: Wadsworth Pub. Co., 1989.  
Chap. 5.5, pp. 201–210.
- [Hum89] David Hume.  
“On Reason and the Emotions: the Fact/Value Distinction”.  
In: *Ethical theory : classical and contemporary readings*.  
Ed. by Louis P. Pojman.  
Belmont, Calif.: Wadsworth Pub. Co., 1989.  
Chap. 8.1, pp. 405–412.
- [IMR97] Masao Ito, Yasushi Miyashita, and Edmund T. Rolls.  
*Cognition, computation, and consciousness*.  
Oxford; New York: Oxford University Press, 1997.
- [JIS09] JISC infoNet.  
*Records and Information Management Maturity Model*.  
Excel Spreadsheet.  
<http://www.jiscinfonet.ac.uk/records-management/measuring-impact/maturity-model/index.htm>  
2009.  
(Visited on 11/15/2009).
- [Jon07] Catherine. Jones.  
*Institutional repositories : content and culture in an open access environment*.  
Chandos information professional series.  
Oxford: Chandos, 2007.
- [Jon91] Thomas M. Jones.  
“Ethical Decision Making by Individuals in Organizations: An Issue-Contingent Model”.  
In: *The Academy of Management Review* 16.2 (1991).  
<http://www.jstor.org/stable/258867>, pp. 366–395.
- [KRT04] Mark Kelman, Yuval Rottenstreich, and Amos Tversky.  
“Context-Dependence in Legal Decision Making”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 40, pp. 963–995.
- [Kan07] Immanuel Kant.  
*Critique of Pure Reason*.  
Basingstoke, New York: Palgrave Macmillan, 2007.

- [Kim00] Jong-Young Kim.  
*Social Interaction in Computer-Mediated Communication*.  
<http://www.asis.org/Bulletin/Mar-00/kim.html>.  
Mar. 2000.
- [Lat05] Bruno Latour.  
*Reassembling the Social: An Introduction to Actor Network Theory*.  
Washington, D.C.: Oxford University Press, 2005,  
P. 301.
- [Les06] Lawrence Lessig.  
*Code : version 2.0 / Lawrence Lessig*.  
Previous ed.: 1999. Includes bibliographical references and index.  
New York: Basic Books, 2006.
- [Lev01] David M. Levy.  
*Scrolling Forward : making sense of documents in the digital age*.  
New York: Arcade Publishing, 2001,  
P. 212.
- [MS97] James H. McMillan and Sally. Schumacher.  
*Research in education : a conceptual introduction*.  
New York: Longman, 1997.
- [MSC08] MSCI Barra.  
*Global Industry Classification Standard*.  
List of code and description, Microsoft Excel format.  
[http://www.msibarra.com/resources/xls/GICS\\_map2008.xls](http://www.msibarra.com/resources/xls/GICS_map2008.xls).  
Aug. 2008.  
(Visited on 11/02/2009).
- [Mur97] Mark C. Murphy.  
“Surrender of Judgment and the Consent Theory of Political Authority”.  
In: *Law and Philosophy* 16.2 (1997). <http://www.jstor.org/stable/3505022>,  
pp. 115–143.
- [Nor02] Donald A. Norman.  
*The Design of Everyday Things*.  
New York: Basic Books, 2002.
- [PP02] Michael Quinn. Patton and Michael Quinn. Patton.  
*Qualitative research and evaluation methods*.  
Thousand Oaks, Calif.: Sage Publications, 2002.
- [Psp] *GNU PSPP*.  
<http://www.gnu.org/software/pspp/>.  
2010.  
(Visited on 10/12/2010).
- [QT04] George A. Quattrone and Amos Tversky.  
“Self-Deception and the Voter’s Illusion”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 33, pp. 825–845.

- [RT04] Donald A. Redelmeier and Amos Tversky.  
“Discrepancy between Medical Decisions for Individual Patients and for Groups”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 36, pp. 887–895.
- [Rap63] Anatol Rapoport.  
“Formal Games as Probing Tools for Investigating Behavior Motivated by Trust and Suspicion”.  
In: *The Journal of Conflict Resolution* 7.3 (1963).  
<http://www.jstor.org/stable/173066>, pp. 570–579.
- [Raz75] Joseph Raz.  
“Permissions and Supererogation”.  
In: *American Philosophical Quarterly* 12.2 (1975).  
<http://www.jstor.org/stable/20009570>, pp. 161–168.
- [Raz85] Joseph Raz.  
“Review: The Morality of Obedience”.  
In: *Michigan Law Review* 83.4 (1985). <http://www.jstor.org/stable/1288770>,  
pp. 732–749.
- [Raz90] Joseph Raz.  
“Facing Diversity: The Case of Epistemic Abstinence”.  
In: *Philosophy and Public Affairs* 19.1 (1990).  
<http://www.jstor.org/stable/2265361>, pp. 3–46.
- [SL05] Bridget. Somekh and Cathy. Lewin.  
*Research methods in the social sciences*.  
London; Thousand Oaks, Calif.: SAGE Publications, 2005.
- [SST04] Eldar Shafir, Itamar Simonson, and Amos Tversky.  
“Reason-Based Choice”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 39, pp. 937–963.
- [ST04] Eldar Shafir and Amos Tversky.  
“Thinking through Uncertainty: Nonconsequential Reasoning and Choice”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 28, pp. 703–729.

- [Sto99] Allucquere Rosanne Stone.  
“Will the Real Body Please Stand Up? Boundary Stories about Virtual Cultures”.  
In: *Cybersexualities: a reader on feminist theory cyborgs and cyberspace*.  
Ed. by Jenny Wolmark.  
Includes bibliographical references and index. Edited and with an introduction by  
Jenny Wolmark.  
Edinburgh University Press, 1999.  
Chap. 4, pp. 69–98.
- [Sug98] Robert Sugden.  
“The Role of Inductive Reasoning in the Evolution of Conventions”.  
In: *Law and Philosophy* 17.4 (July 1998). Issue Title: Rationality, Conventions,  
and Law / Full publication date: Jul., 1998 / Copyright 1998 Springer,  
pp. 377–410.  
URL: <http://www.jstor.org/stable/3505086>.
- [TG04] Amos Tversky and Thomas Gilovich.  
“The Cold Facts about the ”Hot Hand” in Basketball”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 10, pp. 257–267.
- [TK04a] Amos Tversky and Daniel Kahneman.  
“Extensional vs. Intuitive Reasoning: The Conjunction Fallacy in Probability  
Judgment”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 9, pp. 221–257.
- [TK04b] Amos Tversky and Daniel Kahneman.  
“Judgment under Uncertainty: Heuristics and Biases”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 8, pp. 203–220.
- [TK04c] Amos Tversky and Daniel Kahneman.  
“Rational Choice and the Framing of Decisions”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 24, pp. 593–621.
- [Tur84] Sherry Turkle.  
*The second self : computers and the human spirit*.  
Cambridge, Mass: MIT Press, 1984.
- [Tur97] Sherry Turkle.  
*Life on the screen : identity in the age of the internet*.  
New York: Touchstone, 1997.

- [Tve04] Amos Tversky.  
“Features of Similarity”.  
In: *Preference, Belief, and Similarity : Selected Writings*.  
Ed. by Eldar Shafir.  
Cambridge Mass.: MIT Press, 2004.  
Chap. 1, pp. 7–41.
- [U.K00] U.K. Office for National Statistics.  
*Standard Occupational Classification, 2000*.  
Excel Spreadsheet.  
<http://www.ons.gov.uk/ons/index.html>.  
2000.  
(Visited on 11/02/2009).
- [U.S09] U.S. Office of Personnel Management.  
*Position Classification Standards for White Collar Work*.  
Web Page.  
<http://www.opm.gov/fedclass/html/gsseries.asp>.  
2009.  
(Visited on 11/02/2009).
- [Uni06] United States Census Bureau.  
*North American Industry Classification System, 2007*.  
List of code and description, Microsoft Excel format.  
The URL provided contains links to several separate documents. The 2007 code list, full hierarchy, was utilised in surveys. The correlation tables were utilised to bridge statistics available in the 2002 version to match those of the 2007 version of the code list.  
May 2006.  
URL: <http://www.census.gov/epcd/naics07/> (visited on 10/26/2009).
- [Uni09] United Nations Statistics Division.  
*International Classification of Activities for Time-Use Statistics*.  
Web page.  
<http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=231&Lg=1>.  
2009.  
(Visited on 11/02/2009).
- [Vau98] Diane Vaughan.  
“Rational Choice, Situated Action, and the Social Control of Organizations”.  
In: *Law & Society Review* 32.1 (1998). <http://www.jstor.org/stable/827748>,  
pp. 23–61.
- [WT95] Jane Webster and Linda Klebe Trevino.  
“Rational and Social Theories as Complementary Explanations of Communication Media Choices: Two Policy-Capturing Studies”.  
In: *The Academy of Management Journal* 38.6 (1995).  
<http://www.jstor.org/stable/256843>, pp. 1544–1572.

- [Wan+09] J. Wang et al.  
“MHS: A distributed metadata management strategy”.  
In: *The Journal of Systems & Software* (2009).  
<http://linkinghub.elsevier.com/retrieve/pii/S0164121209001551>.  
(Visited on 10/18/2009).
- [Wil11] Scott S. Wiltermuth.  
“Cheating more when the spoils are split”.  
In: *Organizational Behavior and Human Decision Processes* 115.2 (2011).  
<http://www.sciencedirect.com/science/article/pii/S0749597810000841>, pp. 157–168.  
DOI: DOI:10.1016/j.obhdp.2010.10.001.
- [Woo99] Kathleen Woodward.  
“From Virtual Cyborgs to Biological Time Bombs: Technocriticism and the Material Body”.  
In: *Cybersexualities: a reader on feminist theory cyborgs and cyberspace*.  
Ed. by Jenny Wolmark.  
Includes bibliographical references and index. Edited and with an introduction by Jenny Wolmark.  
Edinburgh University Press, 1999.  
Chap. 15, pp. 280–294.

# Appendices

## Appendix A

---

# Survey

---

### A.1 Subjectivity in Information Retention

I hope that you don't find it too tedious, and that the questions are clear and understandable. If you have any problems, please let me know - I want to get this RIGHT! If you are the least bit confused about what I'm asking, please let me know - others may be confused, as well.

You may contact me via email at: [d.macknet.1@research.gla.ac.uk](mailto:d.macknet.1@research.gla.ac.uk) . Further information about my research may be found at:

<http://www.hatii.arts.gla.ac.uk/staff/dm.html>

If you have any concerns about this survey, or if you feel uncomfortable in contacting me directly, you may contact my research supervisor, Dr. Ian Anderson. He is available via email at: [I.Anderson@hatii.arts.gla.ac.uk](mailto:I.Anderson@hatii.arts.gla.ac.uk) . Further contact information for Ian is available at:

<http://www.hatii.arts.gla.ac.uk/staff/ia.html>

This is a study about how people interact with the information on their computers. It has been designed for people who use a computer in the home as well as at work, and whose home cultures are those of the US, the British Isles, and Canada. This survey examines differences between personal computer usage and work computer usage. Because of this, some of the questions are asked twice: once to consider how you use your home computer, and once to consider how you use your work computer. *If you do not use a different computer at home and at work, please do not take this survey.* In testing this survey, most respondents completed the survey in approximately 25 minutes. Some were found to have completed it in as little as 17 minutes, and some took as long as 45 minutes. You may save your survey and return to complete it later. Obviously, though, the sooner you complete the survey, the sooner your responses can be analyzed, so you may receive a reminder email if you have only completed part of the survey. This survey protects your privacy:

- All data about you (your name and email address) is kept in a separate part of the system.
- Nobody looking at the response data will connect your personal information with your responses.
- The analysis and publication of the response data will be entirely anonymous.
- Personal information will be destroyed no later than September, 2011.
- The anonymous response data will be held on secure network storage and be retained to facilitate further research.

It is necessary, though, to log in using a valid email address, for several reasons:

- If you get tired of taking the survey, you can save your progress and come back later.
- If you don't finish your survey, we'd like to send you a reminder to finish (incomplete surveys can't be considered in the research).
- If you wish, and if it is necessary, we would like to be able to ask you to participate in further research.
- When this research is complete, you will be sent an email informing you of any interesting results.

There are 57 questions in this survey

## 1/12: Getting Started

### 1 [0001]Where is “home” for you?

This doesn't mean, “where is your house?” Rather, this is asking about which region you feel best reflects your upbringing.

Please choose **only one** of the following:

- United States     
  British Isles     
  Canada     
  Other

### 2 [0002]What is your home state?

Only answer this question if the following conditions are met:

° Answer was 'United States' at question '1 [0001]' (Where is “home” for you?)

Please choose **only one** of the following:

- |  |                                      |  |
|--|--------------------------------------|--|
| <input type="radio"/> Alabama              | <input type="radio"/> Maryland       | <input type="radio"/> South Carolina                 |
| <input type="radio"/> Alaska               | <input type="radio"/> Massachusetts  | <input type="radio"/> South Dakota                   |
| <input type="radio"/> Arizona              | <input type="radio"/> Michigan       | <input type="radio"/> Tennessee                      |
| <input type="radio"/> Arkansas             | <input type="radio"/> Minnesota      | <input type="radio"/> Texas                          |
| <input type="radio"/> California           | <input type="radio"/> Mississippi    | <input type="radio"/> Utah                           |
| <input type="radio"/> Colorado             | <input type="radio"/> Missouri       | <input type="radio"/> Vermont                        |
| <input type="radio"/> Connecticut          | <input type="radio"/> Montana        | <input type="radio"/> Virginia                       |
| <input type="radio"/> Delaware             | <input type="radio"/> Nebraska       | <input type="radio"/> Washington                     |
| <input type="radio"/> District of Columbia | <input type="radio"/> Nevada         | <input type="radio"/> West Virginia                  |
| <input type="radio"/> Florida              | <input type="radio"/> New Hampshire  | <input type="radio"/> Wisconsin                      |
| <input type="radio"/> Georgia              | <input type="radio"/> New Jersey     | <input type="radio"/> Wyoming                        |
| <input type="radio"/> Hawaii               | <input type="radio"/> New Mexico     | <input type="radio"/> American Samoa                 |
| <input type="radio"/> Idaho                | <input type="radio"/> New York       | <input type="radio"/> Federated States of Micronesia |
| <input type="radio"/> Illinois             | <input type="radio"/> North Carolina | <input type="radio"/> Guam                           |
| <input type="radio"/> Indiana              | <input type="radio"/> North Dakota   | <input type="radio"/> Marshall Islands               |
| <input type="radio"/> Iowa                 | <input type="radio"/> Ohio           | <input type="radio"/> Northern Mariana Islands       |
| <input type="radio"/> Kansas               | <input type="radio"/> Oklahoma       | <input type="radio"/> Palau                          |
| <input type="radio"/> Kentucky             | <input type="radio"/> Oregon         | <input type="radio"/> Puerto Rico                    |
| <input type="radio"/> Louisiana            | <input type="radio"/> Pennsylvania   | <input type="radio"/> U.S. Minor Outlying Islands    |
| <input type="radio"/> Maine                | <input type="radio"/> Rhode Island   | <input type="radio"/> Virgin Islands of the U.S.     |

### 3 [0003]What is your home region?

Only answer this question if the following conditions are met:

° Answer was 'British Isles' at question '1 [0001]' (Where is “home” for you?)

Please choose **only one** of the following:

- England     
  Scotland     
  Wales  
 Northern Ireland     
  Republic of Ireland     
  Channel Islands  
 Isle of Man



**2/12: Your Home Computer Use**

This section will examine how you use your computer at home.

**6 [0006] Do you use a computer at home?**

Please choose **only one** of the following:

- Yes  
 No

**7 [0007] Do you share this computer with anyone else?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '6 [0006]' (Do you use a computer at home?)

Please choose **only one** of the following:

- Yes  
 No

**8 [0008] Do you also share your user account / login?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '6 [0006]' (Do you use a computer at home?)

*and* Answer was 'Yes' at question '7 [0007]' (Do you share this computer with anyone else?)

In other words, When you log into your computer, do you click on or enter your username, or do you use the same security account as others use? Please choose **only one** of the following:

- Yes  
 No

**9 [0009] Which operating system(s) do you use, on your home computer(s)?**

Please choose **all** that apply:

- Microsoft Windows  
 Macintosh (OS X)  
 Macintosh (older OS)  
 Linux

Other: \_\_\_\_\_

**10 [0010]On your home computer, how often do you “clean up” or organize the following:**

Please choose the appropriate response for each item:

	Every day	Once a week	Once a month	Several times a year	Every year	Hardly ever	Never (or Not Applicable)
Your home email	<input type="radio"/>						
Your email contacts	<input type="radio"/>						
Documents on your hard-drive	<input type="radio"/>						
Your social networking statuses or posts	<input type="radio"/>						
Your text messages	<input type="radio"/>						
Your mobile phone contacts	<input type="radio"/>						

**11 [0011]What percentage of your home computing time every week is spent doing each of the following?**

It is perfectly reasonable for these to add up to more than 100%. We understand that some things (e.g., Creating or editing online content and Documenting personal information about yourself) may overlap. This is OK!  
Please write your answer(s) here:

- \_\_\_\_\_ Watching Video (excluding Video Conferencing).
- \_\_\_\_\_ Conference calling (including video conferencing).
- \_\_\_\_\_ Searching the Internet.
- \_\_\_\_\_ Engaging with Social Media (e.g., Facebook, Twitter).
- \_\_\_\_\_ Creating or editing content not on the web.
- \_\_\_\_\_ Creating or editing online content.
- \_\_\_\_\_ Composing or reading email.
- \_\_\_\_\_ Playing Games.
- \_\_\_\_\_ Reading / distance learning.
- \_\_\_\_\_ Recording financial information about yourself.
- \_\_\_\_\_ Recording other financial information.
- \_\_\_\_\_ Documenting personal information about yourself.
- \_\_\_\_\_ Documenting other personal information.
- \_\_\_\_\_ Chatting (e.g., Instant Messenger, Google Chat, etc.)

**3/12: Things you keep on your Home Computer**

This section is all about what sorts of things you keep or save when using your home computer.

**12 [0012]Do you use the following built-in locations for storage?**

Please choose the appropriate response for each item:

	Yes	Uncertain	No
Documents	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Downloads	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Library	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Movies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Music	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pictures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have created your own folder structure for pictures, for example (not the folder which is default on your computer), then you should answer 'No'.

If you keep all of your pictures in the Pictures (or My Pictures) folder, then you should answer 'Yes'.

**13 [0013]Do you file things in particular folders, in your...**

Please choose **all** that apply:

- email?
- documents on your hard-drive?
- music files?
- pictures?
- video files?
- personal records (e.g., tax documents)?

**14 [0014]Which of the following describe reasons you have for keeping things on your home computer? Please consider all possible locations such as in your personal email, on your network, or on the web (e.g., Google Docs).**

Check any that apply. If you wish, provide an example of the kind of thing you use this reason for.

- I might need it again. \_\_\_\_\_
- It reminds me of something important. \_\_\_\_\_
- I want to spend time on it. \_\_\_\_\_
- It's important to the household. \_\_\_\_\_
- Somebody else might need it. \_\_\_\_\_
- I spent a lot of time on it. \_\_\_\_\_
- Somebody else spent a lot of time on it. \_\_\_\_\_
- It's important for legal reasons. \_\_\_\_\_
- I'm emotionally attached to it. \_\_\_\_\_
- It's entertaining / interesting. \_\_\_\_\_
- It doesn't take up much space. \_\_\_\_\_
- It's my original creation. \_\_\_\_\_
- To prove something about my actions. \_\_\_\_\_



### 5/12: Things you delete on your Home Computer

This section is all about what sorts of things you delete or erase when using your home computer.

#### 21 [0021] Which of the following describe reasons you have for deleting /erasing things on your home computer? Please consider all possible locations such as in your personal email, on your network, or on the web (e.g., Google Docs).

If you wish, provide an example of the kind of thing you use this reason for.

- I don't need it any longer. \_\_\_\_\_
- It would be wrong to keep it. \_\_\_\_\_
- No one else needs it. \_\_\_\_\_
- It gets in the way when I'm looking for things. \_\_\_\_\_
- It's old. \_\_\_\_\_
- Somebody else has a copy. \_\_\_\_\_
- It's too much trouble to keep. \_\_\_\_\_
- I have no good way to save it. \_\_\_\_\_
- I'm emotionally troubled by it. \_\_\_\_\_
- It's not important to me. \_\_\_\_\_
- It takes up too much space. \_\_\_\_\_
- It's confidential / sensitive. \_\_\_\_\_
- To protect my privacy or someone else's privacy. \_\_\_\_\_

#### 22 [0022] Which of the following helps you decide what to delete?

This question is not asking whether you think this might be useful. Rather, this question is asking whether you actually, in deciding to delete something, use these to help you decide.  The folder it's in.

- Its file type.
- Its file name.
- Its creation date.
- Its modified date.
- In email, the sender's name.
- In email, the subject.
- In email, whether the message is 'flagged'.
- In email, whether the message has attachments.
- The content of the item.
- Other: \_\_\_\_\_

**6/12: Sharing between Work and Personal Computers**

This section is about the things you might share between your personal computer and your work computer. For example, you make take work files home, or have personal files you also keep at work.

**23 [0023]Do you keep things on your home computer that you believe belong to your work?**

Please choose **only one** of the following:

- Yes  
 No

**24 [0024]If you were to leave your work, how likely are you to keep these things?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '23 [0023]' (Do you keep things on your home computer that you believe belong to your work?)

Please choose **only one** of the following:

- Very Likely  
 Likely  
 Somewhat Likely  
 Perhaps  
 Maybe Not  
 Most Likely Not  
 Not At All Likely

**25 [0025]Do you believe that you should keep these things?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '23 [0023]' (Do you keep things on your home computer that you believe belong to your work?) *and* Answer was NOT 'Not At All Likely' at question '24 [0024]' (If you were to leave your work, how likely are you to keep these things?)

Please choose **only one** of the following:

- Yes  
 No

In other words, is it right for you to keep these things?

**26 [0026]Why do you believe that you should keep these things?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '23 [0023]' (Do you keep things on your home computer that you believe belong to your work?) *and* Answer was NOT 'Not At All Likely' at question '24 [0024]' (If you were to leave your work, how likely are you to keep these things?) *and* Answer was 'Yes' at question '25 [0025]' (Do you believe that you should keep these things?)

Please write your answer here:

**27 [0027]If you believe that you will keep these things and also believe that you shouldn't, please explain why you believe that you will do so?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '23 [0023]' (Do you keep things on your home computer that you believe belong to your work?) *and* Answer was NOT 'Not At All Likely' at question '24 [0024]' (If you were to leave your work, how likely are you to keep these things?) *and* Answer was 'No' at question '25 [0025]' (Do you believe that you should keep these things?)

Please write your answer here:

**28 [0028]Are there things that you keep on your work computer which you feel that you own, personally?**

Please choose **only one** of the following:

- Yes  
 No

**29 [0029]Would you be upset if your workplace were to examine these things?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '28 [0028]' (Are there things that you keep on your work computer which you feel that you own, personally?)

Please choose **only one** of the following:

- Yes  
 No

**30 [0030]Please explain your feelings.**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '28 [0028]' (Are there things that you keep on your work computer which you feel that you own, personally?)

Please write your answer here:

**7/12: Your Work Computer Use**

This section will examine how you use your computer at work.

**31 [0031] Do you use a computer at work?**

Please choose **only one** of the following:

- Yes  
 No

**32 [0032] Do you share this computer with anyone else?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '31 [0031]' (Do you use a computer at work?)

Please choose **only one** of the following:

- Yes  
 No

**33 [0033] Do you also share your user account / login?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '31 [0031]' (Do you use a computer at work?) *and* Answer was 'Yes' at question '32 [0032]' (Do you share this computer with anyone else?)

Please choose **only one** of the following:

- Yes  
 No

In other words, When you log into your computer, do you click on or enter your username, or do you use the same security account as others use?

**34 [0034] Please estimate what percentage of your job skills come from the following:**

This question is about where you learned the computer skills you use on the job. There is no way to be precise, but please think carefully about your estimate.

Please write your answer(s) here:

\_\_\_\_\_ On the job

\_\_\_\_\_ On your own (during work time)

\_\_\_\_\_ On your own (outside of work)

\_\_\_\_\_ Formal education / training (during work time)

\_\_\_\_\_ Formal education / training (outside of work)

\_\_\_\_\_ Seminars and conferences

\_\_\_\_\_ Other

**35 [0035]What percentage of your work computing time is spent doing each of the following?**

It is perfectly reasonable for these to add up to more than 100%. We understand that some things (e.g., Creating or editing online content and Documenting personal information about yourself) may overlap. This is OK!  
Please write your answer(s) here:

- \_\_\_\_\_ Watching Video (excluding Video Conferencing).
- \_\_\_\_\_ Conference calling (including video conferencing).
- \_\_\_\_\_ Searching the Internet.
- \_\_\_\_\_ Engaging with Social Media (e.g., Facebook, Twitter).
- \_\_\_\_\_ Creating or editing content not on the web.
- \_\_\_\_\_ Creating or editing online content.
- \_\_\_\_\_ Composing or reading email.
- \_\_\_\_\_ Playing Games.
- \_\_\_\_\_ Reading / distance learning.
- \_\_\_\_\_ Recording financial information about yourself.
- \_\_\_\_\_ Recording other financial information.
- \_\_\_\_\_ Documenting personal information about yourself.
- \_\_\_\_\_ Documenting other personal information.
- \_\_\_\_\_ Chatting (e.g., Instant Messenger, Google Chat, etc.)

**8/12: Things you keep on your Work Computer**

This section is all about what sorts of things you keep or save when using your work computer.

**36 [0036] Which of the following describe reasons you have for keeping things on your work computer? Please consider all possible locations such as in your personal email, on your network, or on the web (e.g., Google Docs, if it is used for your work).**

If you wish, provide an example of the kind of thing you use this reason for.

Please choose all that apply and provide a comment:

- I might need it again. \_\_\_\_\_
- It reminds me of something important. \_\_\_\_\_
- I want to spend time on it. \_\_\_\_\_
- It's important to the business. \_\_\_\_\_
- Somebody else might need it. \_\_\_\_\_
- I spent a lot of time on it. \_\_\_\_\_
- Somebody else spent a lot of time on it. \_\_\_\_\_
- It's important for legal reasons. \_\_\_\_\_
- I'm emotionally attached to it. \_\_\_\_\_
- It's entertaining / interesting. \_\_\_\_\_
- It doesn't take up much space. \_\_\_\_\_
- It's my original creation. \_\_\_\_\_
- To prove something about my actions. \_\_\_\_\_

**37 [0037] Does your workplace provide you with email?**

Please choose **only one** of the following:

- Yes
- No

**38 [0038]Email practices at your work.**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '37 [0037]' (Does your workplace provide you with email?)

Please choose the appropriate response for each item:

	Yes	Uncertain	No
When sending an email, do you save a copy?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If you save a copy, do you file the copy in a folder immediately?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you "auto-archive" your email?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you move email from your email application to some other storage?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**39 [0039]Do you work with documents which are managed centrally?**

Please choose **only one** of the following:

- Yes
- No

For example, do you work with patient medical records, personnel records, invoices, financial statements, official reports, trouble tickets, or other things which are tracked by your business in some special manner, or in a dedicated system?

**40 [0040]How frequently do you keep copies of any of these documents?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '39 [0039]' (Do you work with documents which are managed centrally?)

For example, if you send these documents on to someone else, do you keep an electronic copy in email, or on your computer?

Please choose **only one** of the following:

- Every day
- Once a week
- Once a month
- Several times a year
- Every year
- Hardly ever
- Never (or Not Applicable)

**41 [0041]Why do you keep these copies?**

Only answer this question if the following conditions are met:

° Answer was 'Yes' at question '39 [0039]' (Do you work with documents which are managed centrally?) *and* Answer was NOT 'Never (or Not Applicable)' at question '40 [0040]' (How frequently do you keep copies of any of these documents?)

Please write your answer here:

**42 [0042]Do you think it's important to keep documents and email organized on your work computer? Why, or why not?**

Please write your answer here:



**46 [0046] Please answer the following about document control / records management in your workplace.**

Please choose the appropriate response for each item:

	Yes	Uncertain	No
Is there someone in your workplace who is responsible for managing documents and records?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Does your workplace have a document / records management policy?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have you ever been instructed about what things to keep, create, or delete?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Does your workplace provide formal training on records management to all employees?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Does your workplace undergo periodic audits of its documents?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**10/12: Things you delete on your Work Computer**

This section is all about what sorts of things you delete or erase when using your work computer.

**47 [0047] Which of the following describe reasons you have for deleting / erasing things on your work computer? Please consider all possible locations such as in your personal email, on your network, or on the web (e.g., Google Docs, if it is used for your work).**

If you wish, provide an example of the kind of thing you use this reason for.

- I don't need it any longer. \_\_\_\_\_
- It would be wrong to keep it. \_\_\_\_\_
- No one else needs it. \_\_\_\_\_
- It gets in the way when I'm looking for things. \_\_\_\_\_
- It's old. \_\_\_\_\_
- Somebody else has a copy. \_\_\_\_\_
- It's too much trouble to keep. \_\_\_\_\_
- I have no good way to save it. \_\_\_\_\_
- I'm emotionally troubled by it. \_\_\_\_\_
- It's not important to me. \_\_\_\_\_
- It takes up too much space. \_\_\_\_\_
- It's confidential / sensitive. \_\_\_\_\_
- To protect my privacy or someone else's privacy. \_\_\_\_\_
- The business doesn't need it any longer. \_\_\_\_\_

**48 [0048] On your work computer, which of the following helps you decide what to delete?**

This question is not asking whether you think this might be useful. Rather, this question is asking whether you actually, in deciding to delete something, use these to help you decide.

- The folder it's in.
- Its file type.
- Its file name.
- Its creation date.
- Its modified date.
- In email, the sender's name.
- In email, the subject.
- In email, whether the message is 'flagged'.
- In email, whether the message has attachments.
- The content of the item.
- Other: \_\_\_\_\_

**49 [0049]How would you feel if your workplace gave you a particular way to keep things organized, on your work computer?**

For example, if you were told that certain kinds of documents needed to go in particular folders.  
Please write your answer here:

**11/12: Information about your Work Position**

These are some general questions about your workplace and about the work you perform there.

**50 [0050] Do you work as a permanent employee?**

Please choose **only one** of the following:

- Yes  
 No

**51 [0051] How many hours do you work in an average week?**

Please write your answer here:

**52 [0052] How long have you worked for this particular company?**

Please choose **only one** of the following:

- Less than 6 months  
 6 months to 1 year  
 1 year or more, but less than 2 years  
 2 years or more, but less than 5 years  
 5 years or more, but less than 10 years  
 10 years or more

**53 [0053] How long have you done this same type of work?**

Please choose **only one** of the following:

- Less than 6 months  
 6 months to 1 year  
 1 year or more, but less than 2 years  
 2 years or more, but less than 5 years  
 5 years or more, but less than 10 years  
 10 years or more

Including the same or similar work in previous positions.

**54 [0054] What percentage of your work involves supervising others?**

Please choose **only one** of the following:

- None  
 1-25%  
 26-50%  
 51-75%  
 76-100%

**55 [0055] What is your occupation?**

It may be that you have several roles in your organization. Please provide as much information as you feel gives a good idea of what you do.

Please write your answer here:

**12/12: Finishing Up**

Just a few more questions and you're done!

**56 [0056]What is your gender?**

Please choose **only one** of the following:

- Female
- Male
- Transperson
- Decline to state

**57 [0057]What is your age?**

Please choose **only one** of the following:

- 0-14 Years
- 15-19 Years
- 20-24 Years
- 25-29 Years
- 30-34 Years
- 35-39 Years
- 40-44 Years
- 45-49 Years
- 50-54 Years
- 55-59 Years
- 60-64 Years
- 65-69 Years
- 70-74 Years
- 75-79 Years
- 80 and over

Submit your survey. Thank you for completing this survey.

## Appendix B

---

# Data Analysis Scripts

---

### Excel Macro for Data Cleanup

The following macro, written in Microsoft Visual Basic for Applications, performs a series of replacements within tabular data as exported from LimeSurvey. The purpose of these replacements is to achieve an uniformity of datatype: LimeSurvey, for example, exports skill levels as a mixture of numeric (codified) and textual data, all within a single column. For the purposes of data analysis, only the numeric portion of this was necessary or desired (for the purposes of database analysis, numeric data is much more rapidly compared than textual data, due to the type of database engine used for analysis).

```
Sub MassageData()  
Columns("L:L").Select  
Selection.Replace What:="1 - Expert", Replacement:="1", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="4 - Average", Replacement:="4", LookAt:=xlWhole _  
    , SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="7 - Beginner", Replacement:="7", LookAt:=xlWhole _  
    , SearchOrder:=xlByRows, MatchCase:=False  
Columns("M:S").Select  
Range("S1").Activate  
Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="N/A", Replacement:="2", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
ActiveWindow.SmallScroll ToRight:=31  
Columns("AO:BA").Select  
Range("BA1").Activate  
Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
Selection.Replace What:="Uncertain", Replacement:="3", LookAt:=xlWhole, _  
    SearchOrder:=xlByRows, MatchCase:=False  
ActiveWindow.SmallScroll ToRight:=14  
Columns("BC:BZ").Select  
Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _
```

```

        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    ActiveWindow.LargeScroll ToRight:=0
    ActiveWindow.SmallScroll ToRight:=6
    Columns("CC:CE").Select
    Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="N/A", Replacement:="2", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    ActiveWindow.SmallScroll ToRight:=8
    Columns("CF:CF").Select
    Selection.Replace What:="1 - Expert", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="4 - Average", Replacement:="4", LookAt:=xlWhole _
        , SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="7 - Beginner", Replacement:="7", LookAt:=xlWhole _
        , SearchOrder:=xlByRows, MatchCase:=False
    Columns("CG:EB").Select
    Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="N/A", Replacement:="2", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    ActiveWindow.LargeScroll ToRight:=2
    ActiveWindow.SmallScroll ToRight:=-3
    Columns("EX:GC").Select
    Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="N/A", Replacement:="2", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="Uncertain", Replacement:="3", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    ActiveWindow.SmallScroll ToRight:=14
    Columns("GM:GM").Select
    Selection.Replace What:="1 - Expert", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="4 - Average", Replacement:="4", LookAt:=xlWhole _
        , SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="7 - Beginner", Replacement:="7", LookAt:=xlWhole _
        , SearchOrder:=xlByRows, MatchCase:=False
    Columns("GN:IG").Select
    Selection.Replace What:="Yes", Replacement:="1", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="No", Replacement:="0", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="N/A", Replacement:="2", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False
    Selection.Replace What:="Uncertain", Replacement:="3", LookAt:=xlWhole, _
        SearchOrder:=xlByRows, MatchCase:=False

```

**End Sub**

## MSSQL Scripts

The following series of scripts are written within Microsoft's Transact-SQL language. Comments as to the purpose of the script or portion of script may be found within each script, contained within `/* ... */` series (where `/*` indicates the beginning of a region of comments, and `*/` denotes the end of such a region); such regions may span multiple lines. Single-line comments may also be found—or regions of code which have been removed from the executable script—denoted by a pair of dashes; for the purposes of clarity—in order that there be some record that a particular path was examined and discarded—such regions of inactive code have been left intact.

### Table Creation Scripts

The following script creates the basic tables required to store data for manipulation by the remaining MSSQL scripts. 'SurveyData\_in' is the initial repository for a data import which has been extracted from LimeSurvey and scrubbed by the Excel Macro (section B). 'SurveyInvalid' records individual surveys which were found to be invalid, along with the reason for their invalidity. 'SurveyTags' is for recording codes (tags) as determined during the qualitative analysis period; these codes, because they are the result of the researcher reading comments made by survey participants and of an unknown initial number, were kept in this table, allowing for many codes to relate to a single survey.

```
CREATE TABLE SurveyData_in
(
  [ID] INT NOT NULL
  , Completed VARCHAR (1) NULL
  , LastActivityDate SMALLDATETIME NULL
  , DateStarted SMALLDATETIME NULL
  , IPAddress VARCHAR (16) NULL
  , ReferringURL VARCHAR (256) NULL
  , HomeCountry VARCHAR (16) NULL
  , OtherHomeCountry VARCHAR (9) NULL
  , HomeState VARCHAR (14) NULL
  , HomeRegion VARCHAR (19) NULL
  , HomeProvince VARCHAR (16) NULL
  , OverallComputerSkill VARCHAR(12) NULL
  , UseComputerAtHome VARCHAR (9) NULL
  , HomeComputerShared VARCHAR (9) NULL
  , HomeAccountShared VARCHAR (9) NULL
  , HomeOSWindows VARCHAR (9) NULL
  , HomeOSMac10 VARCHAR (9) NULL
  , HomeOSMacOlder VARCHAR (9) NULL
  , HomeOSLinux VARCHAR (9) NULL
  , HomeOSOther VARCHAR (48) NULL
  , HomeCleanupEmail VARCHAR (28) NULL
  , HomeCleanupContacts VARCHAR (28) NULL
  , HomeCleanupDocuments VARCHAR (28) NULL
  , HomeCleanupSocialNetworkStatus VARCHAR (28) NULL
  , HomeCleanupTextMessages VARCHAR (28) NULL
  , HomeCleanupMobileContacts VARCHAR (28) NULL
  , HomeWatchVideo INT NULL
  , HomeConferenceCalling INT NULL
  , HomeSearchInternet INT NULL
  , HomeSocialNetworking INT NULL
  , HomeCreateNonWebContent INT NULL
  , HomeCreateWebContent INT NULL
)
```

, HomeEmail **INT NULL**  
, HomePlayGames **INT NULL**  
, HomeReadDistanceLearning **INT NULL**  
, HomePersonalFinance **INT NULL**  
, HomeOtherFinance **INT NULL**  
, HomePersonal **INT NULL**  
, HomeOtherPersonal **INT NULL**  
, HomeChatting **INT NULL**  
, HomeUseDefaultDocuments **VARCHAR (9) NULL**  
, HomeUseDefaultDownloads **VARCHAR (9) NULL**  
, HomeUseDefaultLibrary **VARCHAR (9) NULL**  
, HomeUseDefaultMovies **VARCHAR (9) NULL**  
, HomeUseDefaultMusic **VARCHAR (9) NULL**  
, HomeUseDefaultPictures **VARCHAR (9) NULL**  
, HomeFileEmail **VARCHAR (9) NULL**  
, HomeFileDocuments **VARCHAR (9) NULL**  
, HomeFileMusic **VARCHAR (9) NULL**  
, HomeFilePictures **VARCHAR (9) NULL**  
, HomeFileVideos **VARCHAR (9) NULL**  
, HomeFileRecords **VARCHAR (9) NULL**  
, HomeKeepMightNeed **VARCHAR (9) NULL**  
, HomeKeepMightNeedComment **VARCHAR (255) NULL**  
, HomeKeepImportantMemory **VARCHAR (9) NULL**  
, HomeKeepImportantMemoryComment **VARCHAR (255) NULL**  
, HomeKeepWantToWorkOn **VARCHAR (9) NULL**  
, HomeKeepWantToWorkOnComment **VARCHAR (255) NULL**  
, HomeKeepImportantToHousehold **VARCHAR (9) NULL**  
, HomeKeepImportantToHouseholdComment **VARCHAR (255) NULL**  
, HomeKeepSomebodyMightNeed **VARCHAR (9) NULL**  
, HomeKeepSomebodyMightNeedComment **VARCHAR (255) NULL**  
, HomeKeepSpentTimeOnIt **VARCHAR (9) NULL**  
, HomeKeepSpentTimeOnItComment **VARCHAR (255) NULL**  
, HomeKeepSomebodySpentTimeOnIt **VARCHAR (9) NULL**  
, HomeKeepSomebodySpentTimeOnItComment **VARCHAR (255) NULL**  
, HomeKeepLegalReasons **VARCHAR (9) NULL**  
, HomeKeepLegalReasonsComment **VARCHAR (255) NULL**  
, HomeKeepEmotionallyAttached **VARCHAR (9) NULL**  
, HomeKeepEmotionallyAttachedComment **VARCHAR (255) NULL**  
, HomeKeepInteresting **VARCHAR (9) NULL**  
, HomeKeepInterestingComment **VARCHAR (255) NULL**  
, HomeKeepTakesLittleSpace **VARCHAR (9) NULL**  
, HomeKeepTakesLittleSpaceComment **VARCHAR (255) NULL**  
, HomeKeepMyCreation **VARCHAR (9) NULL**  
, HomeKeepMyCreationComment **VARCHAR (255) NULL**  
, HomeKeepProveActions **VARCHAR (9) NULL**  
, HomeKeepProveActionsComment **VARCHAR (255) NULL**  
, HowManyComputersAtHome **INT NULL**  
, HowManyPeopleAtHome **INT NULL**  
, HomeComputersShareInternet **VARCHAR (9) NULL**  
, HomeComputersCopyFiles **VARCHAR (9) NULL**  
, HomeComputersUseNAS **VARCHAR (9) NULL**  
, HomeComparativeComputerSkill **VARCHAR(16) NULL**  
, HomeDeleteDontNeed **VARCHAR (9) NULL**  
, HomeDeleteDontNeedComment **VARCHAR (255) NULL**  
, HomeDeleteWrongToKeep **VARCHAR (9) NULL**  
, HomeDeleteWrongToKeepComment **VARCHAR (255) NULL**  
, HomeDeleteNobodyElseNeeds **VARCHAR (9) NULL**  
, HomeDeleteNobodyElseNeedsComment **VARCHAR (255) NULL**

, HomeDeleteGetsInWay **VARCHAR** (9) **NULL**  
 , HomeDeleteGetsInWayComment **VARCHAR** (255) **NULL**  
 , HomeDeleteItsOld **VARCHAR** (9) **NULL**  
 , HomeDeleteItsOldComment **VARCHAR** (255) **NULL**  
 , HomeDeleteSomebodyHasCopy **VARCHAR** (9) **NULL**  
 , HomeDeleteSomebodyHasCopyComment **VARCHAR** (255) **NULL**  
 , HomeDeleteTooMuchTroubleToKeep **VARCHAR** (9) **NULL**  
 , HomeDeleteTooMuchTroubleToKeepComment **VARCHAR** (255) **NULL**  
 , HomeDeleteNoWayToSave **VARCHAR** (9) **NULL**  
 , HomeDeleteNoWayToSaveComment **VARCHAR** (255) **NULL**  
 , HomeDeleteEmotionallyTroublesome **VARCHAR** (9) **NULL**  
 , HomeDeleteEmotionallyTroublesomeComment **VARCHAR** (255) **NULL**  
 , HomeDeleteNotImportantToMe **VARCHAR** (9) **NULL**  
 , HomeDeleteNotImportantToMeComment **VARCHAR** (255) **NULL**  
 , HomeDeleteTakesUpSpace **VARCHAR** (9) **NULL**  
 , HomeDeleteTakesUpSpaceComment **VARCHAR** (255) **NULL**  
 , HomeDeleteConfidential **VARCHAR** (9) **NULL**  
 , HomeDeleteConfidentialComment **VARCHAR** (255) **NULL**  
 , HomeDeletePrivacy **VARCHAR** (9) **NULL**  
 , HomeDeletePrivacyComment **VARCHAR** (255) **NULL**  
 , HomeHelpDeleteFolder **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteFileType **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteFileName **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteCreationDate **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteModifiedDate **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteEmailSenders **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteEmailSubject **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteEmailFlagged **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteEmailAttachments **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteContent **VARCHAR** (9) **NULL**  
 , HomeHelpDeleteOtherComment **VARCHAR** (255) **NULL**  
 , HomeKeepWorkObjects **VARCHAR** (9) **NULL**  
 , HomeKeepWorkObjectsAfterLeavingCompany **VARCHAR** (9) **NULL**  
 , HomeKeepWorkObjectsAfterLeavingCompanyShouldKeep **VARCHAR** (9) **NULL**  
 , HomeKeepWorkObjectsAfterLeavingCompanyShouldKeepComment **VARCHAR** (512) **NULL**  
 , HomeKeepWorkObjectsAfterLeavingCompanyShouldntKeepComment **VARCHAR** (255) **NULL**  
 , WorkKeepHomeObjects **VARCHAR** (9) **NULL**  
 , WorkKeepHomeObjectsUpsetIfReviewed **VARCHAR** (9) **NULL**  
 , WorkKeepHomeObjectsUpsetIfReviewedComment **VARCHAR** (1024) **NULL**  
 , UseComputerAtWork **VARCHAR** (9) **NULL**  
 , ShareWorkComputer **VARCHAR** (9) **NULL**  
 , WorkAccountShared **VARCHAR** (9) **NULL**  
 , JobSkillAcquiredOJT **INT** **NULL**  
 , JobSkillAcquiredOnOwnAtWork **INT** **NULL**  
 , JobSkillAcquiredOnOwnOutsideWork **INT** **NULL**  
 , JobSkillAcquiredFormallyDuringWorkTime **INT** **NULL**  
 , JobSkillAcquiredFormallyOutsideWork **INT** **NULL**  
 , JobSkillAcquiredSeminarsConferences **INT** **NULL**  
 , JobSkillAcquiredOther **INT** **NULL**  
 , WorkWatchVideo **INT** **NULL**  
 , WorkConferenceCalling **INT** **NULL**  
 , WorkSearchInternet **INT** **NULL**  
 , WorkSocialNetworking **INT** **NULL**  
 , WorkCreateNonWebContent **INT** **NULL**  
 , WorkCreateWebContent **INT** **NULL**  
 , WorkEmail **INT** **NULL**  
 , WorkPlayGames **INT** **NULL**  
 , WorkReadDistanceLearning **INT** **NULL**

, WorkPersonalFinance INT NULL  
, WorkOtherFinance INT NULL  
, WorkPersonal INT NULL  
, WorkOtherPersonal INT NULL  
, WorkChatting INT NULL  
, WorkKeepMightNeed VARCHAR (9) NULL  
, WorkKeepMightNeedComment VARCHAR (255) NULL  
, WorkKeepImportantMemory VARCHAR (9) NULL  
, WorkKeepImportantMemoryComment VARCHAR (255) NULL  
, WorkKeepWantToWorkOn VARCHAR (9) NULL  
, WorkKeepWantToWorkOnComment VARCHAR (255) NULL  
, WorkKeepImportantToBusiness VARCHAR (9) NULL  
, WorkKeepImportantToBusinessComment VARCHAR (255) NULL  
, WorkKeepSomebodyMightNeed VARCHAR (9) NULL  
, WorkKeepSomebodyMightNeedComment VARCHAR (255) NULL  
, WorkKeepSpentTimeOnIt VARCHAR (9) NULL  
, WorkKeepSpentTimeOnItComment VARCHAR (255) NULL  
, WorkKeepSomebodySpentTimeOnIt VARCHAR (9) NULL  
, WorkKeepSomebodySpentTimeOnItComment VARCHAR (255) NULL  
, WorkKeepLegalReasons VARCHAR (9) NULL  
, WorkKeepLegalReasonsComment VARCHAR (255) NULL  
, WorkKeepEmotionallyAttached VARCHAR (9) NULL  
, WorkKeepEmotionallyAttachedComment VARCHAR (255) NULL  
, WorkKeepInteresting VARCHAR (9) NULL  
, WorkKeepInterestingComment VARCHAR (255) NULL  
, WorkKeepTakesLittleSpace VARCHAR (9) NULL  
, WorkKeepTakesLittleSpaceComment VARCHAR (255) NULL  
, WorkKeepMyCreation VARCHAR (9) NULL  
, WorkKeepMyCreationComment VARCHAR (255) NULL  
, WorkKeepProveActions VARCHAR (9) NULL  
, WorkKeepProveActionsComment VARCHAR (255) NULL  
, WorkEmailProvided VARCHAR (9) NULL  
, WorkEmailSaveCopy VARCHAR (9) NULL  
, WorkEmailFileImmediately VARCHAR (9) NULL  
, WorkEmailAutoArchive VARCHAR (9) NULL  
, WorkEmailMoveToOtherSystem VARCHAR (9) NULL  
, WorkCentrallyManagedDocuments VARCHAR (9) NULL  
, WorkCentrallyManagedDocumentsKeepCopy VARCHAR (9) NULL  
, WorkCentrallyManagedDocumentsKeepCopyComment VARCHAR (255) NULL  
, WorkWhyOrganizeDocuments VARCHAR(1088) NULL  
, WorkHowManyPeopleInCompany VARCHAR (9) NULL  
, WorkHowManyPeopleInDepartment VARCHAR (9) NULL  
, WorkHowManyPeoplePerformSameFunctionInCompany VARCHAR (9) NULL  
, WorkHowManyPeoplePerformSameFunctionInDepartment VARCHAR (9) NULL  
, WorkPrimaryIndustry VARCHAR (9) NULL  
, WorkPrimaryIndustryOther VARCHAR (9) NULL  
, WorkComparativeComputerSkill INT NULL  
, WorkHasRM VARCHAR (9) NULL  
, WorkHasRMPolicies VARCHAR (9) NULL  
, WorkInstructedOnRM VARCHAR (9) NULL  
, WorkAllEmpsInstructedOnRM VARCHAR (9) NULL  
, WorkAuditsDocs VARCHAR (9) NULL  
, WorkDeleteDontNeed VARCHAR (9) NULL  
, WorkDeleteDontNeedComment VARCHAR (255) NULL  
, WorkDeleteWrongToKeep VARCHAR (9) NULL  
, WorkDeleteWrongToKeepComment VARCHAR (255) NULL  
, WorkDeleteNobodyElseNeeds VARCHAR (9) NULL  
, WorkDeleteNobodyElseNeedsComment VARCHAR (255) NULL

```

, WorkDeleteGetsInWay VARCHAR (9) NULL
, WorkDeleteGetsInWayComment VARCHAR (255) NULL
, WorkDeleteItsOld VARCHAR (9) NULL
, WorkDeleteItsOldComment VARCHAR (255) NULL
, WorkDeleteSomebodyHasCopy VARCHAR (9) NULL
, WorkDeleteSomebodyHasCopyComment VARCHAR (255) NULL
, WorkDeleteTooMuchTroubleToKeep VARCHAR (9) NULL
, WorkDeleteTooMuchTroubleToKeepComment VARCHAR (255) NULL
, WorkDeleteNoWayToSave VARCHAR (9) NULL
, WorkDeleteNoWayToSaveComment VARCHAR (255) NULL
, WorkDeleteEmotionallyTroublesome VARCHAR (9) NULL
, WorkDeleteEmotionallyTroublesomeComment VARCHAR (255) NULL
, WorkDeleteNotImportantToMe VARCHAR (9) NULL
, WorkDeleteNotImportantToMeComment VARCHAR (255) NULL
, WorkDeleteTakesUpSpace VARCHAR (9) NULL
, WorkDeleteTakesUpSpaceComment VARCHAR (255) NULL
, WorkDeleteConfidential VARCHAR (9) NULL
, WorkDeleteConfidentialComment VARCHAR (255) NULL
, WorkDeletePrivacy VARCHAR (9) NULL
, WorkDeletePrivacyComment VARCHAR (255) NULL
, WorkDeleteBusinessDoesntNeed VARCHAR (9) NULL
, WorkDeleteBusinessDoesntNeedComment VARCHAR (255) NULL
, WorkHelpDeleteFolder VARCHAR (9) NULL
, WorkHelpDeleteFileType VARCHAR (9) NULL
, WorkHelpDeleteFileName VARCHAR (9) NULL
, WorkHelpDeleteCreationDate VARCHAR (9) NULL
, WorkHelpDeleteModifiedDate VARCHAR (9) NULL
, WorkHelpDeleteEmailSenders VARCHAR (9) NULL
, WorkHelpDeleteEmailSubject VARCHAR (9) NULL
, WorkHelpDeleteEmailFlagged VARCHAR (9) NULL
, WorkHelpDeleteEmailAttachments VARCHAR (9) NULL
, WorkHelpDeleteContent VARCHAR (9) NULL
, WorkHelpDeleteOtherComment VARCHAR (255) NULL
, WorkHowFeelIfGivenOrganizationSystem VARCHAR (480) NULL
, WorkAsPermanentEmployee VARCHAR (9) NULL
, WorkHoursWorkedPerWeek INT NULL
, WorkHowLongAtCurrentCompany VARCHAR (16) NULL
, WorkHowLongAtTypeOfWork VARCHAR (16) NULL
, WorkPercentTimeSupervising VARCHAR (9) NULL
, WorkOccupation VARCHAR(560) NULL
, Gender VARCHAR (9) NULL
, Age VARCHAR (9) NULL
, FirstName VARCHAR(17) NULL
, LastName VARCHAR(20) NULL
, EmailAddress VARCHAR(40) NULL
, Token VARCHAR(16) NULL
)

```

GO

```

ALTER TABLE SurveyData_in ADD CONSTRAINT uci_ID@SurveyData_in
UNIQUE CLUSTERED ( [ID] )

```

GO

```

CREATE TABLE dbo.SurveyInvalid
(
  [ID] INT NOT NULL

```

```

, ReasonInvalid VARCHAR (256) NOT NULL
)
GO

CREATE TABLE dbo.SurveyTags
(
[ID] INT NOT NULL
, TagAs VARCHAR (128) NOT NULL
)
GO

CREATE INDEX IX_SurveyTags ON dbo.SurveyTags ([ID])

```

## Separating Quantitative from Qualitative Questions

Variables which are subject to quantitative analysis were considered separately from those subject to qualitative analysis—e.g., textual comments required coding by the researcher, so are considered qualitative for this purpose. Because of the quantity of questions allowing for comment, separating the qualitative / textual responses allowed for easier coding of these responses.

```

IF EXISTS ( SELECT * FROM SYSOBJECTS
WHERE ID = OBJECT_ID(N'SurveyQuantitative'))
    DROP TABLE SurveyQuantitative
GO
SELECT
    [ID],
    case when HomeCountry is null then OtherHomeCountry
         else HomeCountry end as HomeCountry,
    case
        when HomeState is not null then HomeState
        when HomeRegion is not null then HomeRegion
        when HomeProvince is not null then HomeProvince
        when HomeCountry = 'Other' then OtherHomeCountry
        end as HomeState,
    OverallComputerSkill ,
    UseComputerAtHome ,
    HomeComputerShared ,
    HomeAccountShared ,
    HomeOSWindows ,
    HomeOSMac10 ,
    HomeOSMacOlder ,
    HomeOSLinux ,
    HomeOSOther ,
    HomeCleanupEmail ,
    HomeCleanupContacts ,
    HomeCleanupDocuments ,
    HomeCleanupSocialNetworkStatus ,
    HomeCleanupTextMessages ,
    HomeCleanupMobileContacts ,
    HomeWatchVideo ,
    HomeConferenceCalling ,
    HomeSearchInternet ,
    HomeSocialNetworking ,
    HomeCreateNonWebContent ,
    HomeCreateWebContent ,

```

HomeEmail ,  
HomePlayGames ,  
HomeReadDistanceLearning ,  
HomePersonalFinance ,  
HomeOtherFinance ,  
HomePersonal ,  
HomeOtherPersonal ,  
HomeChatting ,  
HomeUseDefaultDocuments ,  
HomeUseDefaultDownloads ,  
HomeUseDefaultLibrary ,  
HomeUseDefaultMovies ,  
HomeUseDefaultMusic ,  
HomeUseDefaultPictures ,  
HomeFileEmail ,  
HomeFileDocuments ,  
HomeFileMusic ,  
HomeFilePictures ,  
HomeFileVideos ,  
HomeFileRecords ,  
HomeKeepMightNeed ,  
HomeKeepImportantMemory ,  
HomeKeepWantToWorkOn ,  
HomeKeepImportantToHousehold ,  
HomeKeepSomebodyMightNeed ,  
HomeKeepSpentTimeOnIt ,  
HomeKeepSomebodySpentTimeOnIt ,  
HomeKeepLegalReasons ,  
HomeKeepEmotionallyAttached ,  
HomeKeepInteresting ,  
HomeKeepTakesLittleSpace ,  
HomeKeepMyCreation ,  
HomeKeepProveActions ,  
HowManyComputersAtHome ,  
HowManyPeopleAtHome ,  
HomeComputersShareInternet ,  
HomeComputersCopyFiles ,  
HomeComputersUseNAS ,  
HomeComparativeComputerSkill ,  
HomeDeleteDontNeed ,  
HomeDeleteWrongToKeep ,  
HomeDeleteNobodyElseNeeds ,  
HomeDeleteGetsInWay ,  
HomeDeleteItsOld ,  
HomeDeleteSomebodyHasCopy ,  
HomeDeleteTooMuchTroubleToKeep ,  
HomeDeleteNoWayToSave ,  
HomeDeleteEmotionallyTroublesome ,  
HomeDeleteNotImportantToMe ,  
HomeDeleteTakesUpSpace ,  
HomeDeleteConfidential ,  
HomeDeletePrivacy ,  
HomeHelpDeleteFolder ,  
HomeHelpDeleteFileType ,  
HomeHelpDeleteFileName ,  
HomeHelpDeleteCreationDate ,  
HomeHelpDeleteModifiedDate ,  
HomeHelpDeleteEmailSenders ,

HomeHelpDeleteEmailSubject ,  
HomeHelpDeleteEmailFlagged ,  
HomeHelpDeleteEmailAttachments ,  
HomeHelpDeleteContent ,  
HomeKeepWorkObjects ,  
HomeKeepWorkObjectsAfterLeavingCompany ,  
HomeKeepWorkObjectsAfterLeavingCompanyShouldKeep ,  
WorkKeepHomeObjects ,  
WorkKeepHomeObjectsUpsetIfReviewed ,  
UseComputerAtWork ,  
ShareWorkComputer ,  
WorkAccountShared ,  
JobSkillAcquiredOJT ,  
JobSkillAcquiredOnOwnAtWork ,  
JobSkillAcquiredOnOwnOutsideWork ,  
JobSkillAcquiredFormallyDuringWorkTime ,  
JobSkillAcquiredFormallyOutsideWork ,  
JobSkillAcquiredSeminarsConferences ,  
JobSkillAcquiredOther ,  
WorkWatchVideo ,  
WorkConferenceCalling ,  
WorkSearchInternet ,  
WorkSocialNetworking ,  
WorkCreateNonWebContent ,  
WorkCreateWebContent ,  
WorkEmail ,  
WorkPlayGames ,  
WorkReadDistanceLearning ,  
WorkPersonalFinance ,  
WorkOtherFinance ,  
WorkPersonal ,  
WorkOtherPersonal ,  
WorkChatting ,  
WorkKeepMightNeed ,  
WorkKeepImportantMemory ,  
WorkKeepWantToWorkOn ,  
WorkKeepImportantToBusiness ,  
WorkKeepSomebodyMightNeed ,  
WorkKeepSpentTimeOnIt ,  
WorkKeepSomebodySpentTimeOnIt ,  
WorkKeepLegalReasons ,  
WorkKeepEmotionallyAttached ,  
WorkKeepInteresting ,  
WorkKeepTakesLittleSpace ,  
WorkKeepMyCreation ,  
WorkKeepProveActions ,  
WorkEmailProvided ,  
WorkEmailSaveCopy ,  
WorkEmailFileImmediately ,  
WorkEmailAutoArchive ,  
WorkEmailMoveToOtherSystem ,  
WorkCentrallyManagedDocuments ,  
WorkCentrallyManagedDocumentsKeepCopy ,  
WorkHowManyPeopleInCompany ,  
WorkHowManyPeopleInDepartment ,  
WorkHowManyPeoplePerformSameFunctionInCompany ,  
WorkHowManyPeoplePerformSameFunctionInDepartment ,  
WorkPrimaryIndustry ,

```

WorkPrimaryIndustryOther ,
WorkComparativeComputerSkill ,
WorkHasRM,
WorkHasRMPolicies ,
WorkInstructedOnRM ,
WorkAllEmpsInstructedOnRM ,
WorkAuditsDocs ,
WorkDeleteDontNeed ,
WorkDeleteWrongToKeep ,
WorkDeleteNobodyElseNeeds ,
WorkDeleteGetsInWay ,
WorkDeleteItsOld ,
WorkDeleteSomebodyHasCopy ,
WorkDeleteTooMuchTroubleToKeep ,
WorkDeleteNoWayToSave ,
WorkDeleteEmotionallyTroublesome ,
WorkDeleteNotImportantToMe ,
WorkDeleteTakesUpSpace ,
WorkDeleteConfidential ,
WorkDeletePrivacy ,
WorkDeleteBusinessDoesntNeed ,
WorkHelpDeleteFolder ,
WorkHelpDeleteFileType ,
WorkHelpDeleteFileName ,
WorkHelpDeleteCreationDate ,
WorkHelpDeleteModifiedDate ,
WorkHelpDeleteEmailSenders ,
WorkHelpDeleteEmailSubject ,
WorkHelpDeleteEmailFlagged ,
WorkHelpDeleteEmailAttachments ,
WorkHelpDeleteContent ,
WorkAsPermanentEmployee ,
WorkHoursWorkedPerWeek ,
WorkHowLongAtCurrentCompany ,
WorkHowLongAtTypeOfWork ,
WorkPercentTimeSupervising ,
Gender ,
Age
INTO SurveyQuantitative
FROM SurveyData_in
WHERE
Completed = 'Y'
GO

IF EXISTS ( SELECT * FROM SYSOBJECTS
WHERE ID = OBJECT_ID(N'SurveyQualitative'))
DROP TABLE SurveyQualitative
GO
SELECT
[ID] ,
HomeKeepMightNeedComment ,
HomeKeepImportantMemoryComment ,
HomeKeepWantToWorkOnComment ,
HomeKeepImportantToHouseholdComment ,
HomeKeepSomebodyMightNeedComment ,
HomeKeepSpentTimeOnItComment ,
HomeKeepSomebodySpentTimeOnItComment ,
HomeKeepLegalReasonsComment ,

```

```

HomeKeepEmotionallyAttachedComment ,
HomeKeepInterestingComment ,
HomeKeepTakesLittleSpaceComment ,
HomeKeepMyCreationComment ,
HomeKeepProveActionsComment ,
HomeDeleteDontNeedComment ,
HomeDeleteWrongToKeepComment ,
HomeDeleteNobodyElseNeedsComment ,
HomeDeleteGetsInWayComment ,
HomeDeleteItsOldComment ,
HomeDeleteSomebodyHasCopyComment ,
HomeDeleteTooMuchTroubleToKeepComment ,
HomeDeleteNoWayToSaveComment ,
HomeDeleteEmotionallyTroublesomeComment ,
HomeDeleteNotImportantToMeComment ,
HomeDeleteTakesUpSpaceComment ,
HomeDeleteConfidentialComment ,
HomeDeletePrivacyComment ,
HomeHelpDeleteOtherComment ,
HomeKeepWorkObjectsAfterLeavingCompanyShouldKeepComment ,
HomeKeepWorkObjectsAfterLeavingCompanyShouldntKeepComment ,
WorkKeepHomeObjectsUpsetIfReviewedComment ,
WorkKeepMightNeedComment ,
WorkKeepImportantMemoryComment ,
WorkKeepWantToWorkOnComment ,
WorkKeepImportantToBusinessComment ,
WorkKeepSomebodyMightNeedComment ,
WorkKeepSpentTimeOnItComment ,
WorkKeepSomebodySpentTimeOnItComment ,
WorkKeepLegalReasonsComment ,
WorkKeepEmotionallyAttachedComment ,
WorkKeepInterestingComment ,
WorkKeepTakesLittleSpaceComment ,
WorkKeepMyCreationComment ,
WorkKeepProveActionsComment ,
WorkCentrallyManagedDocumentsKeepCopyComment ,
WorkWhyOrganizeDocuments ,
WorkDeleteDontNeedComment ,
WorkDeleteWrongToKeepComment ,
WorkDeleteNobodyElseNeedsComment ,
WorkDeleteGetsInWayComment ,
WorkDeleteItsOldComment ,
WorkDeleteSomebodyHasCopyComment ,
WorkDeleteTooMuchTroubleToKeepComment ,
WorkDeleteNoWayToSaveComment ,
WorkDeleteEmotionallyTroublesomeComment ,
WorkDeleteNotImportantToMeComment ,
WorkDeleteTakesUpSpaceComment ,
WorkDeleteConfidentialComment ,
WorkDeletePrivacyComment ,
WorkDeleteBusinessDoesntNeedComment ,
WorkHelpDeleteOtherComment ,
WorkHowFeelIfGivenOrganizationSystem ,
WorkOccupation
INTO SurveyQualitative
FROM SurveyData_in
WHERE
Completed = 'Y'

```

GO

```
ALTER TABLE SurveyQuantitative
ADD CONSTRAINT PKC_SurveyQuantitative
PRIMARY KEY CLUSTERED ([ID])
```

```
ALTER TABLE SurveyQualitative
ADD CONSTRAINT PKC_SurveyQualitative
PRIMARY KEY CLUSTERED ([ID])
```

## Denormalised Code Structure

After coding of responses was complete, these codes needed to be incorporated into the survey data in such a manner as to allow for them to be considered as variables alongside the survey response variables. Codes were kept in a table which allowed for a survey ID to have multiple Code records—the codes were somewhat *normal*, in the database sense of the word. In order to analyse them alongside survey response variables, codes were denormalised—they were flattened into single rows corresponding to each survey ID, along with multiple columns of data, each column representing a code value, along with a value denoting whether that code was applied.

```
create table SurveyTagsFlat(
[ID] int not null ,
    constraint pk_surveytagsflat
    primary key clustered ([id]) ,
Administrative bit not null default 0,
Archivist bit not null default 0,
Aware_Of_Legal_Retention_Period_In_Deleting bit not null default 0,
Busines_Person bit not null default 0,
Casual_Attitude bit not null default 0,
Civil_Servant bit not null default 0,
Curator bit not null default 0,
Does_Not_Believe_Digital_Can_Be_Owned bit not null default 0,
Does_Not_DeleteAnything bit not null default 0,
Ecologist bit not null default 0,
Educator bit not null default 0,
Encrypts_Private_Files_At_Work bit not null default 0,
Impossible_NOT_to_do_Personal_Things_At_Work bit not null default 0,
Journalist bit not null default 0,
Justifies_Because_IP_Not_Valuable bit not null default 0,
Keeps_Copy_Of_Centrally_Managed_Documents bit not null default 0,
Keeps_For_Reference bit not null default 0,
Keeps_Personal_On_Work_Computer bit not null default 0,
Keeps_To_CYA bit not null default 0,
Keeps_Work_IP bit not null default 0,
Keeps_Work_Materials bit not null default 0,
Lawyer bit not null default 0,
Librarian bit not null default 0,
Machiavellian bit not null default 0,
Marketing bit not null default 0,
Medical_Professional bit not null default 0,
Mixes_Home_And_Work bit not null default 0,
Nostalgia bit not null default 0,
Organizational_System_Not_Receptive bit not null default 0,
Organizational_System_Qualified_Acceptance bit not null default 0,
Organizational_System_Receptive bit not null default 0,
Organize_Because_Documents_Important bit not null default 0,
```

```

Organize_Because_Organization_Important bit not null default 0,
Organize_For_Deletion bit not null default 0,
Organize_For_Efficiency bit not null default 0,
Organize_For_Finding bit not null default 0,
Organize_For_Posterity bit not null default 0,
Personal_Files_On_Separate_Drive bit not null default 0,
Privacy_Concerns bit not null default 0,
Project_Manager bit not null default 0,
Protects_Home_From_Work bit not null default 0,
Protects_Work_From_Home bit not null default 0,
Records_Manager bit not null default 0,
Recruiter bit not null default 0,
Regards_Work_IP_As_Their_Property bit not null default 0,
Secretarial bit not null default 0,
Social_Work bit not null default 0,
Student bit not null default 0,
Technologist bit not null default 0,
Too_Much_Trouble_To_Separate_Home_From_Work bit not null default 0,
Unjustified bit not null default 0,
Upset_If_Personal_Files_Examined_At_Work bit not null default 0
)

```

## View for High-Level Quantitative Analysis

The following view of data is a database query, joining a few different pieces of information together to be examined. It accomplishes this unification of data and also provides some codification in the process. This view will be used to facilitate examination of the data, comparing different variables to one another.

```

CREATE VIEW vwQuantitativeHL
AS

```

```

SELECT
    SurveyQuantitative.ID
    ,SurveyQuantitative.HomeCountry
    ,SurveyQuantitative.HomeState
    ,SurveyQuantitative.HomeComputerShared
    ,SurveyQuantitative.HomeAccountShared
    ,SurveyQuantitative.HomeOSWindows
    ,SurveyQuantitative.HomeComputersShareInternet
    ,SurveyQuantitative.HomeComputersCopyFiles
    ,SurveyQuantitative.HomeComputersUseNAS
    ,SurveyQuantitative.ShareWorkComputer
    ,SurveyQuantitative.WorkAccountShared
    ,SurveyQuantitative.WorkAsPermanentEmployee
    ,SurveyQuantitative.WorkHowLongAtCurrentCompany
    ,SurveyQuantitative.WorkHowLongAtTypeOfWork
    ,SurveyQuantitative.WorkPercentTimeSupervising
    ,SurveyQuantitative.Gender
    ,SurveyQuantitative.Age
    , Age2 = CASE
        WHEN SurveyQuantitative.Age = '15-19 Yea' THEN '15-29 Years'
        WHEN SurveyQuantitative.Age = '20-24 Yea' THEN '15-29 Years'
        WHEN SurveyQuantitative.Age = '25-29 Yea' THEN '15-29 Years'
        WHEN SurveyQuantitative.Age = '30-34 Yea' THEN '30-49 Years'
        WHEN SurveyQuantitative.Age = '35-39 Yea' THEN '30-49 Years'
        WHEN SurveyQuantitative.Age = '40-44 Yea' THEN '30-49 Years'

```

```

        WHEN SurveyQuantitative.Age = '45-49 Yea' THEN '30-49 Years'
        WHEN SurveyQuantitative.Age = '50-54 Yea' THEN '50-74 Years'
        WHEN SurveyQuantitative.Age = '55-59 Yea' THEN '50-74 Years'
        WHEN SurveyQuantitative.Age = '60-64 Yea' THEN '50-74 Years'
        WHEN SurveyQuantitative.Age = '65-69 Yea' THEN '50-74 Years'
        WHEN SurveyQuantitative.Age = '70-74 Yea' THEN '50-74 Years'
    END
    , Age3 = CASE
        WHEN SurveyQuantitative.Age = '15-19 Yea' THEN '15-34 Years'
        WHEN SurveyQuantitative.Age = '20-24 Yea' THEN '15-34 Years'
        WHEN SurveyQuantitative.Age = '25-29 Yea' THEN '15-34 Years'
        WHEN SurveyQuantitative.Age = '30-34 Yea' THEN '15-34 Years'
        WHEN SurveyQuantitative.Age = '35-39 Yea' THEN '35-54 Years'
        WHEN SurveyQuantitative.Age = '40-44 Yea' THEN '35-54 Years'
        WHEN SurveyQuantitative.Age = '45-49 Yea' THEN '35-54 Years'
        WHEN SurveyQuantitative.Age = '50-54 Yea' THEN '35-54 Years'
        WHEN SurveyQuantitative.Age = '55-59 Yea' THEN '55-74 Years'
        WHEN SurveyQuantitative.Age = '60-64 Yea' THEN '55-74 Years'
        WHEN SurveyQuantitative.Age = '65-69 Yea' THEN '55-74 Years'
        WHEN SurveyQuantitative.Age = '70-74 Yea' THEN '55-74 Years'
    END
    , Profession = CASE
        WHEN SurveyTagsFlat.Archivist = 1 THEN 'Archivist'
        WHEN SurveyTagsFlat.Busines_Person = 1 THEN 'Business Person'
        WHEN SurveyTagsFlat.Civil_Servant = 1 THEN 'Civil Servant'
        WHEN SurveyTagsFlat.Curator = 1 THEN 'Curator'
        WHEN SurveyTagsFlat.Ecologist = 1 THEN 'Ecologist'
        WHEN SurveyTagsFlat.Educator = 1 THEN 'Educator'
        WHEN SurveyTagsFlat.Journalist = 1 THEN 'Journalist'
        WHEN SurveyTagsFlat.Lawyer = 1 THEN 'Lawyer'
        WHEN SurveyTagsFlat.Librarian = 1 THEN 'Librarian'
        WHEN SurveyTagsFlat.Marketing = 1 THEN 'Marketing'
        WHEN SurveyTagsFlat.Medical_Professional = 1 THEN 'Medical
            Professional'
        WHEN SurveyTagsFlat.Project_Manager = 1 THEN 'Project Manager'
        WHEN SurveyTagsFlat.Records_Manager = 1 THEN 'Records Manager'
        WHEN SurveyTagsFlat.Recruiter = 1 THEN 'Recruiter'
        WHEN SurveyTagsFlat.Administrative = 1 THEN 'Administrative'
        WHEN SurveyTagsFlat.Secretarial = 1 THEN 'Secretarial'
        WHEN SurveyTagsFlat.Social_Work = 1 THEN 'Social Work'
        WHEN SurveyTagsFlat.Student = 1 THEN 'Student'
        WHEN SurveyTagsFlat.Technologist = 1 THEN 'Technologist'
        ELSE 'Other'
    END
END
/*
Following are 'Effect' variables.
*/
,[Cleanup Home] = (
CASE
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Every day' THEN 6
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Once a week' THEN 5
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Once a month' THEN 4
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Several times a year' THEN 3
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Every year' THEN 2
    WHEN SurveyQuantitative.HomeCleanupEmail = 'Hardly ever' THEN 1
    ELSE 0
END
+ CASE

```

```

    WHEN SurveyQuantitative.HomeCleanupContacts = 'Every day' THEN 6
    WHEN SurveyQuantitative.HomeCleanupContacts = 'Once a week' THEN 5
    WHEN SurveyQuantitative.HomeCleanupContacts = 'Once a month' THEN 4
    WHEN SurveyQuantitative.HomeCleanupContacts = 'Several times a year' THEN
        3
    WHEN SurveyQuantitative.HomeCleanupContacts = 'Every year' THEN 2
    WHEN SurveyQuantitative.HomeCleanupContacts = 'Hardly ever' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Every day' THEN 6
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Once a week' THEN 5
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Once a month' THEN 4
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Several times a year'
        THEN 3
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Every year' THEN 2
    WHEN SurveyQuantitative.HomeCleanupDocuments = 'Hardly ever' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Every day' THEN
        6
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Once a week'
        THEN 5
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Once a month'
        THEN 4
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Several times a
        year' THEN 3
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Every year'
        THEN 2
    WHEN SurveyQuantitative.HomeCleanupSocialNetworkStatus = 'Hardly ever'
        THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Every day' THEN 6
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Once a week' THEN 5
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Once a month' THEN 4
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Several times a year'
        THEN 3
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Every year' THEN 2
    WHEN SurveyQuantitative.HomeCleanupTextMessages = 'Hardly ever' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Every day' THEN 6
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Once a week' THEN 5
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Once a month' THEN 4
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Several times a year'
        THEN 3
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Every year' THEN 2
    WHEN SurveyQuantitative.HomeCleanupMobileContacts = 'Hardly ever' THEN 1
    ELSE 0
END
)
,[Deletion Affordances Home] = (
CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFolder = 'Yes' THEN 1

```

```

        ELSE 0
    END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFileType = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFileName = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteCreationDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteModifiedDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailSenders = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailSubject = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailFlagged = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailAttachments = 'Yes' THEN 1
    ELSE 0
END
)
,[Deletion Affordances Home 2] = (
CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFolder = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFileType = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteFileName = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteCreationDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteModifiedDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailSenders = 'Yes' THEN 1

```

```

        ELSE 0
    END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailSubject = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailFlagged = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteEmailAttachments = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeHelpDeleteContent = 'Yes' THEN 1
    ELSE 0
END
)
,[Deletion Affordances Work] = (
CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFolder = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFileType = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFileName = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteCreationDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteModifiedDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailSenders = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailSubject = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailFlagged = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailAttachments = 'Yes' THEN 1
    ELSE 0
END
)
,[Deletion Affordances Work 2] = (

```

```

CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFolder = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFileType = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteFileName = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteCreationDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteModifiedDate = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailSenders = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailSubject = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailFlagged = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteEmailAttachments = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHelpDeleteContent = 'Yes' THEN 1
    ELSE 0
END
)
,[Email Management Activity Work] = (
CASE
    WHEN SurveyQuantitative.WorkEmailProvided = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkEmailSaveCopy = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkEmailFileImmediately = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkEmailMoveToOtherSystem = 'Yes' THEN 1
    ELSE 0
END
)

```

```

)
,[Filing Activities Home] = (
CASE
    WHEN SurveyQuantitative.HomeFileEmail = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeFileDocuments = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeFileMusic = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeFilePictures = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeFileVideos = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeFileRecords = 'Yes' THEN 1
    ELSE 0
END
)
,[Functional Reasons Home] = (
CASE
    WHEN SurveyQuantitative.HomeDeleteGetsInWay = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteItsOld = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteSomebodyHasCopy = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteTooMuchTroubleToKeep = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteNoWayToSave = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteTakesUpSpace = 'Yes' THEN 1
    ELSE 0
END
)
,[Functional Reasons Work] = (
CASE
    WHEN SurveyQuantitative.WorkDeleteGetsInWay = 'Yes' THEN 1
    ELSE 0
END

```

```

+ CASE
  WHEN SurveyQuantitative.WorkDeleteItsOld = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkDeleteSomebodyHasCopy = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkDeleteTooMuchTroubleToKeep = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkDeleteNoWayToSave = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkDeleteTakesUpSpace = 'Yes' THEN 1
  ELSE 0
END
)
,[Home Work Mixing] = (
CASE
  WHEN SurveyQuantitative.HomeKeepWorkObjects = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Very
    Like' THEN 1
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Likely'
    THEN 1
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = '
    Somewhat ' THEN 1
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Perhaps
    ' THEN 0
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Maybe
    Not' THEN -1
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Most
    Like' THEN -1
  WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Not At
    All' THEN -1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkKeepHomeObjects = 'Yes' THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkPersonalFinance > 0 THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkPersonal > 0 THEN 1
  ELSE 0
END
+ CASE
  WHEN SurveyQuantitative.WorkOtherPersonal > 0 THEN 1
  ELSE 0

```

```

END
)
,[Personal Reasons Home] = (
CASE
    WHEN SurveyQuantitative.HomeKeepMightNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepImportantMemory = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepWantToWorkOn = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepSpentTimeOnIt = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepSomebodySpentTimeOnIt = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepLegalReasons = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepEmotionallyAttached = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepInteresting = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepMyCreation = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepProveActions = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteDontNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteWrongToKeep = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteEmotionallyTroublesome = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteNotImportantToMe = 'Yes' THEN 1
    ELSE 0

```

```

END
+ CASE
    WHEN SurveyQuantitative.HomeDeletePrivacy = 'Yes' THEN 1
    ELSE 0
END
)
,[Personal Reasons Work] = (
CASE
    WHEN SurveyQuantitative.WorkKeepMightNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepImportantMemory = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepWantToWorkOn = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepSpentTimeOnIt = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepSomebodySpentTimeOnIt = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepLegalReasons = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepEmotionallyAttached = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepInteresting = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepMyCreation = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepProveActions = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteDontNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteWrongToKeep = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteEmotionallyTroublesome = 'Yes' THEN 1
    ELSE 0

```

```

END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteNotImportantToMe = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeletePrivacy = 'Yes' THEN 1
    ELSE 0
END
)
,[Reasons Involving Others Home] = (
CASE
    WHEN SurveyQuantitative.HomeKeepImportantToHousehold = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepSomebodyMightNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteNobodyElseNeeds = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeDeleteConfidential = 'Yes' THEN 1
    ELSE 0
END
)
,[Reasons Involving Others Work] = (
CASE
    WHEN SurveyQuantitative.WorkKeepImportantToBusiness = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepSomebodyMightNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteNobodyElseNeeds = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteConfidential = 'Yes' THEN 1
    ELSE 0
END
)
,[Reasons Involving Others Work 2] = (
CASE
    WHEN SurveyQuantitative.WorkKeepImportantToBusiness = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkKeepSomebodyMightNeed = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteNobodyElseNeeds = 'Yes' THEN 1
    ELSE 0

```

```

END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteConfidential = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkDeleteBusinessDoesntNeed = 'Yes' THEN 1
    ELSE 0
END
)
,[Record Generating Activity Home] = (
SurveyQuantitative.HomeEmail
+ SurveyQuantitative.HomePersonalFinance
+ SurveyQuantitative.HomeOtherFinance
)
,[Record Generating Activity Work] = (
+ SurveyQuantitative.WorkEmail
+ SurveyQuantitative.WorkPersonalFinance
+ SurveyQuantitative.WorkOtherFinance
)
),[RM Violation] = (
CASE
    WHEN SurveyQuantitative.HomeKeepWorkObjects = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Very
    Like' THEN 1
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Likely'
    THEN 1
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = '
    Somewhat ' THEN 1
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Perhaps
    ' THEN 0
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Maybe
    Not' THEN -1
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Most
    Like' THEN -1
    WHEN SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany = 'Not At
    All' THEN -1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Very
    Like' THEN 1
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Likely'
    THEN 1
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Somewhat
    ' THEN 1
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Perhaps'
    THEN 0
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Maybe
    Not' THEN -1
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Most
    Like' THEN -1
    WHEN SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy = 'Not At
    All' THEN -1
    ELSE 0

```

```

END
)
,[Training Organisation] = (
CASE
    WHEN SurveyQuantitative.JobSkillAcquiredFormallyDuringWorkTime > 0 THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.JobSkillAcquiredFormallyOutsideWork > 0 THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.JobSkillAcquiredSeminarsConferences > 0 THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkInstructedOnRM = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkAllEmpsInstructedOnRM = 'Yes' THEN 1
    ELSE 0
END
)
,[Training Self] = (
CASE
    WHEN SurveyQuantitative.JobSkillAcquiredOJT > 0 THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.JobSkillAcquiredOnOwnAtWork > 0 THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.JobSkillAcquiredOnOwnOutsideWork > 0 THEN 1
    ELSE 0
END
)
,[Use Default Stores Home] = (
CASE
    WHEN SurveyQuantitative.HomeUseDefaultDocuments = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeUseDefaultDownloads = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeUseDefaultLibrary = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeUseDefaultMovies = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.HomeUseDefaultMusic = 'Yes' THEN 1
    ELSE 0

```

```

END
+ CASE
    WHEN SurveyQuantitative.HomeUseDefaultPictures = 'Yes' THEN 1
    ELSE 0
END
)
,[Work RM System] = (
CASE
    WHEN SurveyQuantitative.WorkEmailMoveToOtherSystem = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkCentrallyManagedDocuments = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHasRM = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkHasRMPolicies = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkInstructedOnRM = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkAllEmpsInstructedOnRM = 'Yes' THEN 1
    ELSE 0
END
+ CASE
    WHEN SurveyQuantitative.WorkAuditsDocs = 'Yes' THEN 1
    ELSE 0
END
)
FROM          SurveyQuantitative
LEFT JOIN     SurveyTagsFlat
ON           SurveyQuantitative.ID = SurveyTagsFlat.ID
LEFT JOIN     SurveyInvalid
ON           SurveyQuantitative.ID = SurveyInvalid.ID
WHERE        SurveyInvalid.ID IS NULL

GO

```

### Coding the High-Level Quantitative Analysis

This section represents an exploratory examination of the data—essentially, fumbling around with it, attempting to find whether there are any relationships between response variables.

```

/*
sp_help vwQuantitativeHL

select * from sysobjects where name = 'vwQuantitativeHL'
select name, colorder from syscolumns where id = 1730105204 order by colorder
*/

```

```

/*
column 1 = ID (count this, so we know what we're averaging)
Columns 2 through 20 are potentially 'causal' variables,
so we'll mash those against columns 21-42, calculating
the sum, average, and standard deviation of columns 21-42 for
each grouping
*/

SET NOCOUNT ON

IF EXISTS ( SELECT * FROM SYSOBJECTS WHERE ID = OBJECT_ID(N'HLProbability'))
    DROP TABLE HLProbability
GO

CREATE TABLE HLProbability (
    CausalVarName      varchar(48) not null ,
    CausalVarVal       varchar(48) not null ,
    DependentVarName   varchar(48) not null ,
    CntVar             INT not null ,
    SumVar             INT not null ,
    MinVar             INT not null ,
    MaxVar             INT not null ,
    AvgVar             real not null ,
    StDevVar           real null )

GO

/*
Give ourselves a temporary (denoted by the hash) stored procedure to iterate
through that column & every subsequent column. Since we know column 21 is the
first "dependent" variable, start there (ColOrder is the place of the
column from left-to-right, so that tells us the column number, essentially).
*/
CREATE PROCEDURE #DoHLProbability(
    @ColName          varchar(48)
)
AS

/*
A variable to hold the second column's name
*/
DECLARE @ColName2 varchar(48)

/*
A cursor to spool through the columns as we go, considering only
those columns which might be "dependent" for this go.
*/
DECLARE CUR_DEPVARS CURSOR
FOR SELECT SYSCOLUMNS.NAME
FROM SYSCOLUMNS
INNER JOIN SYSOBJECTS
ON SYSCOLUMNS.ID = SYSOBJECTS.ID
WHERE SYSOBJECTS.NAME = 'vwQuantitativeHL'
AND SYSCOLUMNS.ColOrder >= 21

OPEN CUR_DEPVARS
FETCH NEXT FROM CUR_DEPVARS INTO @ColName2

/*

```

*For so long as we have our cursor in a good state, spool through, feeding aggregated records into HLPProbability. These are dynamic SQL statements because we don't know what columns are in the view, yet need to aggregate them (well, we know the columns, but it would be very tedious to write them all out as separate queries).*

```

*/
WHILE @@FETCHSTATUS = 0
    BEGIN
        EXEC( 'SELECT ''' + @ColName + ''' AS CausalVarName, [ ' + @ColName
            +
            ' ] AS CausalVarVal, ''' + @ColName2 + ''' AS DependentVarName,
COUNT([ID]) AS CntVar, SUM([ ' + @ColName2 + ' ]) AS SumVar, MIN([ ' +
@ColName2 + ' ]) AS MinVar, MAX([ ' + @ColName2 + ' ]) AS MaxVar, AVG([ ' +
@ColName2 + ' ]) AS AvgVar, StDev([ ' + @ColName2 + ' ]) AS StDevVar
INTO ##WKG
FROM vwQuantitativeHL
GROUP BY [ ' + @ColName + ' ]')

        EXEC( 'INSERT INTO HLPProbability ( CausalVarName , CausalVarVal ,
DependentVarName , CntVar , SumVar , MinVar , MaxVar , AvgVar , StDevVar )
SELECT CausalVarName , CausalVarVal ,
DependentVarName , CntVar , SumVar , MinVar , MaxVar , AvgVar , StDevVar
FROM ##WKG')

        EXEC( 'DROP TABLE ##WKG')

        PRINT @ColName + ' & ' + @ColName2

        FETCH NEXT FROM CUR_DEPVARS INTO @ColName2
    END

CLOSE CUR_DEPVARS
DEALLOCATE CUR_DEPVARS

GO

/*
Give ourselves a temporary stored procedure to call the previous
temporary stored procedure. This is the "outer" loop. Yes, we
could have put them into one, looping within a loop. It made my
head hurt, and confused the process, and this simplifies the thinking
about it.
*/

CREATE PROCEDURE #DriveHLPProbability
AS
/*
A variable to hold the column name.
*/
DECLARE @ColName varchar(48)

/*
A cursor to spool through the appropriate columns of the view
(not column 1, which is the ID column, and only through column
20, which is the last "driving" variable).
*/
DECLARE CUR_CAUSALVARS CURSOR

```

```

FOR SELECT SYSCOLUMNS.NAME
FROM SYSCOLUMNS
INNER JOIN SYSOBJECTS
ON SYSCOLUMNS.ID = SYSOBJECTS.ID
WHERE SYSOBJECTS.NAME = 'vwQuantitativeHL'
AND SYSCOLUMNS.ColOrder < 21
AND SYSCOLUMNS.ColOrder > 1

OPEN CUR_CAUSALVARS

FETCH NEXT FROM CUR_CAUSALVARS INTO @ColName

/*
Spool through, executing the analysis for each column.
*/
WHILE @@FETCHSTATUS = 0
    BEGIN
        EXEC #DoHLProbability @ColName
        FETCH NEXT FROM CUR_CAUSALVARS INTO @ColName
    END

CLOSE CUR_CAUSALVARS
DEALLOCATE CUR_CAUSALVARS

GO

EXEC #DriveHLProbability

GO

DROP PROC #DoHLProbability
DROP PROC #DriveHLProbability

```

### Coding the Sorted-Histograms

The following code was used to establish a count of each variable value's frequency, focusing on dependent variables.

```

/*
select * from sysobjects where name = 'vwQuantitativeHL'
select name, colorder from syscolumns where id = 1730105204 order by colorder

Get data for our bubble-charts from all columns, but we're mostly interested
in columns 21-42.
*/

SET NOCOUNT ON

IF EXISTS ( SELECT * FROM SYSOBJECTS WHERE ID = OBJECT_ID(N'HLHistograms') )
    DROP TABLE HLHistograms
GO

CREATE TABLE HLHistograms (
    VariableName          varchar(48) not null ,
    VariableVal           INT not null ,
    VariableCnt           INT not null
)

```

```

GO

DECLARE @ColName varchar(48)

DECLARE CUR_HIST CURSOR
FOR SELECT SYSCOLUMNS.NAME
FROM SYSCOLUMNS
INNER JOIN SYSOBJECTS
ON SYSCOLUMNS.ID = SYSOBJECTS.ID
WHERE SYSOBJECTS.NAME = 'vwQuantitativeHL'
AND SYSCOLUMNS.ColOrder >= 21

OPEN CUR_HIST
FETCH NEXT FROM CUR_HIST INTO @ColName

WHILE @@FETCHSTATUS = 0
    BEGIN
        EXEC( 'SELECT ''' + @ColName + ''' AS VariableName, [ ' + @ColName
            +
            ' ] AS VariableVal, COUNT([ID]) AS VariableCnt
        INTO ##AWKG FROM vwQuantitativeHL GROUP BY [ ' + @ColName + ' ]' )

        EXEC( 'INSERT INTO HLHistograms ( VariableName , VariableVal ,
            VariableCnt )
        SELECT VariableName , VariableVal , VariableCnt FROM ##AWKG' )

        EXEC( 'DROP TABLE ##AWKG' )

        FETCH NEXT FROM CUR_HIST INTO @ColName
    END

CLOSE CUR_HIST
DEALLOCATE CUR_HIST

select * from HLHistograms order by VariableName , VariableVal

```

### View for Quantitative Analysis

This view of the survey data represents data which has been first exported from LimeSurvey into .csv format, manipulated in Excel, imported into MSSQL, codified, and subsequently had qualitative codes applied. This view is for the purposes of data export into .csv format yet again, and for import into PSPP for further analysis.

```

CREATE VIEW vwQuantitative
AS

SELECT SurveyQuantitative.HomeCountry
    , SurveyQuantitative.HomeState
    , SurveyQuantitative.OverallComputerSkill
    , SurveyQuantitative.UseComputerAtHome
    , SurveyQuantitative.HomeComputerShared
    , SurveyQuantitative.HomeAccountShared
    , SurveyQuantitative.HomeOSWindows
    , SurveyQuantitative.HomeOSMac10
    , SurveyQuantitative.HomeOSMacOlder
    , SurveyQuantitative.HomeOSLinux

```

,SurveyQuantitative.HomeOSOther  
,SurveyQuantitative.HomeCleanupEmail  
,SurveyQuantitative.HomeCleanupContacts  
,SurveyQuantitative.HomeCleanupDocuments  
,SurveyQuantitative.HomeCleanupSocialNetworkStatus  
,SurveyQuantitative.HomeCleanupTextMessages  
,SurveyQuantitative.HomeCleanupMobileContacts  
,SurveyQuantitative.HomeWatchVideo  
,SurveyQuantitative.HomeConferenceCalling  
,SurveyQuantitative.HomeSearchInternet  
,SurveyQuantitative.HomeSocialNetworking  
,SurveyQuantitative.HomeCreateNonWebContent  
,SurveyQuantitative.HomeCreateWebContent  
,SurveyQuantitative.HomeEmail  
,SurveyQuantitative.HomePlayGames  
,SurveyQuantitative.HomeReadDistanceLearning  
,SurveyQuantitative.HomePersonalFinance  
,SurveyQuantitative.HomeOtherFinance  
,SurveyQuantitative.HomePersonal  
,SurveyQuantitative.HomeOtherPersonal  
,SurveyQuantitative.HomeChatting  
,SurveyQuantitative.HomeUseDefaultDocuments  
,SurveyQuantitative.HomeUseDefaultDownloads  
,SurveyQuantitative.HomeUseDefaultLibrary  
,SurveyQuantitative.HomeUseDefaultMovies  
,SurveyQuantitative.HomeUseDefaultMusic  
,SurveyQuantitative.HomeUseDefaultPictures  
,SurveyQuantitative.HomeFileEmail  
,SurveyQuantitative.HomeFileDocuments  
,SurveyQuantitative.HomeFileMusic  
,SurveyQuantitative.HomeFilePictures  
,SurveyQuantitative.HomeFileVideos  
,SurveyQuantitative.HomeFileRecords  
,SurveyQuantitative.HomeKeepMightNeed  
,SurveyQuantitative.HomeKeepImportantMemory  
,SurveyQuantitative.HomeKeepWantToWorkOn  
,SurveyQuantitative.HomeKeepImportantToHousehold  
,SurveyQuantitative.HomeKeepSomebodyMightNeed  
,SurveyQuantitative.HomeKeepSpentTimeOnIt  
,SurveyQuantitative.HomeKeepSomebodySpentTimeOnIt  
,SurveyQuantitative.HomeKeepLegalReasons  
,SurveyQuantitative.HomeKeepEmotionallyAttached  
,SurveyQuantitative.HomeKeepInteresting  
,SurveyQuantitative.HomeKeepTakesLittleSpace  
,SurveyQuantitative.HomeKeepMyCreation  
,SurveyQuantitative.HomeKeepProveActions  
,SurveyQuantitative.HowManyComputersAtHome  
,SurveyQuantitative.HowManyPeopleAtHome  
,SurveyQuantitative.HomeComputersShareInternet  
,SurveyQuantitative.HomeComputersCopyFiles  
,SurveyQuantitative.HomeComputersUseNAS  
,SurveyQuantitative.HomeComparativeComputerSkill  
,SurveyQuantitative.HomeDeleteDontNeed  
,SurveyQuantitative.HomeDeleteWrongToKeep  
,SurveyQuantitative.HomeDeleteNobodyElseNeeds  
,SurveyQuantitative.HomeDeleteGetsInWay  
,SurveyQuantitative.HomeDeleteItsOld  
,SurveyQuantitative.HomeDeleteSomebodyHasCopy

,SurveyQuantitative.HomeDeleteTooMuchTroubleToKeep  
,SurveyQuantitative.HomeDeleteNoWayToSave  
,SurveyQuantitative.HomeDeleteEmotionallyTroublesome  
,SurveyQuantitative.HomeDeleteNotImportantToMe  
,SurveyQuantitative.HomeDeleteTakesUpSpace  
,SurveyQuantitative.HomeDeleteConfidential  
,SurveyQuantitative.HomeDeletePrivacy  
,SurveyQuantitative.HomeHelpDeleteFolder  
,SurveyQuantitative.HomeHelpDeleteFileType  
,SurveyQuantitative.HomeHelpDeleteFileName  
,SurveyQuantitative.HomeHelpDeleteCreationDate  
,SurveyQuantitative.HomeHelpDeleteModifiedDate  
,SurveyQuantitative.HomeHelpDeleteEmailSenders  
,SurveyQuantitative.HomeHelpDeleteEmailSubject  
,SurveyQuantitative.HomeHelpDeleteEmailFlagged  
,SurveyQuantitative.HomeHelpDeleteEmailAttachments  
,SurveyQuantitative.HomeHelpDeleteContent  
,SurveyQuantitative.HomeKeepWorkObjects  
,SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompany  
,SurveyQuantitative.HomeKeepWorkObjectsAfterLeavingCompanyShouldKeep  
,SurveyQuantitative.WorkKeepHomeObjects  
,SurveyQuantitative.WorkKeepHomeObjectsUpsetIfReviewed  
,SurveyQuantitative.UseComputerAtWork  
,SurveyQuantitative.ShareWorkComputer  
,SurveyQuantitative.WorkAccountShared  
,SurveyQuantitative.JobSkillAcquiredOJT  
,SurveyQuantitative.JobSkillAcquiredOnOwnAtWork  
,SurveyQuantitative.JobSkillAcquiredOnOwnOutsideWork  
,SurveyQuantitative.JobSkillAcquiredFormallyDuringWorkTime  
,SurveyQuantitative.JobSkillAcquiredFormallyOutsideWork  
,SurveyQuantitative.JobSkillAcquiredSeminarsConferences  
,SurveyQuantitative.JobSkillAcquiredOther  
,SurveyQuantitative.WorkWatchVideo  
,SurveyQuantitative.WorkConferenceCalling  
,SurveyQuantitative.WorkSearchInternet  
,SurveyQuantitative.WorkSocialNetworking  
,SurveyQuantitative.WorkCreateNonWebContent  
,SurveyQuantitative.WorkCreateWebContent  
,SurveyQuantitative.WorkEmail  
,SurveyQuantitative.WorkPlayGames  
,SurveyQuantitative.WorkReadDistanceLearning  
,SurveyQuantitative.WorkPersonalFinance  
,SurveyQuantitative.WorkOtherFinance  
,SurveyQuantitative.WorkPersonal  
,SurveyQuantitative.WorkOtherPersonal  
,SurveyQuantitative.WorkChatting  
,SurveyQuantitative.WorkKeepMightNeed  
,SurveyQuantitative.WorkKeepImportantMemory  
,SurveyQuantitative.WorkKeepWantToWorkOn  
,SurveyQuantitative.WorkKeepImportantToBusiness  
,SurveyQuantitative.WorkKeepSomebodyMightNeed  
,SurveyQuantitative.WorkKeepSpentTimeOnIt  
,SurveyQuantitative.WorkKeepSomebodySpentTimeOnIt  
,SurveyQuantitative.WorkKeepLegalReasons  
,SurveyQuantitative.WorkKeepEmotionallyAttached  
,SurveyQuantitative.WorkKeepInteresting  
,SurveyQuantitative.WorkKeepTakesLittleSpace  
,SurveyQuantitative.WorkKeepMyCreation

,SurveyQuantitative.WorkKeepProveActions  
,SurveyQuantitative.WorkEmailProvided  
,SurveyQuantitative.WorkEmailSaveCopy  
,SurveyQuantitative.WorkEmailFileImmediately  
,SurveyQuantitative.WorkEmailAutoArchive  
,SurveyQuantitative.WorkEmailMoveToOtherSystem  
,SurveyQuantitative.WorkCentrallyManagedDocuments  
,SurveyQuantitative.WorkCentrallyManagedDocumentsKeepCopy  
,SurveyQuantitative.WorkHowManyPeopleInCompany  
,SurveyQuantitative.WorkHowManyPeopleInDepartment  
,SurveyQuantitative.WorkHowManyPeoplePerformSameFunctionInCompany  
,SurveyQuantitative.WorkHowManyPeoplePerformSameFunctionInDepartment  
,SurveyQuantitative.WorkPrimaryIndustry  
,SurveyQuantitative.WorkPrimaryIndustryOther  
,SurveyQuantitative.WorkComparativeComputerSkill  
,SurveyQuantitative.WorkHasRM  
,SurveyQuantitative.WorkHasRMPolicies  
,SurveyQuantitative.WorkInstructedOnRM  
,SurveyQuantitative.WorkAllEmpsInstructedOnRM  
,SurveyQuantitative.WorkAuditsDocs  
,SurveyQuantitative.WorkDeleteDontNeed  
,SurveyQuantitative.WorkDeleteWrongToKeep  
,SurveyQuantitative.WorkDeleteNobodyElseNeeds  
,SurveyQuantitative.WorkDeleteGetsInWay  
,SurveyQuantitative.WorkDeleteItsOld  
,SurveyQuantitative.WorkDeleteSomebodyHasCopy  
,SurveyQuantitative.WorkDeleteTooMuchTroubleToKeep  
,SurveyQuantitative.WorkDeleteNoWayToSave  
,SurveyQuantitative.WorkDeleteEmotionallyTroublesome  
,SurveyQuantitative.WorkDeleteNotImportantToMe  
,SurveyQuantitative.WorkDeleteTakesUpSpace  
,SurveyQuantitative.WorkDeleteConfidential  
,SurveyQuantitative.WorkDeletePrivacy  
,SurveyQuantitative.WorkDeleteBusinessDoesntNeed  
,SurveyQuantitative.WorkHelpDeleteFolder  
,SurveyQuantitative.WorkHelpDeleteFileType  
,SurveyQuantitative.WorkHelpDeleteFileName  
,SurveyQuantitative.WorkHelpDeleteCreationDate  
,SurveyQuantitative.WorkHelpDeleteModifiedDate  
,SurveyQuantitative.WorkHelpDeleteEmailSenders  
,SurveyQuantitative.WorkHelpDeleteEmailSubject  
,SurveyQuantitative.WorkHelpDeleteEmailFlagged  
,SurveyQuantitative.WorkHelpDeleteEmailAttachments  
,SurveyQuantitative.WorkHelpDeleteContent  
,SurveyQuantitative.WorkAsPermanentEmployee  
,SurveyQuantitative.WorkHoursWorkedPerWeek  
,SurveyQuantitative.WorkHowLongAtCurrentCompany  
,SurveyQuantitative.WorkHowLongAtTypeOfWork  
,SurveyQuantitative.WorkPercentTimeSupervising  
,SurveyQuantitative.Gender  
,SurveyQuantitative.Age  
,SurveyTagsFlat.Administrative  
,SurveyTagsFlat.Archivist  
,SurveyTagsFlat.Aware\_Of\_Legal\_Retention\_Period\_In\_Deleting  
,SurveyTagsFlat.Busines\_Person  
,SurveyTagsFlat.Casual\_Attitude  
,SurveyTagsFlat.Civil\_Servant  
,SurveyTagsFlat.Curator

```

, SurveyTagsFlat . Does_Not_Believe_Digital_Can_Be_Owned
, SurveyTagsFlat . Does_Not_DeleteAnything
, SurveyTagsFlat . Ecologist
, SurveyTagsFlat . Educator
, SurveyTagsFlat . Encrypts_Private_Files_At_Work
, SurveyTagsFlat . Impossible_NOT_to_do_Personal_Things_At_Work
, SurveyTagsFlat . Journalist
, SurveyTagsFlat . Justifies_Because_IP_Not_Valuable
, SurveyTagsFlat . Keeps_Copy_Of_Centrally_Managed_Documents
, SurveyTagsFlat . Keeps_For_Reference
, SurveyTagsFlat . Keeps_Personal_On_Work_Computer
, SurveyTagsFlat . Keeps_To_CYA
, SurveyTagsFlat . Keeps_Work_IP
, SurveyTagsFlat . Keeps_Work_Materials
, SurveyTagsFlat . Lawyer
, SurveyTagsFlat . Librarian
, SurveyTagsFlat . Machiavellian
, SurveyTagsFlat . Marketing
, SurveyTagsFlat . Medical_Professional
, SurveyTagsFlat . Mixes_Home_And_Work
, SurveyTagsFlat . Nostalgia
, SurveyTagsFlat . Organizational_System_Not_Receptive
, SurveyTagsFlat . Organizational_System_Qualified_Acceptance
, SurveyTagsFlat . Organizational_System_Receptive
, SurveyTagsFlat . Organize_Because_Documents_Important
, SurveyTagsFlat . Organize_Because_Organization_Important
, SurveyTagsFlat . Organize_For_Deletion
, SurveyTagsFlat . Organize_For_Efficiency
, SurveyTagsFlat . Organize_For_Finding
, SurveyTagsFlat . Organize_For_Posterity
, SurveyTagsFlat . Personal_Files_On_Separate_Drive
, SurveyTagsFlat . Privacy_Concerns
, SurveyTagsFlat . Project_Manager
, SurveyTagsFlat . Protects_Home_From_Work
, SurveyTagsFlat . Protects_Work_From_Home
, SurveyTagsFlat . Records_Manager
, SurveyTagsFlat . Recruiter
, SurveyTagsFlat . Regards_Work_IP_As_Their_Property
, SurveyTagsFlat . Secretarial
, SurveyTagsFlat . Social_Work
, SurveyTagsFlat . Student
, SurveyTagsFlat . Technologist
, SurveyTagsFlat . Too_Much_Trouble_To_Separate_Home_From_Work
, SurveyTagsFlat . Unjustified
, SurveyTagsFlat . Upset_If_Personal_Files_Examined_At_Work
FROM          SurveyQuantitative
LEFT JOIN     SurveyTagsFlat
ON           SurveyQuantitative.ID = SurveyTagsFlat.ID
LEFT JOIN     SurveyInvalid
ON           SurveyQuantitative.ID = SurveyInvalid.ID
WHERE        SurveyInvalid.ID IS NULL
GO

```

## Probability Analysis - Calculating Basic Probabilities

This section essentially bashes each variable against the others, dumping the results into a table for further consideration. It calculates the probability of a pair of variables with particular

values occurring in conjunction. For example, if we have two questions with yes / no answers, this solves for:

- What is the probability that question 1 answers ‘yes’ and question 2 answers ‘yes’?
- What is the probability that question 1 answers ‘yes’ and question 2 answers ‘no’?
- What is the probability that question 1 answers ‘no’ and question 2 answers ‘yes’?
- What is the probability that question 1 answers ‘no’ and question 2 answers ‘no’?

The logic is not that simple, however, particularly when considering responses which have the possibility of answering, e.g., in the range of 1 to 7. This script must take the range of possible answers for each variable under consideration and determine the probability of each possible response of the first variable’s intersection with the probability of each possible response of the second variable.

```

/*
CREATE TABLE RawProbability (
    ColName varchar(48) NOT NULL ,
    ColVal varchar(48) NOT NULL ,
    Cnt INT NOT NULL ,
    Probability FLOAT NOT NULL )
*/

CREATE TABLE SplitSetProbability (
    FirstColName varchar(48) not null ,
    FirstColVal varchar(48) not null ,
    SecondColName varchar(48) not null ,
    SecondColVal varchar(48) not null ,
    FirstColQty int null default 0 ,
    IntersectionQty int null default 0 )

GO

CREATE PROCEDURE #DoSplitProbability
(
    @ColName          varchar(48)
    , @ColOrder       int
)
AS
PRINT @ColName

DECLARE @ColName2 VARCHAR(48)

DECLARE CUR_COLUMNS2 CURSOR
FOR SELECT SYSCOLUMNS.NAME
FROM SYSCOLUMNS
INNER JOIN SYSOBJECTS
ON SYSCOLUMNS.ID = SYSOBJECTS.ID
WHERE SYSCOLUMNS.NAME <> 'ID'
AND SYSOBJECTS.NAME = 'vwQuantitative'
AND SYSCOLUMNS.ColOrder <> @ColOrder
ORDER BY SYSCOLUMNS.COLORDER

OPEN CUR_COLUMNS2

FETCH NEXT FROM CUR_COLUMNS2 INTO @ColName2

```

```

WHILE @@FETCHSTATUS = 0
    BEGIN
        exec('INSERT INTO SplitSetProbability (FirstColName, FirstColVal,
            SecondColName, SecondColVal, IntersectionQty)
SELECT ''' + @ColName + ''', vwQuantitative.' + @ColName + ', ''' +
    @ColName2 + ''', vwQuantitative.' + @ColName2 + ', COUNT([ID])
FROM vwQuantitative
WHERE vwQuantitative.' + @ColName + ' IS NOT NULL
AND vwQuantitative.' + @ColName2 + ' IS NOT NULL
GROUP BY vwQuantitative.' + @ColName + ', vwQuantitative.' + @ColName2)

        exec('SELECT vwQuantitative.' + @ColName + ', COUNT([ID]) AS CNT
INTO #wkg
FROM vwQuantitative
WHERE vwQuantitative.' + @ColName + ' IS NOT NULL
AND vwQuantitative.' + @ColName2 + ' IS NOT NULL
GROUP BY vwQuantitative.' + @ColName + ')

UPDATE SplitSetProbability SET FirstColQty = CNT
FROM #wkg
INNER JOIN SplitSetProbability
ON #wkg.' + @ColName + ' = SplitSetProbability.FirstColVal
WHERE SplitSetProbability.FirstColName = ''' + @ColName + ''')

        FETCH NEXT FROM CUR_COLUMNS2 INTO @ColName2
    END

CLOSE CUR_COLUMNS2
DEALLOCATE CUR_COLUMNS2

GO

DECLARE @ColName1 VARCHAR(48)
DECLARE @COLORDER1 INT

DECLARE CUR_COLUMNS1 CURSOR
FOR SELECT SYSCOLUMNS.NAME, SYSCOLUMNS.COLORDER
FROM SYSCOLUMNS
INNER JOIN SYSOBJECTS
ON SYSCOLUMNS.ID = SYSOBJECTS.ID
WHERE SYSCOLUMNS.NAME <> 'ID'
AND SYSOBJECTS.NAME = 'vwQuantitative'
/*
—Include this critereon if you're only testing the algorith.
AND SYSCOLUMNS.COLORDER = 1
*/
ORDER BY SYSCOLUMNS.COLORDER

OPEN CUR_COLUMNS1

FETCH NEXT FROM CUR_COLUMNS1 INTO @ColName1, @COLORDER1

WHILE @@FETCHSTATUS = 0
    BEGIN
        EXEC #DoSplitProbability @ColName1, @COLORDER1
        FETCH NEXT FROM CUR_COLUMNS1 INTO @ColName1, @COLORDER1
    END
END

```

```
CLOSE CUR_COLUMNS1
DEALLOCATE CUR_COLUMNS1

GO
/*
Run the procedure, then drop it 'cause we don't need it any more.
*/
EXEC #DoSplitProbability
DROP PROCEDURE #DoSplitProbability

/*
Calculate the intersection probabilities based upon p1 / p2.
*/

ALTER TABLE SplitSetProbability ADD Probability FLOAT NULL

UPDATE SplitSetProbability
    SET Probability = ROUND(CAST(IntersectionQty AS FLOAT)
    / CAST(FirstColQty AS FLOAT), 2)

/*
Index the table so that we can query it more efficiently. Since
we're not adding anything else, ever, these indexes can be
100 percent full, not leaving any room at the level of leaf
nodes for new rows to be added to the binary (technically B+) trees.
*/
CREATE CLUSTERED INDEX IX_Cols@SplitSetProbability
ON SplitSetProbability ( FirstColName, SecondColName )
WITH FILLFACTOR = 100

CREATE INDEX IX_Probability@SplitSetProbability
ON SplitSetProbability ( Probability )
WITH FILLFACTOR = 100

CREATE INDEX IX_All@SplitSetProbability
ON SplitSetProbability ( FirstColName,
    SecondColName,
    FirstColVal,
    SecondColVal,
    Probability )
WITH FILLFACTOR = 100

GO
```

---

# Index

---

- "Whatever I am, Wherever I am, How does it Matter?... Why does it Matter?" Egocasting in-between Identities*, 62, 147
- A NASA Technical White Paper: Implementing the NASA Taxonomy Through Service Oriented Architectures to Promote Knowledge Sharing and Increased Mission Success*, 5, 147
- Evolution of scientific and technical information distribution*, 5, 147
- MHS: A distributed metadata management strategy*, 5, 154
- A Belief-Based Account of Decision under Uncertainty*, 47–49, 147
- Accountability, 26
- Actor, 15, 16
- Affirmative Action, 21
- Affordance, 4, 15
- Age, 9
- Ageing, 10
- Agency, 12, 13, 16, 17
- Ambiguity, 53
- Ambiguity Aversion, 53
- Ambiguity Aversion and Comparative Ignorance*, v, 48, 53, 148
- Amoral, 11, 13, 16
- Amoral Agent, 11, 13
- Amoral Calculator, 19
- Analogue, 5, 14
- Anchoring, 47
- Arendt, Hannah, 26–28, 146
- Artistry, 11
- Asymmetric Dominance, 31, 32
- Asymmetry
- of Blame, 52
  - of Credit, 52
- Audit, 22
- Authenticity
- Catalytic, 76
  - Educative, 76
  - Ontological, 76
- Authority, 19
- Avatar, 11
- Bailey, Steve, 1, 4, 5, 146
- Balsamo, Anne, 18, 146
- Being and Time*, 5, 149
- Bias, 71, 77, 80, 81
- Big Bucket, 5
- Binark, Mutlu, 61, 146
- Blame, 19, 20, 28, 52, 53, 58
- Blog Post, 9, 59
- Bovens, Mark, 20, 21, 26, 27, 146
- Brink, Hilla., 81, 146
- Bryman, Alan, 70–74, 76, 79–83, 86, 92, 93, 147
- Bureaucracy, 26
- Business Management, 21, 61
- Causal, 54, 55
- Causation, 58, 70
- Cheating more when the spoils are split*, 28, 29, 89, 154
- Clustering, 35, 38
- Code : version 2.0 / Lawrence Lessig.*, 19, 150
- Cognition, computation, and consciousness*, 36, 149
- Cognitive Error, 2, 35, 132, 133, 137
- Cohen, Laura B., 4, 147
- Common Features, 33
- Community of Practise, 11
- Comparative Ignorance Hypothesis, 53
- Comparison, 33
- Comparison Error, 32, 132–134
- Compeau, Deborah R., 83, 147
- Complex Organisation, 26, 27
- Compliance, 20
- Compromise Effect, 132, 134
- Compromise Effects, 41, 42
- Compromise Verdict, 41
- Computer Self-Efficacy: Development of a Measure and Initial Test*, 83, 147
- Conformity, 27
- Conjunction Error, 38
- Conscience, 27, 28
- Contamination, 23
- Content, 14, 15
- Location, 7
  - Private, 7
  - Shared, 7
- Context, 30, 36

- Context Effect, 32  
*Context-Dependence in Legal Decision Making*, 41, 43, 149
- Contrast Effect, 132, 134
- Contrast Effects, 42
- Copy, 5
- Corporate Decision, 26
- Cost Control, 10
- Credibility Testing, 93  
*Critique of Pure Reason*, 28, 149
- Cultural Change, 17
- Curation, 7
- Cyborg, 17, 18
- Database Records, 3
- Decision Under Adversity, 57
- Delegation, 13–16
- Dervin, Fred, 62, 147
- Desktop, 7
- Destruction, 10
- Difference, 31, 33, 37
- Digital Junk, 6
- Digital Objects, 13
- Digital Records, 10
- Direction  
of Asymmetry, 37
- Directionality  
of Choice, 37  
of Comparison, 37
- Discrepancy between Medical Decisions for Individual Patients and for Groups*, 29, 151
- Discrimination, 21
- Disjunction Effect, 56
- Dissimilarity, 36, 38, 39
- Distinctive Features, 33
- Dutra, J., 5, 147
- Duty, 2, 11, 12, 19, 30, 34, 61, 143  
to Others, 11, 32, 34, 35, 143  
to Preserve, 59  
to Self, 143
- Edmundson, William A., 19, 147
- Effective Context, 39
- Efficiency, 10
- Email, 2, 5, 59, 60, 95–97, 100, 101, 103, 105, 110, 111, 116, 132, 136, 142
- Email Contacts, 2, 95, 110
- Engineering, 11
- Enriched Option, 30, 132, 134  
*Envisioning Cyborg Bodies: Notes from Current Research*, 17, 148
- Esler, S.L., 5, 147  
*Ethical Decision Making by Individuals in Organizations: An Issue-Contingent Model*, 23–25, 27, 138, 139, 149  
*Ethical Standards within Organizational Behavior*, 21, 22, 27, 148
- Ethics, 16, 61
- Ethnography, 78
- Evil, 18, 27
- Extension of the Self, 10–12, 16, 17  
*Extensional vs. Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 44, 46, 57, 58, 152
- Extremeness Aversion, 32, 42
- Facing Diversity: The Case of Epistemic Abstinence*, 19, 151
- Fear, 10
- Fear of Ageing, 10
- Feature Representations, 37
- Features  
Distinguishing, 32  
Irrelevant, 32  
*Features of Similarity*, 36–39, 153
- Fiduciary Duty, 3
- Focus Groups, 80
- Folksonomy, 5
- Foot, Philippa, 25, 148
- Form, 14  
*Formal Games as Probing Tools for Investigating Behavior Motivated by Trust and Suspicion*, 22–24, 27, 59, 138, 139, 151
- Fox, Craig R., v, 47–49, 53, 147, 148
- Fox, Robin, 57, 58, 148
- Free Will, 58  
*From Virtual Cyborgs to Biological Time Bombs: Technocriticism and the Material Body*, 10, 154
- Function, 14  
*Fundamentals of research methodology for health care professionals*, 81, 146
- Future Self, 3
- Games Theory, 22
- Gati, Itamar, 33, 90, 148
- Gender, 18
- Gender Role, 18
- Gilovich, Thomas, 34, 35, 152
- Glinow, Mary Ann Von, 21, 22, 27, 148  
*Global Industry Classification Standard*, 85, 150  
*GNU PSPP*, 93, 150
- Gonzalez, Jennifer, 17, 148
- Good, 18
- Grounded Theory, 73, 74, 92
- Grouping, 39
- Guilt, 27, 41
- Habermas, Jurgen, 25, 138, 139, 148
- Hacker, 11
- Hard Masters, 11
- HCI Design, 143, 144
- Heath, Chip, 52, 148
- Heidegger, Martin, 5, 149
- Higgins, Christopher A., 83, 147
- Holocaust, 26, 27
- Home Culture, 3

- Hospers, John, 25, 149  
Hume, David, 25, 149  
Humean Theory of Motivation, 25  
Hybrid, 17
- Immoral, 11  
Immoral Agent, 11  
Impoverished Option, 30, 132, 134  
Indicative, 54, 55  
Inferior Option, 30, 31  
Information Load, 38  
Information Management, 59  
Information Science, 61  
Innocence, 27  
Innovation, 10  
*Institutional repositories : content and culture in an open access environment*, 44, 149  
Intention, 69  
Internalised Moral Norm, 20, 21  
*International Classification of Activities for Time-Use Statistics*, 85, 153  
Interviews, 74  
    Semi-structured, 79  
    Structured, 79  
Intimacy, 11, 12  
Irrelevant Information, 56  
Ito, Masao, 36, 149
- JISC infoNet, 85, 149  
Jones, Catherine., 44, 149  
Jones, Thomas M., 23–25, 27, 138, 139, 149  
*Judgment under Uncertainty: Heuristics and Biases*, 43–46, 152
- Kahneman, Daniel, v, 43–46, 50, 51, 57, 58, 152  
Kant, Immanuel, 28, 149  
Kelman, Mark, 41, 43, 149  
Kim, Jong-Young, 19, 150  
Knowledge Management, 61
- Latour, Bruno, 15, 16, 150  
Lay Records Manager, 4  
Legal Environment, 7  
Lessig, Lawrence, 19, 150  
Levy, David M., 5, 13–15, 150  
Lewin, Cathy., 81, 151  
*Library 2.0 initiatives in academic libraries*, 4, 147  
*Life on the screen : identity in the age of the internet*, 10, 152  
Linguistic Perception, 34  
Lock-in, 22  
Loss Averse, 39  
Loss Seeking, 39
- Magic, 54, 55  
Mala In Se, 20  
Mala Prohibita, 20  
*Managing the crowd : rethinking records management for the Web 2.0 world*, 1, 4, 5, 146
- Mastery, 11  
McMillan, James H., 81, 150  
Mediation, 14  
Metadata, 14, 15, 134, 136, 137  
Mixed Methods Study, 71  
Miyashita, Yasushi, 36, 149  
Moral, 16  
Moral Development, 25  
Morality, 11, 12, 19, 61  
    Moral Agency, 12, 13  
    Moral Agent, 20  
    Moral Development, 24  
    Moral Duty, 3, 12  
    Moral Field, 23  
    Moral Intensity, 23  
    Moral Judgment, 2  
    Moral Obligation, 12  
    Moral Philosophy, 27  
    Normative Engagement, 138
- Movies, 2  
MSCI Barra, 85, 150  
Multi-national Organisation, 26  
Murphy, Mark C., 19, 25, 150  
Music, 105  
My Documents, 7, 100, 108
- Nash Equilibrium, 20, 22, 23  
Natural Person, 20  
Nazi, 26  
Negative Feature, 30  
Nelson, M.L., 5, 147  
Neurosis, 11  
Nonconformity, 27  
Norman, Donald A., 4, 15, 150  
Normative Structure, 27  
Normativity, 12, 19–22, 56  
*North American Industry Classification System, 2007*, 84, 153  
Novelli Luke, Jr., 21, 22, 27, 148
- Obligation, 2, 61  
Old, 9, 10  
*On Reason and the Emotions: the Fact/Value Distinction*, 25, 149  
Opaqueness, 26  
Operating System, 15, 16  
Ordering, 36  
Organisational Behaviour, 21  
Organisational Complexity, 26  
Organisational Culture, 27  
Original, 5  
Ought Statement, 29
- Participant Observation, 77  
Pattern Recognition, 35  
Patton, Michael Quinn., 81, 150  
Peer-group Pressure, 21  
Perceived Competence, 52  
Perfect Friendship, 11

- Perfect Love, 11
- Performance, 13
- Permissions and Supererogation*, 19, 151
- Personal Archive, 13
- Phenomenology, 19
- Photograph, 105
- Pictorial Evidence, 33
- Position Classification Standards for White Collar Work*, 85, 153
- Positive Feature, 30
- Posterity, 3, 5, 7, 138
- Practical Morality, 24
- Practicing Identity in the Digital Game World: The Turkish Tribes' Community Practices in "Silkroad Online"*, 61, 146
- Praise, 52, 53
- Preference and Belief: Ambiguity and Competence in Choice under Uncertainty*, 52, 148
- Prejudice, 9
- Prejudice and the Unfinished Mind: A New Look at an Old Failing*, 57, 58, 148
- Prisoners' Dilemma, 22
- Probability
- Anchoring, 45, 46
  - Calculation, 49
  - Calculation Errors, 47
  - Cause, 46
  - Chance, 45
  - Conjunction Error, 46, 47
  - Conjunctive Event, 46, 47
  - Current Probability, 43
  - Disjunctive Event, 46, 47
  - Effect, 46
  - Estimating Risk, 44
  - Estimation, 47, 49
  - Estimation Error, 51
  - Gambler's Fallacy, 47
  - Imaginability, 44
  - Known Probability, 43
  - Prediction, 43
  - Prior Probability, 43, 44
  - Worthless Evidence, 43
- Programming, 11
- Property-enriched, 38
- Property-poor, 38
- Prospect Theory, 50, 51
- Loss Aversion, 50
  - Possible Gain, 50
  - Possible Loss, 50
  - Risk Aversion, 51
  - Willingness to Bet, 50
- Prudential Duty, 3
- Punishment, 19
- Qualitative Methods, 70, 71
- Qualitative research and evaluation methods*, 81, 150
- Qualitative Study, 72–74, 77
- Quantitative Methods, 70
- Quantitative Study, 72
- Quasi-magical Thinking, 55, 56
- Quattrone, George A., 54, 150
- Questions
- Closed, 73, 80, 92, 94
  - Open, 73, 92, 93
- Race, 21
- Rapoport, Anatol, 22–24, 27, 59, 138, 139, 151
- Rational and Social Theories as Complementary Explanations of Communication Media Choices: Two Policy-Capturing Studies*, 24, 25, 27, 64, 138, 139, 153
- Rational Choice and the Framing of Decisions*, v, 50, 51, 152
- Rational Choice Theory, 19
- Rational Choice, Situated Action, and the Social Control of Organizations*, 19, 20, 22–24, 26, 27, 59, 138, 139, 153
- Rational Intent, 19
- Raz, Joseph, 19, 25, 151
- Reading Cyborgs, Writing Feminism*, 18, 146
- Reason-Based Choice*, 30–32, 40–42, 136, 151
- Reassembling the Social: An Introduction to Actor Network Theory*, 15, 16, 150
- Reciprocal Relationship, 14
- Record, 7, 105
- Records and Information Management Maturity Model*, 85, 149
- Redelmeier, Donald A., 29, 151
- Regulation, 20–22
- Regulatory Environment, 7
- Rejection, 30, 31
- Relational Data, 35
- Rendering Process, 14
- Rensburg, Gisela. Van, 81, 146
- Research in education : a conceptual introduction*, 81, 150
- Research methods in the social sciences*, 81, 151
- Response Coding, 74
- Axial Coding, 74
  - Open Coding, 74
  - Selective Coding, 74
- Responsibility, 12, 14, 21, 26, 27, 34
- Responsibility and Judgment*, 26–28, 146
- Review: Rethinking Exclusionary Reasons: A Second Edition of Joseph Raz's "Practical Reason and Norms"*, 19, 147
- Review: The Morality of Obedience*, 19, 25, 151
- Right, 27
- Riikonen, Tanja, 62, 147
- Risk, 47
- Rolls, Edmund T., 36, 149
- Rottenstreich, Yuval, 41, 43, 149
- Rule-Utilitarianism*, 25, 149
- Sütcü, Günseli Bayraktutan, 61, 146
- Sample Size, 35

- Sampling  
 Criterion, 81  
 Purposive, 92  
 Quota, 92  
 Random, 71  
 Snowball, 71, 81  
 Theoretical, 73, 74, 81
- Schumacher, Sally., 81, 150
- Scrolling Forward : making sense of documents in the digital age*, 5, 13–15, 150
- Second Self, 11
- Selection, 10, 30, 31
- Self-Deception and the Voter's Illusion*, 54, 150
- Self-interest, 22
- Sensuality, 11
- Sequential Data, 35
- Shafir, Eldar, 30–32, 40–42, 55–57, 136, 151
- Similarity, 31, 33, 36–39
- Simonson, Itamar, 30–32, 40–42, 136, 151
- Social Change, 17
- Social Context, 15, 16, 19, 31, 32
- Social Control, 19, 20, 22, 61
- Social Interaction in Computer-Mediated Communication*, 19, 150
- Social Networking, 59
- Social research methods*, 70–74, 76, 79–83, 86, 92, 93, 147
- Sociology, 15
- Soft Masters, 11
- Software  
 LimeSurvey, 94  
 Microsoft Access, 93  
 Microsoft Excel, 94  
 Microsoft SQL Server, 93, 94  
 Microsoft Visual Basic for Applications, 94  
 NVivo, 93  
 PSPP, 93, 94  
 SPSS, 94  
 TweakUI, 109
- Somekh, Bridget., 81, 151
- Sorting, 36, 38, 39
- Standard Occupational Classification, 2000*, 85, 153
- Stated Belief, 29
- Stone, Allucquere Rosanne, 13, 16, 58–60, 152
- Sugden, Robert, 58, 152
- Superior Option, 30, *see also* Enriched Option, 31
- Surrender of Judgment and the Consent Theory of Political Authority*, 19, 25, 150
- Survey, 74, 80
- Synchronic, 17
- Taboo, 11, 18
- Taxonomy, 5, 33
- Technics, 13
- Technosociality, 13
- Text Message, 2, 110
- The Cold Facts about the Hot Hand in Basketball*, 34, 35, 152
- The Design of Everyday Things*, 4, 15, 150
- The quest for responsibility : accountability and citizenship in complex organisations*, 20, 21, 26, 27, 146
- The Role of Inductive Reasoning in the Evolution of Conventions*, 58, 152
- The second self : computers and the human spirit*, 10–12, 16, 18, 152
- Theoretical Saturation, 74, 81, 93
- Thinking Machines, 11
- Thinking through Uncertainty: Nonconsequential Reasoning and Choice*, 55–57, 151
- Thinking-With Machines, 11
- Third Reich, 26–28
- Transcription, 79
- Treasure Chest, 13
- Trevino, Linda Klebe, 24, 25, 27, 64, 138, 139, 153
- Triangulation, 80
- Truth and Justification*, 25, 138, 139, 148
- Turkle, Sherry, 10–12, 16, 18, 152
- Tversky, Amos, v, 29–58, 90, 136, 147–153
- U.K. Office for National Statistics, 85, 153
- U.S. Office of Personnel Management, 85, 153
- Unconscious Decision-making, 33, 36
- Unconscious Knowledge, 23
- United Nations Statistics Division, 85, 153
- United States Census Bureau, 84, 153
- Utilitarianism and the Virtues*, 25, 148
- Valueless Features, 40
- Vaughan, Diane, 19, 20, 22–24, 26, 27, 59, 138, 139, 153
- Verbal Evidence, 33
- Video, 105
- Virtual Culture, 13
- Visual Perception, 34
- Walt, C. Van der, 81, 146
- Wang, J., 5, 154
- Warranted Norm, 25
- Webster, Jane, 24, 25, 27, 64, 138, 139, 153
- Weighting Common and Distinctive Features in Perceptual and Conceptual Judgments*, 33, 90, 148
- Wiki Page, 7
- Will the Real Body Please Stand Up? Boundary Stories about Virtual Cultures*, 13, 16, 58–60, 152
- Wiltermuth, Scott S., 28, 29, 89, 154
- Woodward, Kathleen, 10, 154
- Wrong, 27
- Xiao, Q., 5, 147
- Youth, 9