Task, Interrupted: Understanding the Effect of Time Costs on Task Interruptions During Data Entry

Candidate:
Judith BORGHOUTS

Supervisors:
Dr. Duncan BRUMBY
Prof. Anna COX

August 9, 2018

A dissertation submitted in partial fulfilment of the requirements for the degree of:

Doctor of Philosophy of University College London

UCL Interaction Centre,
Department of Psychology,
University College London

Declaration

I, Judith Borghouts, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Computer-based work often involves looking up information from different sources. Though these interruptions are required to progress with work, switching away from a task can be disruptive: it slows people down, increases errors and it is challenging to remain focused on work. This thesis investigates how interruption management tools can better support people in managing these types of work-required interruptions in the context of data entry work.

The first part of the thesis reports two qualitative studies looking at understanding data entry in an office setting. They demonstrate that physical interruptions are postponed until a convenient moment in the task if they are expected to take time, but digital interruptions are addressed immediately as these are presumed to be quick to deal with. The second part of the thesis reports three controlled experiments to test the hypothesis that people manage interruptions by avoiding time costs. Results show that if people are able to learn the expected time costs of digital interruptions, they avoid interruptions with a high time cost by reducing the number of these interruptions and postponing them until later in the task, addressing interruptions with low time costs first. The third part of the thesis reports an online experiment and a field study that evaluate a design intervention showing people the duration of their interruptions. These studies demonstrate that showing people the time costs of digital interruptions makes people reflect on what they were doing during an interruption, reduces the duration of interruptions, and made people faster and more accurate in completing data entry tasks.

Taken together, this thesis demonstrates that people manage interruptions based on expected time costs, and that giving people feedback on the time they spend on interruptions can help them manage their interruptions better. It makes a theoretical contribution by showing how people adapt to small changes in time costs by reducing the number and duration of interruptions, and postponing them until later in a task. It makes a practical contribution by showing that giving people feedback on time costs can help them to reduce the duration of interruptions, and improve their focus on the task at hand.

Impact statement

The work presented in this thesis has impacts both inside and outside academia.

Within academia, this thesis has increased our understanding of how time costs affect how people manage self-interruptions to look up information. These interruptions are addressed straight away if they are presumed to be quick and it is considered part of the activity, but it is hard for people to predict time costs outside of a controlled setting. Interruptions can take much longer than expected, which increases their disruptiveness to work. My thesis demonstrates that showing people how long they go away from a specific task helps people reduce the length of their interruptions, which can inform future work on understanding and controlling interruption behaviour. The findings have been published and presented at HCI conferences, universities and research labs.

The findings also make an impact to the methodology of future data entry research, by demonstrating that for some types of data entry work, a major component is collecting data which impacts how data is entered. If data entry interfaces are to be used in situations where information is not readily available, they should be evaluated by requiring participants to collect data from the environment. As part of this thesis I developed a new experimental task, which can be used in future data entry studies to investigate time costs of collecting data during a data entry task.

Outside academia, the findings of this thesis have a direct impact on developers of interruption management tools. This thesis contributes to our understanding of how awareness of time spent on technology can help people manage their use of time. Time management of technology use is a current issue: various technology companies have started to introduce features in 2018 to show how much time is spent in their applications, to encourage users to adopt healthier technology habits. My work contributes to addressing this issue by demonstrating how time awareness can not only help to regulate work-irrelevant technology use, but also to manage task-relevant interruptions during work. As part of my thesis, I have developed and evaluated a browser extension showing how much time the user spends away from a task: the extension is publicly available online to download and use.

Acknowledgements

This is where the acknowledgements will go.

CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Chapter outline

In this chapter, I introduce the problem which forms the motivation of this thesis. I outline the scope and the research questions it aims to answer, give an outline of the structure of the thesis, and present the proposed contribution.

1.1 Motivation

Imagine you have just come back from a business trip and have to claim back your expenses. You collect your receipts from your wallet, open the expenses system, and start entering the items and prices into the system. The prices are in a foreign currency, so you leave the expenses system to go to a currency converter website and convert the prices. You then need to enter a budget code, which was sent via email a few months ago. Upon opening your email inbox, you see a new incoming email that captures your attention. How many times did you stop entering to go and look up certain information? And how long did this take you? You may have entered information that was easy to retrieve first, and left information that would take time to find until the end. And did you retrieve the budget code from your inbox and return to the expenses system straight away - or did you open that unread email instead? Whatever way you chose to complete the task, it involved making decisions on how to manage interruptions to look up required information for the entry task.

Interruptions like these, called inquiries (Jin & Dabbish, 2009), happen when a person goes to look up information to aid the completion of a primary task. These interruptions are common, and many computer tasks require switching between different information sources (Cangiano & Hollan, 2009). The longer and more frequent these interruptions are, the more disruptive they can be: it takes time to resume the main task, it is challenging to remain focused, and it increases likelihood of errors. Furthermore, people can be triggered to further self-interrupt their work for other off-task activities (Jin & Dabbish, 2009). As a result, there now exists a large number of tools that aim to support people in avoiding digital distractions (Lyngs, 2018): for example, a common approach is to temporarily block interruptions (Kim, Cho, & Lee, 2017). However, interruptions to look up information cannot be blocked, as these are needed to progress with the current task. Furthermore, interruptions can also be beneficial: they can boost mood and productivity (Mark, Iqbal, Czerwinski, & Johns, 2014), and when information is found quickly, this can positively impact people's work (Jin & Dabbish, 2009). In addition, a recent study showed that blocking distracting sources at the workplace caused office workers to work longer periods of time without any breaks, with reported higher stress (Mark, Czerwinski, & Iqbal, 2018).

In this thesis, I propose that for inquiry interruptions, it is better to develop tools that allow users to regulate their behaviour themselves, and help them learn and develop useful strategies on how best to allocate their time. To date, there have been insufficient tools suitable to self-regulate inquiries. Commercial tools such as RescueTime and ManicTime track computer usage to allow users to view and reflect on their behaviour, but these tools give information about the user's entire computer usage: interview studies revealed that users feel the data lacks context, and it is often not clear what to do with this data (Collins, Cox, Bird, & Cornish-Tresstail, 2014). Limiting the amount of information has been useful to help regulate non-work activities: a time management tool visualising only the last 30 minutes of users' computer activity reduced time users spent on web browsing and social media (Whittaker, Hollis, & Guydish, 2016). More recently, several technology companies such as Instagram, Facebook and Google have seemed to respond to this trend by making features available showing how much time you spend in their applications, to encourage users to adopt healthy technology habits (Constine, 2018a,b; Lynley, 2018). What is unknown is whether time spent would also affect work-required interruptions, in which several different applications, documents and computer windows may need to be accessed. Information may be a few clicks away, or time has to be spent searching for this information. To what extent are people already aware of the time they spend on interruptions, and do they take this into account when managing interruptions? This thesis investigates how different time costs affect how people manage self-interruptions to look up information for a data entry task, with the aim to inform the design of interruption management tools.

1.2 Aim and scope

The aim of this thesis is to understand how people can better manage the time spent on inquiries needed for a data entry task, with the goal to reduce the disruptiveness of inquiries for the task. To achieve this aim, this thesis addresses the following research question:

How can interruption management tools support people in managing inquiries for a routine data entry task, given variable time costs of required inquiries?

To answer this question, a good understanding is needed of people's current interruption strategies, and to what extent these are influenced by different time costs. Therefore, the research question is addressed by answering the following subquestions across three chapters:

1. How do people manage inquiries for data entry in a finance office setting? (Chapter 3)
2. Do time costs affect:
   - the number of inquiries for data entry? (Chapter 4)
   - the duration of inquiries for data entry? (Chapter 4)
   - the timing of inquiries for data entry? (Chapter 4)
3. Does time feedback affect:
   - the number of inquiries for data entry? (Chapter 5)
   - the duration of inquiries for data entry? (Chapter 5)
   - awareness of time spent on inquiries for data entry? (Chapter 5)

These questions are best answered by using a mixed methods approach. To get a detailed understanding of how people currently manage inquiries, a qualitative approach can get an understanding of how it is conducted in practice as well as people's motivations for their behaviour

(Blandford, Furniss, & Makri, 2016). However, to be able to make claims about whether differences in behaviour are due to different time costs, a more quantitative approach is needed (Cairns & Cox, 2008). Therefore, this thesis uses a combination of the following qualitative and quantitative methods:

- Semi-structured interviews are used to get a grounded understanding of people's data entry work.
- Contextual inquiry interviews, observations and think-aloud protocols are used to get an understanding of how people currently manage inquiries for data entry work and why.
- Laboratory experiments are used to investigate to what extent interruption behaviour can be attributed to a difference in time costs of inquiries.
- An online experiment is used to investigate whether a design intervention that shows the duration of interruptions can affect people's self-interruption behaviour and data entry performance in an uncontrolled setting.
- A two-week field study with semi-structured interviews is used to explore how applicable the design intervention would be for people's own data entry work.

The motivation for choosing these methods are further discussed in the corresponding study chapters. For the scope of this thesis, I primarily focus on office workers in financial administration offices conducting data entry work. This task requires entering different types of information from a variety of sources, such as paper, spreadsheets, emails and databases. It is important the task is done accurately but within a reasonable amount of time as there is time pressure to finish work on time. This task and setting is therefore considered to be an appropriate and interesting example to study further. The task serves as an example of a wider class of routine computer-based tasks, and it can be imagined findings of this thesis can be useful and generalise to other, similar, tasks that require frequently going and out of documents to collect information. For example, people who have to fill in their tax returns have to similarly enter a range of information into a computer system, and have to collect this information from multiple sources with varying time costs.

In this thesis, I am also primarily interested in inquiries, a particular type of self-interruption triggered by the need to look up task-related information. In the qualitative studies, I do consider how participants address interruptions overall, and the proposed design intervention in Chapter 5 may be used for all types of self-interruptions. However, the development of the design intervention was mainly based on the knowledge of how people address inquiries. Different types of interruptions have different triggers and impacts on task performance (Jin & Dabbish, 2009). It was therefore considered important to make a distinction between different self-interruptions, and pose restrictions on the scope of the thesis, to make a valuable contribution.

1.3 Thesis outline

Chapter 2 discusses related literature. It first discusses prior research on interruptions, managing information needs, and technology interventions to address interruptions and information management. It also gives an overview of prior data entry research. This literature review highlights that it is important to get a detailed understanding of data entry work in its situated context, and the type of interruptions that people deal with, in order to support people in managing inquiries during data entry.

Chapter 3-6 describe seven studies conducted to address the research question of this thesis. Chapter 3 describes two qualitative studies aimed to get a detailed understanding of data entry work in a finance office setting. These studies revealed that a major part of data entry is retrieving information from multiple sources with varying time costs. Whereas participants postponed physical interruptions if these were associated with high time costs, people interrupted immediately if they realised they needed digital information. The hypothesis was made that this behaviour was influenced by a difference in time costs, and that participants presumed digital interruptions to have a low time cost. Chapter 4 reports three laboratory experiments to test the hypothesis that time costs influence the timing, number and duration of inquiries for a data entry task. These studies showed that people try to minimise interruptions with a high time cost, and postpone these interruptions to address interruptions with a low time cost first. Chapter 5 describes the development and design of a browser notification which is based on the premise that showing time costs to users may help people manage interruptions outside a controlled setting. The chapter describes an online experiment and a field study to evaluate the effectiveness of this browser notification showing the time costs of interruptions during data entry work. These studies showed that making people more aware of time costs can make them reflect on what they were doing during prior interruptions, reduces the duration of interruptions, and can make people faster and more accurate in completing data entry tasks.

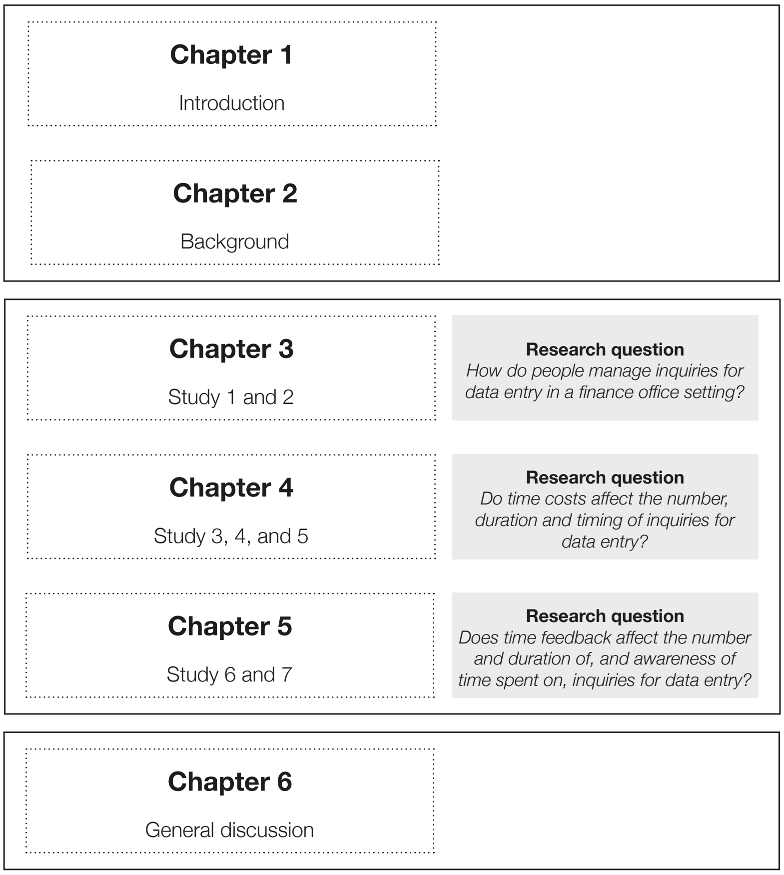


Figure 1.1: Visual overview of the thesis structure.

Chapter 6 summarises the key findings of the studies. It discusses the contribution that these findings make to knowledge and what the practical implications are. It also discusses opportunities for future work.

Figure 1.1 provides a visual overview of the thesis structure, and shows the research question(s) addressed in each study chapter.

1.4 Contribution

The first contribution of my thesis is an increased understanding of the effect of time costs on people's self-interruption behaviour to collect task-related information. Prior work has shown that longer interruptions are more disruptive than short ones, but to date it was unclear how time costs affect people's decisions when and whether to address self-interruptions. My thesis demonstrates that the presumed time cost of an interruption determines whether an interruption is addressed straight away (if it is presumed to be short), or if it is postponed until a more convenient moment in the task (if it is presumed to be long). I also show that in a controlled environment, people are able to learn the time costs of digital interruptions and adapt their strategies to first address interruption with a low time cost, before addressing interruptions with a high time cost. Outside of a controlled setting, these time costs are much harder to predict, and interruptions can end up taking much longer than expected or intended: users may have to go in and out of several computer windows, find the right information within information sources with task-irrelevant information, and get distracted. This means that people cannot always use expected time costs to effectively manage their interruptions.

My thesis also shows that inquiries are handled differently than task-irrelevant self-interruptions: whereas irrelevant interruptions may be ignored, depending on individual differences in ability to self-control interruptions (Lyngs, 2018), inquiries have to be addressed in order to progress with work. Contrary to the idea that focus on an activity can be improved by temporarily blocking distracting sources (Kim et al., 2017), my work shows that inquiries to distracting sources are considered part of the activity, and need different treatment.

These findings have implications for the design of any tools aiming to manage inquiries as well as other self-interruptions and distractions. An applied contribution of this thesis is demonstrating how making time costs of interruptions visible can support people in managing these self-interruptions. Based on the understanding of how time costs affect self-interruption behaviour, a browser notification was developed and evaluated in this thesis. This browser notification showed people how long they go away from their data entry work, to make them more aware of time costs of their interruptions.

Lastly, my work has implications for future data entry research. This thesis has highlighted that for some types of data entry work, a major component of the task is collecting data from various locations. Participants often spent a long time away from the data entry interface to find what they were looking for, which impacts how data is entered: it slows down data entry and increases likelihood of data entry errors. Data entry interfaces should take this into consideration and make it easier to resume a task. If data entry interfaces are intended to be used in situations where information is not readily available, they should be evaluated by requiring participants to first collect data from the environment, to determine how usable they are in this context.

Chapter 2

Background

Chapter outline

In this chapter, I first review previous literature on interruptions, and discuss existing interventions to manage interruptions. The second section discusses research related to finding information as part of work. The last section focuses on prior data entry research.

This chapter provides an overview of previous literature to situate my research on self-interruptions to look up information for data entry. The chapter consists of three sections: I first consider prior research on interruptions, and discuss existing interventions to manage interruptions. This section provides context to the problem of interruptions in the workplace and its effect on task performance, and what factors influence the disruptiveness of interruptions. The section also shows the advantage of using a combination of qualitative and quantitative research methods to study interruptions: while qualitative methods are needed to get an understanding of the type of interruptions that occur and in which context, experiments have been a suitable method to then understand how certain factors influence the disruptiveness of interruptions on task performance. The review of existing interruption management tools gives insight into how other types of interruptions have been managed, and I draw on similarities and differences between inquiries in particular and interruptions in general to argue why current tools are insufficient to support inquiries. The second section discusses research related to managing information as part of work, which demonstrates that the way in which people interrupt a task to look up information depends on the nature of the task. Existing tools to support information management have mostly focused on tasks where information has already been collected and needs to be organised, but does not consider how best to interrupt a task when additional information is needed. The last section focuses on prior data entry research, which highlights that design interventions can reduce data entry errors, but need to be adapted to be suitable for disruptive work environments.

2.1 Terminology

Throughout the thesis, I use the key terms inquiries, interruptions, time cost and information access cost. To avoid confusion, I clarify the definitions of these terms here.

Inquiries are a type of self-interruption, and happen when a person goes away from a task, to look up information that aids the completion of that primary task (Jin & Dabbish, 2009). This type of interruption is the main focus of the thesis. In sections where the terms *interruptions* or *self-interruptions* are used, rather than inquiries, it can be assumed this section refers to all types of interruptions.

Time costs refer to the time involved to complete an action. Prior studies in cognitive psychology who have studied how people use task information have used the term *information access cost*, which is a particular type of time cost, and refers to the time, physical and/or mental effort required to access information (Gray, Sims, Fu, & Schoelles, 2006). Because this thesis looks at a broader definition of time costs to look up information, which can be caused by accessing information, but also searching for information and getting distracted by other information, the term time costs is used for the majority of the thesis. When the term information access cost is used, usually when referring to prior work, it refers to the specific time cost of accessing information.

2.2 Interruptions and fragmentation of work

Computer work frequently gets interrupted: on average, office workers switch between activities every three minutes (Gonzalez & Mark, 2004). Furthermore, tasks themselves are also often fragmented: people have to switch between documents and applications to look up information for their task. These interruptions can be disruptive: it can slow people down, increase errors and cause stress. However, some interruptions can also be beneficial: short breaks can improve mood and restore energy (Mark et al., 2014), and quickly retrieving relevant information can aid completion of a task (Jin & Dabbish, 2009). A range of both field studies and controlled experiments have tried to understand what factors influence disruptiveness of interruptions.

2.2.1 Interruptions during computer work

Field studies have been done to understand interruptions in situated settings such as healthcare (e.g. Grundgeiger, Sanderson, Macdougall, & Venkatesh, 2010; Westbrook, Woods, Rob, Dunsmuir, & Day, 2010) and office environments (e.g. Czerwinski, Horvitz, & Wilhite, 2004; Gonzalez & Mark, 2004). For the scope of this thesis, I mainly discuss work on interruptions during computer-based work. Mark, Gonzalez, & Harris (2005) observed office workers, and found half of all interruptions are self-interruptions, and that self-interrupted tasks were less likely to be resumed later than tasks that were stopped by an external interruption. An interview study on interruption management strategies found major differences in the level of difficulty for users to manage external versus self-interruptions (Kim et al., 2017). Whereas external interruptions may be ignored or deferred, self-interruptions require more self-control, and are experienced as harder to resist and as more distracting. Furthermore, self-interruptions take more time to recover from than external interruptions as they can end up taking much longer than planned. When switching between computer windows, there are numerous opportunities to get distracted and get diverted from the main task. For example when switching to communication tools, users can get tempted to answer unrelated messages instead (Mark, Voida, & Cardello, 2012).

There are also individual differences in people's tendency to attend to distractions and resist interruptions (Lyngs, 2018; Mark, Iqbal, Czerwinski, Johns, & Sano, 2016b). Mark et al. (2016b) conducted a field study with office workers, in which window switching behaviour was measured, and participants were asked to fill in a personality survey at the end of each day. They found a positive relationship between how people scored on a Neuroticism and Impulsivity scale, their switching behaviour, and how productive they felt at the end of the day. This suggests that distractibility could be a personal trait.

Jin & Dabbish (2009) conducted an observational study of self-interruptions during computer tasks. They identified seven categories of self-interruptions: adjustments, breaks, recollections, routines, triggers, waits, and inquiries. Adjustments happen when people try to adjust and improve their work environment, for example by closing irrelevant documents. Breaks are moments when people switch to something else because they want to take a rest from the main task. Recollections occur if people remember they need to perform another task. Routines are self-interruptions that happen out of habit, such as checking social media regularly. A trigger is an external stimulus that triggers the user to switch to another task. Waits are when people switch to something else during a delay in the primary task. Lastly, an inquiry, which is the type of interruption I focus on in this thesis, happens when a person goes to look up information to aid the completion of a primary task. Some of these interruptions may have a positive effect. For example, breaks can be positive: ICT is increasingly used in work breaks, and these self-interruptions may restore energy (Skatova et al., 2016). Furthermore, in Jin and Dabbish's study inquiries positively impacted people's work, as they quickly found what they were looking for. However, Jin and Dabbish speculate that these interruptions may be disruptive if information cannot be found straight away. Overall, all types of interruptions may become disruptive if they happen too often or for too long (Mark et al., 2018).

Given the occurrence of interruptions, studies have tried to understand the consequences of interruptions on work productivity. Studies focusing on university students found that students who made fewer interruptions to social media during class and study sessions had higher grades (Carrier, Rosen, Cheever, & Lim, 2015). While grades can be used to measure students' performance, it is more difficult to objectively measure productivity in the workplace, and most workplace studies have relied on participants' self-assessment of productivity. For example, Mark, Iqbal, Czerwinski, & Johns (2015) instructed participants to assess how productive they felt on a 7-point Likert scale. Self-assessments of productivity were then compared with objective measures of people's screen switches, which were obtained using logging techniques. It was found that the more screen switches people made, the less productive they felt at the end of the day (Mark et al., 2015). These studies suggest that fragmented attention has a negative effect on work productivity, though to study a direct relationship between interruptions and work performance, experiments have been used, which will be discussed next.

2.2.2 Experimental investigations of interruptions

Field studies have given insight in the prevalence of interruptions, the context that leads up to an interruption, and the different type of interruptions that happen. A limitation of qualitative methods is that it is difficult to establish a direct link between various factors influencing the disruptiveness of interruptions. Controlled experiments on the other hand are a useful method to measure the effect of interruptions on task performance.

Two factors that contribute to the disruptiveness of an interruption are its length and timing. This can be explained by the Memory for Goals theory (Altmann & Trafton, 2002): this theory holds that when people are interrupted from a task, the representation of the task in working memory enables them to keep track of where they were in the task, making it easier to resume the task after the interruption. The longer people are interrupted and away from a task, the more likely that this representation weakens and fades from memory (Altmann, Trafton, & Hambrick, 2017; Monk, Trafton, & Boehm-Davis, 2008). Furthermore, it is more disruptive if people are interrupted at high than low workload moments, because people have to hold more information in memory while being interrupted. For instance, interruptions have been shown to be more disruptive if they happen in the middle of a subtask compared to when they happen between subtasks, because people have a higher workload of remembering where they are in the middle of the subtask (Gould, Brumby, & Cox, 2013; Iqbal & Bailey, 2005). Salvucci & Bogunovich (2010) conducted an experiment investigating how people manage interruptions that are deferrable, and found people deferred interruptions until moments of low workload. This suggests that people may be fairly good at focusing on a task and postponing interruptions, though the study only looked at external interruptions, which were triggered by notifications. As discussed above, self-interruptions are perceived to be harder to ignore than external interruptions. Furthermore, in Salvucci and Bogunovich's study people did not have to remember to attend to a task later, and had their working memory free to continue to focus on the main task. Self-interruptions, on the other hand, are often triggered by the users' internal thought (Jin & Dabbish, 2009): holding the intention to interrupt in working memory may be just as cognitively effortful as interrupting a task at a high-workload moment.

Interruptions can be useful if they benefit the task (Jin & Dabbish, 2009), and a range of studies have shown that irrelevant interruptions are more disruptive than relevant ones (e.g. Adamczyk & Bailey, 2004; Gould et al., 2013). Based on this finding, it may seem more important to focus on avoiding irrelevant interruptions. However, task-relevant interruptions can be just as distracting and disruptive if not managed well. Iqbal & Bailey (2008) conducted a lab study, in which participants were interrupted by relevant or irrelevant notifications during work. Selecting a notification took them to a web page with more information. Even though participants wanted relevant interruptions to happen more often than irrelevant interruptions, for both types of interruptions participants were observed entering chains of diversions. This refers to moments where the participant does not return to the main task after an interruption, but instead attends to other activities. This means that both relevant and irrelevant interruptions may trigger people to spend longer on interruptions than necessary.

2.2.3 The effect of information access costs on interruptions

People's interruption behaviour is further influenced by time costs associated with making an interruption. Several studies have looked at how information access costs (IAC), that is the physical, mental and cognitive effort to access task information, influence how people switch away from a task to look up task information. Even though most of these studies do not explicitly label these switches as 'interruptions', window switches can be disruptive as it fragments people's attention (Rule & Youngstrom, 2013). In a typical IAC study, participants are asked to complete an experimental task, such as solving the Tower of Hanoi puzzle (Waldron, Patrick, Morgan, & King, 2007), programming a VCR (Gray & Fu, 2004) or copying a pattern of coloured blocks (Gray et al., 2006). To complete the task, the participant needs to access task information. In the control condition, this information is easily accessible. In experimental conditions, there is a cost to access the information, for example the information is covered by a grey mask, and participants have to hover over the mask with their cursor to reveal the information. A consistent finding across IAC research is that if the cost to access task information increases, people try to minimise time costs by making fewer switches to the information source, and instead rely on information in memory. This adaptive use of memory is explained by the soft constraints hypothesis (Gray et al., 2006), a cognitive theory which holds that people adapt their cognitive strategies to the constraints of a task environment with the aim to optimise task completion time: rather than preserving cognitive resources, people try to minimise time.

Though a memory-based strategy carries the risk that the memorised information is incorrect, several studies have shown that an increased IAC can also have a positive effect on task performance. In a problem-solving task, an increased IAC resulted in people taking the time to carefully memorise task information and plan actions before making any moves, which made them more efficient in completing the task (e.g. Morgan & Patrick, 2012; Morgan, Waldron, King, & Patrick, 2007). A memory-intensive strategy can also be useful for resuming a task after an interruption. Morgan, Patrick, Waldron, King, & Patrick (2009) conducted a study looking at the effect of IAC on a copying task. People had to perform the Blocks World Task (BWT), which involves copying a pattern of coloured blocks, by dragging blocks from a resource window to a target window. They manipulated the cost to access the original source which showed the pattern they had to copy. In the Low IAC condition, the pattern was permanently visible on the screen. In the Medium IAC condition, the pattern was covered by a grey mask and participants had to hover over the mask with their mouse to reveal the pattern. In the High IAC condition, there was an additional time delay before the pattern was revealed. At certain intervals, they would get interrupted and asked to do a secondary task, irrelevant to the primary task. As IAC increased, people made fewer but longer visits to the target pattern and memorised more of the pattern. As a result, following the irrelevant interruption they were faster to resume the primary task, and could copy more blocks before having to revisit the target pattern. This again shows how a strong representation of the task in working memory aids resumption after an interruption.

The soft constraints hypothesis assumes a situation where the user only makes switches between task information and the main task. In this context, a longer interruption time may be used to encode task information in memory, with the effect that people are quicker to resume a task and more efficient to complete it (Morgan & Patrick, 2012; Morgan et al., 2007). In reality, people may need to make several window switches before they retrieve the right information, and people may further switch to other unrelated tasks. In this context, a long interruption time may be caused because people are diverted by other tasks and not attending to the task at all (Iqbal & Bailey, 2008), which can increase errors. As such, while the theory is useful in explaining how people adapt their use of memory to information access costs, it is difficult to map the theory to a broader range of time costs of interruptions in the workplace.

2.2.4 Interruption management tools

Given the occurrence of interruptions and its potential negative effect on work performance, there have been different approaches to support self-interruption management and improve people's focus.

Commercial applications, such as RescueTime and ManicTime, provide users with an overview of all their computer activities, to increase awareness of their use of time. Users can view how much time they spent in particular documents, websites and applications, and during which hours of the day. Little work has evaluated how effective these applications are in improving focus, and interview studies have reported a lack of engagement among users (Collins et al., 2014; Whittaker et al., 2016). An interview study by Collins et al. (2014) on understanding people's use of RescueTime found four barriers to explain people's lack of engagement with the data: the data lacks salience, a lack of context made it difficult to extract work patterns from the data, participants felt it was not a true representation of their actual activities, and they were not sure what actions to take based on the data.

Whittaker et al. (2016) interviewed office workers and students to establish user requirements for a time awareness application, and found users were primarily interested in their current activities rather than long-term behaviour. Therefore, they developed and evaluated an application which presented users with a visualisation of the last 30 minutes of computer activity. The application reduced the time spent in email, browsing and social media, but it did not increase time spent on work and it was unclear whether it improved people's productivity. Whittaker et al. speculated that participants may already have time limits to spend on work, but are more flexible with the amount of time they spend on other online activities.

Other commercial tools such as StayFocusd, Freedom and FocusMe limit access to specific sources. Kim et al. (2017) developed an intervention that allowed people to block applications and websites across devices for a fixed period that they considered distracting. The blocking feature was viewed positively by participants who found it difficult to mitigate self-interruptions themselves. However, many distracting sources, such as web browsers and instant messaging applications, could not be blocked during work because these need to be accessed for the current work task. To investigate how appropriate a blocking approach would be in the workplace, Mark et al. (2018) conducted a field study with office workers using blocking software for one week. Participants installed software that allowed them to disable websites, and were asked to block any websites they considered distracting and nonessential to work. Several participants disliked the feeling that the software was controlling them, and rather wanted to learn how to gain control themselves over their work and interruptions. Furthermore, with distractions blocked, several participants ended up working longer periods of time and forgot to take any breaks. As a result, they reported higher stress during work.

Other interventions suggest that making people aware of their interruption behaviour during a specific task may help focus. Gould, Cox, & Brumby (2016) looked at people's switches to unrelated activities during an online data entry task. They found that an intervention that encouraged people to stay focused after they had self-interrupted reduced the number of switches to unrelated tasks.

2.2.5 Summary

Interruptions in the workplace are common, and both controlled and field studies have found a link between fragmented attention and a decrease in work performance (Bailey, Konstan, & Carlis, 2001; Carrier et al., 2015). The more frequent and longer an interruption is, the more disruptive it is. Controlled studies have shown that as the cost to access information increases, people reduce the number of interruptions to this information. Various tools have been developed aiming to reduce interruptions in natural settings, by giving reflective information on people's behaviour or blocking interruptions during focused work. These tools may be useful to manage non-essential interruptions, but are insufficient to manage interruptions which are required for work, such as looking up task-related information.

2.3 Managing information needs

Though applications are usually designed assuming users stay within a single application for a task, task information is often spread across sources, which requires users to interrupt their work and switch between multiple applications and documents to retrieve information (Cangiano & Hollan, 2009; Czerwinski et al., 2004; Mark et al., 2005; Sellberg & Susi, 2014). To get a better understanding of how users can be supported in this fragmentation of work, a line of studies have studied how people find and re-find information as part of work in healthcare (Reddy & Dourish, 2002) and law offices (Cangiano & Hollan, 2009; Makri, 2008), on a mobile device (Sohn, Li, Griswold, & Hollan, 2008), and how information behaviour differs for different types of tasks (Bondarenko & Janssen, 2005). For the purpose of this thesis, I limit my discussion to how information needs are addressed in practice to aid completion of a particular activity, rather than theoretical models of information seeking as an activity in itself.

2.3.1 Managing inquiries for work

Several studies have used interviews and observations to get a detailed understanding of people's information behaviour at work. For instance, Reddy & Dourish (2002) conducted an ethnographic study in an intensive care unit of a hospital. The focus of this study was on how health-care workers managed and retrieved the information sources needed to support their work throughout the day. Health-care workers used their own and their colleagues' working patterns to plan for future information needs. This planning was necessary because of the nature of work: workers were on the move for most of their working day, they had to get information sources from different locations within the unit, and were reliant on other people to get access to information. These physical constraints and dependencies encouraged workers to think about whether to collect information right away, postpone it until later, or strategically request it for later access. In contrast, office workers are often situated behind a desk, and information may be only a few clicks away. It is uncertain whether office workers also engage in this kind of planning behaviour. Cangiano & Hollan (2009) interviewed and observed people in law offices searching for information. They found that lawyers do not plan for future information needs, but instead have to spend time retracing the steps of a legal case, and that they need to access different sources to find all the relevant information for the case, such as emails, instant messages with colleagues, legal documents, and written reports. They argue that, even though work may appear fragmented, these participants still perceived they were working on the same activity when switching between all these different sources.

2.3.2 Factors influencing inquiries

While some studies have observed that people plan inquiries carefully around their work (Reddy & Dourish, 2002), other studies have shown that people interrupt their task immediately as soon as they think of information they need (Jin & Dabbish, 2009). To get insight into what factors may influence information behaviour, Sohn et al. (2008) conducted a diary study investigating how people address inquiries on a mobile phone. They did not focus on a particular task, but asked participants to keep a record of all instances where they needed information for an activity they were doing. Four main factors determined whether participants looked up the information immediately or whether they postponed to address it later: urgency, importance, situational factors, and costs. The more urgent and important it was to have the information for the activity they were doing, the more likely it was they looked up the information at the moment they realised they needed it. If they were currently involved in an activity that made it difficult to address the information need at that moment, or they did not know where to get the information from, they were more likely to postpone the inquiry. Lastly, the more time or monetary cost was associated with getting the information, the more likely they were to not address it or leave it until later. This suggests that time costs may make people less likely to interrupt a task, although Sohn et al.'s study exclusively looked at information needs on a mobile device, which differs from desktop search: the affordance of easy switching between windows on conventional desktop computers may give the false impression that information is easier to access (Sellen & Harper, 2003). Furthermore, participants mostly reported non-work situations in which information was not essential: the majority of information needs were categorised as 'trivia'. As such, participants may have been more likely to give up an information search if it required effort, compared to if this search was required for work: for 30% of the diary entries information was not accessed later at all, with the main reason being that it was unimportant.

2.3.3 Types of tasks

The way in which people address inquiries is further influenced by the type of task. Bondarenko & Janssen (2005) distinguish between two different types of tasks that information workers engage in: administrative and research tasks. Administrative tasks are routine tasks, of which the steps are usually the same, and are characterised by short and frequent switches between documents. For these tasks, documents often switch between what Bondarenko & Janssen refer to as 'hot', when a document is needed, and 'cold', when a document is not needed and archived. For administrative tasks, documents are only needed for a short period of time, but need to be accessed repeatedly. Examples are filling in a personal information form for a new member of staff. Research tasks, on the other hand, only require a small number of documents, but these need to be accessed for a longer period of time. An example is writing a research paper: the writer may need to consult another paper, and need to read this or keep it close by for a long period of time.

Because of the different nature of these tasks, they need different support for accessing and managing information. Bondarenko & Janssen (2005) conclude that most information management tools only support 'pure' administrative tasks, which are repetitive, structured and predictable, but provide inadequate support for research tasks. However, these two task types should be seen as two extremes rather than a distinct classification, and many tasks may fall somewhere on the spectrum between these two extremes. For example, though data entry shares some of the characteristics of an administrative task and usually follows the same sequence of steps, it does not always follow the same linear path, and as will be demonstrated in the Chapter 3 of this thesis, it does not always require the same information.

2.3.4 Information management tools

Given the fragmented nature of many tasks, several tools have looked at how to make information easier to access and reduce interruptions, so people can focus on their work. For example, GroupBar (Smith et al., 2003) makes it possible to group windows needed for a task in the task bar. Cangiano & Hollan (2009) developed ActivityTrails, a tool which allows people to play back a visual summary of sources which were accessed the last time they were working on an activity. These tools can be useful for work which has to be resumed later on, in which the same sources need to be accessed again, but may be less useful when starting a new activity, in which new sources are needed.

Increased screen space has also been explored as a way to support people in fragmented work, as people will not need to flick back and forth between windows as often, but can have these open and on the screen simultaneously (Czerwinski et al., 2003). Bi & Balakrishnan (2009) conducted a week-long diary study looking at how office workers utilise a large screen and a second screen to access information. Participants were given a large or second screen for the study, and had no prior experience working with this set-up. People arranged information on their available screen space differently in each condition. For two screens, people dedicated one screen for their primary task which filled up the entire screen. They moved all information they did not need at the moment to the second screen and did not bother re-arranging the windows: they would only attend to the second screen when they needed information, and deliberately allocated the second screen for a different purpose than the first screen. For one large screen, participants first spent a certain amount of time optimising the layout of the windows by resizing and re-arranging them. They put all windows needed for the primary task in the center of the screen and placed other windows in the periphery. As participants needed information from this periphery, they dragged the window to the center of the screen rather than interacting with that particular part of the screen. On the other hand, if participants needed information from the second screen, they physically turned to the second screen and interacted with it but did not drag the information to the primary screen, unless they had to interact with it for a longer time.

Though people had to spend some initial time organising information, office workers felt more focused on the task and immersed in their work when surrounded by task-relevant documents. In this study, participants were given large screens for the purpose of the study, so there may have been a novelty effect. In practice, dual screens have been shown to also increase multitasking and fragmented attention (Robertson et al., 2005). In addition, a large screen may reduce the disruption of switching between windows but it can introduce another type of cost, which Robertson et al. (2005) refer to as the *'distal information access cost problem'*: as screen size increases, it becomes harder and more time-consuming to target and select certain buttons and windows.

Despite the popularity of one large screen, Grudin (2001) states dividing screen space up in multiple monitors can sometimes be better. He argued that the main benefit of having a second screen is not so much the increase in screen space, but the partitioning of information into dedicated areas: having multiple screens prompts people to think more about where to put which information. To support his argument, he conducted a field study looking at how office workers use multiple screens to arrange information. Participants positioned information they did not need at the moment on a second screen where they were not distracted by it, but could easily access it when needed. People preferred that information was always in the same known location and referred to the second screen whenever they needed to look information up, even when they were aware they could also access it using their primary screen as well, where the information was sometimes less time-consuming to access (Grudin, 2001). While increased screen size may be useful to organise task information once it has been retrieved, the studies provide little guidance on how to switch between the primary task and retrieving new information.

2.3.5 Summary

Many computer tasks require switching between different documents, applications and windows. How people switch depends on the type of task, the cost to access information, the urgency and importance of an information need, and the current situation people are in. Several tools have tried to make it easier to group information and search, so people can focus on the main task. However, these solutions are best suited for situations where information is known beforehand and needs to be re-used, and does not consider how best to interrupt a task when additional information is needed. An increased screen size can reduce certain access costs, such as mouse clicks to flick back and forth between windows, but may introduce other costs, such as time costs to select the right window, and can increase multitasking to unrelated tasks.

2.4 Data entry

The way in which people interrupt to look up information can depend on the type of task (Bondarenko & Janssen, 2005), so in order to get a detailed understanding of interruption behaviour, it is not only important to consider the task environment but also the scope of the task studied. This thesis focuses specifically on data entry, a core computing task in office scenarios. An example of a data entry task is completing a payment request form, in which information has to be entered into specific fields of the form, such as the person's name, address, and bank account details. Interruptions during this type of task can be particularly disruptive as it not only takes time to resume the task, but it can also increase data entry errors. The consequences of data entry errors can range from an inconvenience to having a more serious outcome: in 2018, up to 1,500 junior doctors' job offers in the UK were withdrawn after a data entry error in a spreadsheet caused the candidates to incorrectly receive a job offer (BBC, 2018). It is therefore important to study how errors can be reduced, and a line of experiments have looked at understanding the different steps of a data entry task, and have demonstrated how making design changes to a data entry interface can reduce errors.

A data entry task can be broken down in four stages, as is shown as a diagram in Figure 2.1. An information source contains the input, i.e. the data to be entered. In the perception stage, the user perceives this data. In the encoding stage, the user encodes these in the mind. In the execution stage, the user enters them into a device which produces the output of the task, namely the data entered. Additionally, there can be a checking stage where the user checks their entered output against the original input to see if it matches. The following four sections will briefly describe each stage of the task in turn, and will discuss research that has been done to reduce errors at this stage.

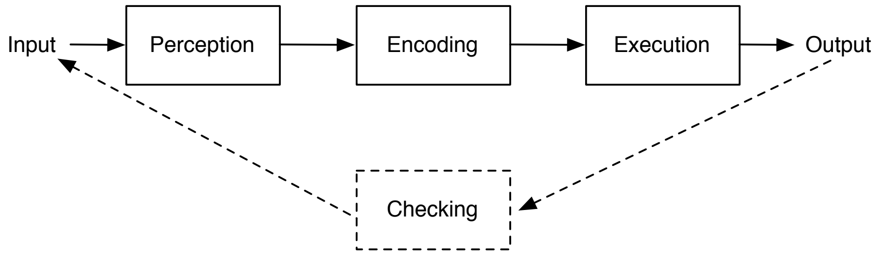


Figure 2.1: The different stages of a routine data entry task: a user perceives data input, encodes these in the mind, executes certain actions to enter data and produce the output, and can check the output against the original input.

2.4.1 The perception stage

A data entry task begins with the user looking at the data that has to be entered on a data source. The way in which data is presented to the user affects how that data is encoded by the user in the internal mind. The stronger data is encoded in memory, the more robust the user is against interruptions, and the less often the user may have to interrupt a task to look back at information. Several studies have shown that making information more difficult to perceive can encourage a deeper encoding in memory (Diemand-Yauman, Oppenheimer, & Vaughan, 2011; Soboczenski, Cairns, & Cox, 2013). Soboczenski et al. (2013) conducted two experiments where people had to transcribe text and numbers that were presented either in a black font colour or a harder-to-read grey font colour. A hard-to-read font forced people to make more of an effort to read and understand the text, and as a result the text was more deeply processed and encoded in the mind. Participants made fewer data transcription errors if data was shown in the harder-to-read font colour, both for transcribing text and numbers. There was no difference in speed between the hard-to-read and normal font colour, suggesting that the improved accuracy was not due to a speed-accuracy trade-off.

The way in which data is perceived and encoded is further influenced by how it is situated in the environment. The distributed cognition approach has been used as a theoretical framework to explain how cognition is 'distributed', meaning that people use a combination of internal information in their mind, and external information in the physical environment, to carry out work (Hollan, Hutchins, & Kirsh, 2000; Hutchins, 1995). This means that to be able to understand how people work, it is not enough to know how the mind processes information, but it is also important to know how task information is situated in the physical world (Hollan et al., 2000). In most data entry studies, the data to enter is given to participants, which might not reflect how data entry is situated in offices. This motivates the need to understand how information for data entry work is spread in an office environment, which is the focus of Chapter 3 in this thesis.

2.4.2 The encoding stage

After the user has perceived the input source in a data entry task, the next stage in the task sequence is the encoding stage, where data is encoded in memory. Prior work has shown several benefits of a deep encoding in memory: it can reduce interruptions to look back at information, but can also make people more robust against other, task-irrelevant, interruptions. As discussed earlier, another factor that influences encoding is the costs associated with accessing data in the task environment. If information access costs increase, people rely on information in memory to avoid incurring costs and avoid making switches to the information source. Morgan et al. (2009) showed that in a copying task, if IAC was increased, participants made more of an effort to memorise the information. After an interruption, they were quicker to resume the task, because the information was still in memory.

It is not just the external representation that influences how strongly something is encoded: in certain settings, certain data items are often re-used and thus are more strongly represented in memory. From text entry literature, it is known that words are easier to transcribe than non-words as they are used more often, they are more meaningful and thus have a stronger representation in memory (Salthouse, 1986). This highlights the importance to get a thorough understanding of the setting in which data entry is conducted. For example, Wiseman, Cox, Brumby, Gould, & Carroll (2013) looked at number entry in hospitals and found some numbers were used far more often than others. Experiments showed that familiar numbers are faster to transcribe, suggesting that these are more strongly represented in memory than random numbers as well (Wiseman, n.d.). This means that data entry interfaces that are intended for specific settings should be evaluated with familiar numbers used in that setting, rather than random numbers, and make it easier to enter commonly used numbers.

2.4.3 The execution stage

The third stage of the data entry task is the execution stage, which is the stage where the user performs the motoric actions to enter data into a device.

The design of the input method influences the speed with which users enter data, which can

(a) A number pad. **(b)** An incremental interface.

Figure 2.2: Two different number entry interfaces tested in Oladimeji et al.'s study.

subsequently affect errors. Oladimeji, Thimbleby, & Cox (2011) compared a number keypad with an incremental interface. The two types of interfaces are shown in Figure 2.2. The number keypad is most common, and is used on calculators and phones. In this interface, each digit is assigned a button and additional buttons are usually a decimal point and a delete key to correct an error, as shown in Figure 2.2a. In an incremental interface, a number is entered by increasing or decreasing the number using up and down keys. The incremental interface used in Oladimeji et al.'s study is shown in Figure 2.2b. The double arrows increase and decrease the number by a larger amount than the single arrows.

Results of the study showed that a number keypad allowed people to enter a number more quickly than an incremental interface, but more errors were made. With the keypad, the visual attention was more on the input keys than the display. In an incremental interface, people were changing an existing value rather than entering a new value, so they had to look at the display to see how their actions changed the current value. This attention on the display may have made it more likely for them to detect errors in time. While an incremental interface may not be feasible when entering large amounts of data as it will slow users down too much, it may be preferrable over a keypad in situations where accuracy is of great importance (Thimbleby, 2011).

2.4.4 The checking stage

Most data transcription models consider the execution stage as the final stage of a data entry task (Card, Moran, & Newell, 1983; Salthouse, 1986), but an additional stage can be a checking stage, where people review what they have entered and compare it with the original input, to see if it is correct. A reason why most models do not include this stage may be that people often do not make the effort to check, and if they do, they are poor at detecting errors (Olsen, 2008).

Olsen (2008) conducted a lab experiment in which he simulated an internet banking tool, and participants were asked to enter account numbers from a paper sheet into a computer. After participants had entered an account number, they were presented with a confirmation screen with the input, and users were asked to check their input on this screen before submitting. Participants confirmed 88 trials where they had entered an incorrect account number. In addition, in 178 trials the simulator changed people's input to another number and this incorrect number was presented on the confirmation screen. Only 5 of these 178 errors were detected and corrected. This large amount of incorrect confirmations again suggests users do not check properly, even if they are explicitly asked to do so. People are even worse at checking their input if they are switching between a data entry task and another task (Wiseman, Borghouts, Grgic, Brumby, & Cox, 2015).

Given the limited effectiveness of confirmation screens, some studies have supplemented these with lockouts, where users have to wait a short period of time before they are able to confirm and submit their input. While this has been shown to reduce errors in a controlled setting (Gould, Cox, Brumby, & Wickersham, 2016), the presence of other tasks and distractions can entice users to switch to something else instead (Gould, Cox, Brumby, & Wickersham, 2016; Katidioti & Taatgen, 2013).

Gould, Cox, Brumby, & Wickersham (2016) studied a number entry task where after each number the submit button would be disabled for a number of seconds, and a text instruction to check input appeared on the input screen. This lockout was an effective method in encouraging people to check and detect errors in a lab setting. When the study was replicated online, a short lockout made people detect errors as well but the longer the lockout duration was, the more likely people were to switch to doing other tasks, and not check anymore. This illustrates the importance of taking the task context into account, and suggests that findings from controlled studies do not always directly translate to an applied setting (Gould, Cox, Brumby, & Wickersham, 2016).

Similar switching behaviour was found by Katidioti & Taatgen (2013). They conducted a lab experiment where people had to copy information and were interrupted by chat messages. Participants were free to choose when they wanted to attend to the messages. When people were locked out in the copying task and had to wait 3 seconds before they could enter the information, they often switched to the chat message, which made them forget the information to copy and slowed them down in completing the task.

2.4.5 Summary

Research has shown how the way data is perceived, encoded, entered and checked all influence data entry performance. Time costs associated with perceiving data can improve encoding, and a better encoding leads to faster and more accurate entry. The majority of data entry design interventions have been evaluated through laboratory experiments, and attempts to study data entry in a multitask setting suggest that people may interact with these interventions differently beyond a controlled setting. For example, in experimental studies the information was given and people were focused on the task, whereas many office settings are fragmented, and people can get distracted. How is data entry situated in an office setting? Are people focused on the task? And do they have the information readily available, or is this fragmented? In order to understand how people can be supported in making inquiries for data entry, it is important to understand how information for data entry is distributed in the environment.

2.5 Conclusion

This chapter reviewed literature on interruptions, managing information needs and data entry. From this review, we learn that the length and timing affects the disruptiveness of interruptions, and that people try to avoid time costs in a controlled setting. As task-irrelevant interruptions are considered more disruptive than task-relevant interruptions, interruption management tools have mostly focused on avoiding these, even though relevant interruptions can be just as disruptive if not managed well. Furthermore, based on prior literature on information management we learn that how people search for information depends on the type of work. While existing tools support tasks where information sources need to be re-used, these provide little guidance on how to interrupt a task to collect information in the first place. Lastly, data entry research has shown that design interventions can reduce data entry errors if people are focused on the task, but that the context needs to be considered: people have workarounds to these design interventions if they interrupt and are exposed to distractions.

This highlights a need to better understand inquiries for a data entry task, and how the disruptiveness of these inquiries can be reduced. The next three chapters report a series of studies aimed to understand how interruption management tools can support people in managing inquiries for a routine data entry task, given variable time costs of required inquiries. To design a suitable tool, an understanding is required of how people currently self-interrupt for data entry work, which will be the focus of the next chapter.

CHAPTER 3

UNDERSTANDING DATA ENTRY IN AN OFFICE SETTING

Chapter outline

This chapter reports an interview study on understanding data entry work in a naturalistic office setting, and a contextual inquiry study on how people self-interrupt in this setting. Participants try to avoid task-irrelevant interruptions, and postpone physical inquiries if they expect them to take a long time, but digital inquiries are addressed immediately. Together these studies suggest that people do not always know how to effectively manage self-interruptions if they are seen as part of the same activity, because of preconceptions about ease of access, and because of lack of awareness of time spent on digital interruptions.

3.1 Introduction

Bondarenko & Janssen (2005) showed that the type of task influences how people manage information and therefore might impact how they address inquiries. To address the aim of this thesis, which is to understand how people can better manage the time spent on inquiries needed for a data entry task, it is therefore important to not only look at people's data entry performance using a well-structured task, but also at data entry in the environment in which these tasks are normally performed. The aim of the studies reported in this chapter was therefore to get a detailed understanding of data entry work in a finance office setting and how people currently manage inquiries for this type of work. In particular, the aim of Study 1 was to get a grounded understanding of data entry work, the physical environment, and the type of information sources needed for this type of work. Office workers were interviewed at their workplace about their data entry work and asked to demonstrate a typical data entry task. The aim of Study 2 was to understand the different time costs associated with information sources, and whether this affected people's self-interruption behaviour.

3.2 Study 1: Understanding data entry work in a financial office

This study and its results have been published in Borghouts, Brumby, & Cox (2017) and were presented at the European conference on Computer-Supported Cooperative Work in 2017.

3.2.1 Introduction

As data entry is a common task and it is important this is done both accurately and efficiently, work has been done to design and optimise data entry interfaces to support fast and accurate data entry (e.g. Oladimeji, Thimbleby, & Cox, 2013; Vertanen, Memmi, Emge, Reyal, & Kristensson, 2015; Wiseman et al., 2013). Studies have shown that creating interfaces to slow down data entry (Gould, Cox, Brumby, & Wickersham, 2016), by requiring additional information (Wiseman et al., 2013) or using alternative input technology (Oladimeji et al., 2011) can all reduce error rates in the lab. However, these solutions do not always work outside of the lab (e.g. Gould, Cox, Brumby, & Wickersham, 2016), as it is not just the data entry interface that determines efficiency and accuracy but also other aspects of the task, such as the environment within which it is conducted (Payne & Howes, 2013; Randall & Rouncefield, 2014). For instance, in lab studies, users are given clear instructions and are given the data to enter. In everyday computer use, data entry tasks might not be so clearly prescribed (Evans & Wobbrock, 2012). To illustrate, Evans & Wobbrock (2012) investigated if people's data entry behaviour in a lab setting was comparable to how they would normally perform these inputting tasks in their everyday life. They remotely observed people's input behaviour on their personal computer, and compared this with their performance on similar tasks in a lab. Participants installed a tool on their personal computer which logged all data entry and mouse pointing behaviour they performed in one work week. Examples of tasks that were carried out were sending personal messages to friends and browsing the web. There were no differences in uncorrected errors or data entry speed between the lab and the field, but they did find that participants corrected more errors in the lab. This study shows that people check and correct their entries more when they are in a controlled environment and are focused on the task, though the measured behaviour on people's personal computers mostly included tasks where accuracy may not have been considered important, such as sending an informal chat message to a friend.

The aim of Study 1 was to get an understanding of people's data entry work in an office setting. As the nature of this first study was exploratory, interviews were considered to be an appropriate method for this purpose. Participants were able to further explain their strategies and discuss challenges they experienced. Furthermore, it enabled to collect participants' experience with past critical incidents, which may not be captured in observational sessions. The study explored people's data entry work overall, and the focus of this study was not specifically on people's information behaviour. As data collection and analysis progressed, information collection was found to be a large and integrated part of people's data entry work, which could potentially influence their performance. After this study, people's information behaviour became the focus of the next study.

The user group of this study were office workers at finance administration offices at two public universities, who conduct data entry tasks as part of their daily work. This user group was chosen as they have a lot of data entry tasks as part of their job, and it is an area where it is important to enter data accurately, but there is also time pressure to finish work on time. Furthermore, it was an accessible user group to approach.

3.2.2 Method

Participants

Nine participants (four male) took part in the study. They were employees from two public universities and their work involved receiving various requests for payment, checking the information of these requests was correct, and entering the information along with administration data into computer systems. Ages ranged from 18 to 52 (two participants wished to not disclose their age). Their level of experience differed, with some participants having just started doing this type of job and other participants working in Finance for 17 years. All but one worked full-

University	P#	Gender	Age	Occupation	Years of experience
A	P1	M	49	Research Services Administration	17
	P2	F	39	Administrator	2
	P3	F	20	Credit Controller	1
	P4	M	-	Assistant Accountant	15
	P5	F	33	Accounts Assistant Expenses	4
B	P6	M	18	Payroll and Pensions Apprentice	1
	P7	F	40	Payroll and Pensions Assistant	10
	P8	M	52	Payroll Supervisor	12
	P9	F	-	Payroll Officer	13

Table 3.1: Participant information.

time. Table 3.1 shows further demographic details of the participants. Typical data entry tasks participants dealt with were checking and entering expense forms sent by staff and students, paying salaries and pensions, controlling research budgets, monitoring university income and expenses and entering employee information. Participants were recruited by sending invitations to opt-in mailing lists of Finance departments, and were reimbursed with a £10 Amazon voucher.

Materials

Materials that were used during the interview were a voice recorder, a paper copy of an interview script with the interview topics and guiding questions, a consent form, an information sheet for the participant and a notebook and pen to make notes. The interview script, information sheet and consent form are included in Appendix B. Each interview covered four guiding topics, which are briefly described in Table 3.2. For each topic, a number of questions were written out beforehand. These questions were used as a starting point to get the participant talking and guide the interview. Based on what the participant was saying follow-up questions were asked. The audio transcription program ExpressScribe was used to transcribe the interviews. The data analysis programs Nvivo and Atlas.ti were used to analyse the data. Nvivo was used to code the interview transcripts and notes. Atlas.ti was used to complement the analysis in Nvivo and allowed to identify relations between codes.

Topic	Description
Job description	A description of the tasks that the interviewee deals with. The purpose of this topic was to start the interview easy and give the interviewee the opportunity to explain what their job entails.
Number transcription	This includes questions on when and how people typically enter numbers for work.
Environment	This topic includes people's physical work environment, and the organisation they are a part of.
Demonstration	Interviewees were asked to give a demonstration of entering data into their system. The aim of this part of the interview is to see the type of data entry tasks people have to do, and also gives a chance to see the information sources and systems people currently use.

Table 3.2: Interview topics.

Data recording

A voice recorder was used to audio record the interviews. One participant wished to not be audio recorded and one interview could not be audio recorded due to technical issues, so for these two interviews notes were taken of the answers. For the remaining seven interviews, notes were only made of observations and not the participants' answers. Notes were made with pen and paper. Photographs were made of the work environment and screenshots of the systems that the interviewees used.

Interviewing procedure

The interviews took place at the participants' workplace. For two interviews, the interviewee's office place was not suitable for talking so the interview took place in a common room nearby, and these participants showed their workplace and completed a demonstration of entering data after the interview. Participants were welcomed and informed about the study. They received a paper information sheet with the outline of the study and contact details of the researcher to keep for future reference. They were also asked to read and sign a consent form. They were asked permission for the interview to be audio recorded, and could still participate if they did not wish not to be audio recorded.

Participants were informed that the data would be used for research purposes only and stored in accordance with the Data Protection Act 1998. They were also informed that their data would be anonymised and when used in a report or academic paper, their data would not be directly identifiable. Names of participants or the universities they were working at were not included in the interview notes and transcripts.

The interviews were semi-structured and took between 20 and 55 minutes. Each interview was reviewed afterwards, and findings sometimes fed into new questions being included or some questions being adapted in subsequent interviews.

Pilot interview

A pilot interview was conducted with an acquaintance of the researcher who worked in Finance, to test out the set-up of the study and questions. The interview took place at the participant's home, notes were taken with pen and paper, and the interview was audio-recorded using iMovie on a Macbook Pro.

Taking notes slowed down the flow of the interview: sometimes the interviewee stopped talking to give the interviewer the opportunity to finish taking notes. Furthermore, taking notes took attention away from what the interviewee was explaining: assumptions made during the interview did not seem to be accurate in later analysis. Therefore, it was decided that note taking would be kept to a minimum. Notes would only be made of observations that could not be taken from audio recordings.

3.2.3 Results

Data analysis

After each interview or set of interviews, a first analysis took place. The audio recording was played back, notes were typed into a digital file and reviewed and the interview was transcribed verbatim. Several non-verbal cues were included in the transcription as well, such as when the interviewee laughed or sighed, as well as descriptions of when the interviewee was demonstrating something. The advantage of doing the transcription shortly after the interview was that it was still easy to remember from listening to the audio recording what was being demonstrated. Interesting findings and initial patterns that were apparent across the data were written down.

After all interviews had been transcribed, the transcriptions and notes were printed and the data was analysed using thematic analysis (Braun & Clarke, 2006). An inductive approach of thematic analysis was used: there was no pre-existing coding scheme, and codes were created based on what emerged from reading over the transcripts and notes. Anything in the data that was considered to be interesting was annotated by hand and labelled with an appropriate code. On reviewing the coding, some codes were grouped together in one code, additional codes were named, and similar codes were grouped under themes. For instance, an initial code was Notifications, such as e-mail notifications. During the second coding iteration, it was identified that people always talked about notifications in the context that they were interrupted (by a notification) rather than about notifications on its own. Therefore, notifications and interruptions were grouped into one code.

The themes were then reviewed, to see if they addressed the purpose of the study. The transcripts and notes were then imported into Nvivo and coded digitally. Atlas.ti was used to complement this analysis and allowed to identify relations between codes.

Themes

In total 51 codes were derived, and these were grouped into 8 main themes, which are listed and described in Table 3.3. If codes or separate quotes did not belong in a certain theme but were still considered relevant, they were grouped in the Other category.

Theme were visualised as diagrams, which are included in Appendix A. Each diagram shows a theme's main codes and relationships between codes, the number of quotes that were grouped under this theme, and the number of interviewees who mentioned it. These diagrams helped to gain insight how codes were connected, and also how prevalent topics were across the collected data. In the following Results section, I report the key findings from the analysis. The diagrams are included in the Appendix to make transparent how data was analysed and to clarify what I base the key findings and conclusions on.

I first describe the data entry work participants dealt with, to provide context to the type of work I focus on in this thesis. I then discuss how people scheduled their work, the information sources they dealt with, and current strategies people described to improve data entry performance.

Type of data entry tasks

A common data entry task was processing expenses. Participants received requests from university staff and students, who had spent money for research purposes and wanted to claim

Theme	Description	Quotations	Participants
Task characteristics	People described things that were particular to their task, for instance how they structured their task, whether they switched tasks, and how long they took to complete tasks.	129	9
Checking	People talked about checking data input as part of their job.	103	9
System	People talked about the computer system they were using to input data	91	9
Environment	People described their environment, for instance they talked about their physical work setting, and the work culture of their organisation.	80	9
Data	People described the data they were dealing with, for instance the type and length of data items, and from which source they copied data.	75	9
Errors	People described situations where errors were made: who made them, why were they made, what were the consequences.	75	9
Strategy	People described the strategies they used to carry out their task.	54	9
Importance of accuracy and paper trails	People talked about the sensitivity of financial data which is why not all people are authorised to approve or access financial data, and the importance of a paper trail for data entries.	35	8
Other	People talked about things that did not fit into any other category but were still considered relevant, such as issues they experienced, or queries they often received.	74	9

Table 3.3: The themes, along with a description. The column Quotations indicates how many times this theme was brought up during interviews, and the column Participants indicates how many participants talked about it.

back the expenses. Participants had to check the information they were given was correct. For instance, they had to check that the amount being claimed back matched the value on the original receipts, and that the expense fell within permitted limits. In addition to transcribing and checking individual numbers, participants also mentioned they often have to perform and check calculations, for example to check that all individual expenses were converted correctly from a foreign currency. After checking the data and calculations, they then had to enter the data, along with other information such as budgeting and staff information, into a computer system. The time spent on processing one expense claim differed: checking data on a short expenses form could be completed in two minutes, while a single manual calculation could take between 20 and 30 minutes. Two participants described dealing with a lot of data entry as time-consuming and tiresome.

Participants primarily dealt with numeric data, such as financial data and IDs. The monetary numbers they dealt with ranged from five to millions of pounds. Participants also entered and checked alphanumeric and non-numeric data such as employee names, addresses and bank account details. Numeric data consisted of individual numbers, as well as groups of numbers that together made up a new number, such as the total amount of money spent on a project. Participants had to both check and transcribe each individual number, and check that the calculation was correct.

Scheduling data entry work

The workload varied throughout the year. Four participants mentioned that their work followed a work cycle and they did different things at different times of the month. One week could be reserved for checking all the data they received from another department, while another week could be spent on solely inputting data. On average participants reported that they dealt with between 30 and 80 expense claims per day. The amount of individual data items to enter averaged to 6,000 items a day.

To complete data entry work efficiently, eight participants saved up data entry tasks, to enter all data in one sequence. If they received new data entry tasks, they often did not work on them immediately, but instead waited until they had several saved up and then completed all of them in a single batch. P2 was the only interviewee that processed forms with numbers to enter as they came in. She only dealt with a couple of claims per day, and did admit that if she had more data to fill in, she would probably do it in a more efficient way.

The reason for batching data entry work was that participants found it quicker and easier to do the same type of task in one sequence, before they started another type of task. Participants did not have specific deadlines set by their organisation to finish certain expenses. Some said they saved them up but did try to finish expenses before the next payroll was due, so that claimants could get their expenses reimbursed on time. P6 explained he postponed processing expense forms until the deadline to submit forms for that month has passed, after which he did all forms in one sequence: *"When I'm doing it lots at a time, I think once you get into sort of the hang of it, it gets done a lot quicker."* (P6)

Interruptions

Participants batched data entry work, and tried to concentrate on the task at hand, and avoid self-interruptions to other tasks: *"I try to concentrate on my task...I try to do one task (i.e. doing all expenses), finish one, and then do another."* (P9). All participants did attend to some external interruptions, which they considered part of their job: P8 had to pause his task immediately if a staff member entered his office and needed his help. Other participants mentioned they primarily tried to concentrate on the task at hand, but did briefly glance over e-mail notifications, to see whether something important needed their attention.

Information sources

While participants avoided task-irrelevant interruptions during a data entry task, they did have to interrupt and leave the data entry system to look up information related to the task and then get back to the system window to enter it. The information needed for one task was usually spread over several windows on the computer, so participants had to flick back and forth or memorise certain information from one window to use in another window: *"Depending on the input, it can be quite complicated, and there are quite a lot of different screens to input."* (P7). All participants had their computer windows maximised, so one window covered the whole screen and participants could not look at the data entry window and information source concurrently. Participants said that if they had to get out of the system to look up information digitally and then get back to the system window to enter it, they preferred to memorise the information, rather than flick back and forth and look it up each time they needed it: *"I wouldn't necessarily have to [memorise it], it's more (. . .) if you have to keep flicking back to different things, it's sometimes just easier to try and remember it. But you can obviously take the long version and keep flicking back to* *the correct screen."* (P3). Data items that had to be entered frequently such as project codes, were memorised even if participants did not deliberately choose to do so.

The data was spread across both digital as well as paper sources, and participants had to switch between computer windows and physical locations to retrieve all data required for a single expense claim. Examples of digital sources were spreadsheets, Word documents, departmental databases and e-mails. Examples of paper sources were receipts, claim forms, and print-outs of spreadsheets. Participants worked with paper sources for two main reasons. First, digital information was printed out to keep it nearby. P7 and P8 had printed out information they frequently needed to look up and had placed this nearby on their desk, so they could easily use this to check if the input they had received was correct. Second, some documents had to be in paper form for auditing purposes, because participants were working with sensitive financial data. Hard copies of receipts and signed paper claim forms had to be archived and were checked by external auditors. People also explained they liked to write out and keep a record of manual calculations, in case someone had any questions on how figures was calculated.

Checking for data entry errors

Even though participants discussed they tried to focus on data entry tasks by avoiding task switches, they did discuss that data entry errors still happened frequently. Errors that were made were typos, miscalculations, or the wrong information altogether. The main explanation people gave for errors was that it is human to make mistakes, but it was also mentioned people are under time pressure, and that people rely on the fact it will be checked by another person, which makes them less careful in entering accurate data. P9 attributed it to having to switch between computer windows. P4 attributed errors to the use of both paper and digital files: he felt it was much easier to make an error and omit figures when transcribing from paper sources.

Errors had negative consequences on people's work. If an error was spotted and was sufficiently small, it could be processed or corrected without negotiation, but if it was a large error, it had to be sent back or forwarded to a higher authority for approval. This considerably slowed the process down. Furthermore, P8 warned that even with extra human checks not all errors get caught, which meant errors could be made with paying back claimants.

The main method to prevent errors was to have data input visually checked by multiple people before it was processed for payment. People's experience with this checking system differed:

P3, a credit controller who was one of the first people on her team to enter the data before it went to another colleague, believed it increased the chances of an error being caught because it goes through so many different checks. In contrast, P8 and P9 argued this made people even less careful about making errors. As payroll supervisors, they were the last persons at their office to check data before it was submitted to the system and processed for payment. They commented that even at this last stage it was still quite common to spot data entry errors: *"The departments actually sometimes treat us as a checking system [laughs], but they shouldn't really. (...) even though we are like a second check, we feel sometimes that we are the first checkpoint."* (P9).

3.2.4 Discussion

The purpose of this study was to gain a better understanding of data entry work in financial offices, and the physical environment in which these are conducted. For this purpose nine interviews were conducted with office workers from two public universities who worked with financial data. The main findings of the study are:

- people batch their work to enter a lot of data entry at once, and minimise interruptions to other tasks.
- people have to interrupt to get the data to enter from multiple information sources
- data entry errors are common, and the current solution is to have data entered and checked by multiple people before it gets submitted to the system

I first compare the findings with prior research, before concluding what we learn from this study.

Batching data entry work

Though participants received data entry tasks on an ad-hoc basis, most of them saved them until a specific moment and then processed them all in a bulk. People stated that it made them faster in entering data, and they preferred to focus on one type of task at once. Prior experiments have shown that people do become faster in data entry over time, but also more erroneous (Healy, Kole, Buck-Gengle, & Bourne, 2004). Data entry interfaces that slow people down have been shown to reduce errors (e.g. Gould et al., 2016; Oladimeji et al., 2011), but these studies tested up to 240 number entries, whereas participants in the current study reported they often had to enter around 6,000 numbers a day. Interfaces that slow people down may therefore not be applicable in the setting of the current study. Healy et al. (2004) recommend regular breaks when entering large amounts of data to maintain accuracy. However as participants in this study were free to schedule their data entry work and deliberately chose to schedule all their data entry in one session, it may be challenging to initiate and take work breaks.

Fragmentation of work

Contrary to prior data entry research, information was not located on one source but was scattered across physical and digital information sources. Furthermore, in data entry experiments people are often only presented with the data they have to enter and sometimes are given only one data item at a time. The sources from which people had to enter data in this study usually contained a lot of data, not all of which was relevant to the task: sometimes participants had to go through a large spreadsheet, before they found the number they needed to copy. The amount of irrelevant data on the sources can increase the time people need to look up the information they need.

This fragmentation of work is consistent with prior workplace studies that found office work often involves switching between different sources (Cangiano & Hollan, 2009; Czerwinski et al., 2004; Mark et al., 2005; Sellberg & Susi, 2014). In the current study, the time cost to access these information sources also differed. For instance, some paper sheets were on people's desk, but some paper sheets had to be retrieved elsewhere. Participants also dealt with multiple windows on their computer screen, and sometimes needed to switch between different windows. Instead of flicking back and forth to view information they had to enter in another window, participants said they preferred to memorise it, even though they could also write the information down or in some cases copy and paste it, which would be a more accurate strategy. This behaviour can be explained by the soft constraints hypothesis, which states that people increasingly rely on information in memory, as more effort is involved in cognitive offloading or accessing external resources (Gray et al., 2006). This strategy allows people to be faster, but carries the risk that they misremember it. In previous studies, trying to hold more items in memory during a copying task increased errors (e.g. Borghouts, Soboczenski, Cairns, & Brumby, 2015; Morgan et al., 2009). However, in these studies participants had to copy unfamiliar data. In the current setting, the information had some meaning to users and some items were entered more often than others. These familiar numbers are more strongly represented in memory (Wiseman, n.d.), in which case a memory-based strategy may be less error prone.

The understanding that entering data is only one part of the broader data entry task flow can inform future data entry research and improve the way data entry tasks are modelled in lab-based experiments. Future lab-based studies could require participants to first collect data from multiple sources, in order to see how it affects data entry performance. Having an experimental task that is more closely modelled to a situated task will give a better understanding to what extent different interventions are applicable. For example, slowing people down in data entry has shown to reduce errors in the lab (Gould, Cox, Brumby, & Wiseman, 2016; Wiseman et al., 2013), but this intervention may not desirable if people are holding items in memory, or entering large volumes of data.

Error checking

Fragmented attention has been shown to have a negative effect on work performance (Bailey et al., 2001; Carrier et al., 2015), and participants in the current study reported errors happened regularly. To detect these errors, data input was visually checked by multiple people. This checking method resembles Reason's (1990) Swiss Cheese model, where multiple checking layers are used to minimise the risk of errors. However, prior research has shown this is a poor checking method to check for errors in data entry (Olsen, 2008; Wiseman et al., 2013). Despite being widely applied in practice, there is no strong support for the effectiveness of double-checking (Li, Cox, Or, & Blandford, 2016). One of the reasons people may not detect errors when checking a colleague's entries is confirmation bias, which occurs when people selectively attend to stimuli that confirm one's belief (Lewis, 1986). People may expect data entries to be correct: participants reported they regularly received erroneous data, which had previously been checked and approved by several people.

Summary

The aim of this study was to get a grounded understanding of people's data entry work in an office setting. Through interviews with data entry workers, I found that a substantial part of this type of work is not just entering the data but also collecting it from multiple sources. This has often been overlooked in previous data entry work, and may have an impact on data entry performance. We also learn that data entry workers treat irrelevant and relevant self-interruptions differently: interview findings suggest that they try to work efficiently and avoid task-irrelevant interruptions during data entry work. People try to batch similar data entry tasks to avoid task switches, but then have to interrupt for the task to collect information sources, which are spread across the task environment with different time costs to access them.

The study relied on people's own explanations of their practices. This gave insight into reasons why people may employ certain strategies, and through this method I was able to discuss critical incidents which would be unlikely to be uncovered through observation alone. A limitation of relying on people's self-reporting however is that they may not do what they say they do (e.g. Randall & Rouncefield, 2014). Though people gave short demonstrations to support their explanations, they were not shadowed doing their work for longer periods of time. Importantly, while we know that information is scattered across the environment, it remains unclear from these interviews alone how people go about accessing these - in other words, how do they manage inquiries with different time costs? The aim of Study 2 was therefore to observe and understand how participants currently address inquiries for data entry work.

3.3 Study 2: Managing inquiries for data entry

This study and a subset of its results are included in Borghouts, Brumby, & Cox (n.d.) which is currently under review as a journal paper.

3.3.1 Introduction

Study 3.2 revealed that a substantial part of data entry work involved collecting required information from paper and digital sources before it could be entered into a computing system. The time and effort it took to locate different kinds of information sources varied: paper sources could be nearby on people's desks or situated further in another physical location, and digital information could be one click away, or people had to switch between multiple documents before they found what they were looking for. What was difficult to gather from interviews was how variability in the sources of information involved, and the time cost of locating them, affected how people decide whether to attend to, or defer, self-interruptions. This is important to understand to inform the design of any interruption management tools aimed to reduce the disruptiveness of inquiries.

The aim of Study 2 was therefore to observe and understand how participants self-interrupt to access information sources for data entry work. A contextual inquiry was conducted with nine office workers from a similar population as in Study 1. In Study 2, the specific focus was on people's inquiry strategies. Participants were observed at their workplace, and asked to first carry out a data entry task while thinking out loud. Next, they were observed while they continued working as they would normally. As a prevalent data entry task that all participants were involved in was processing expense claims, I focused on this task.

Participants in the current study were video recorded while doing their expenses work. The video recordings captured the participants' interactions with the artefacts involved in the task, but the financial data on the information sources could not be identified from these recordings. The video recordings were used to supplement written observation notes, and after the observation part of the study, some video segments were played back to participants to explain what they were doing at certain moments.

3.3.2 Method

My data collection was informed by using the methodological approach of contextual inquiry. Contextual inquiry is a combination of observation and interviewing users and their everyday work, with the aim to use the findings to inform design of the systems they use (Holtzblatt & Beyer, 2014). Holtzblatt & Beyer (2014) argue that observing participants carrying out their work can reveal concrete details, and it can help participants to recall past situations of carrying out work. It is therefore considered to be an appropriate method for the aim of this thesis, to understand how people can be better supported in managing inquiries for a routine data entry task.

Participants

Nine participants (five male) took part in the study. Ages ranged from 27 to 52 (M = 36, SD = 9); three participants wished not to disclose their age. Participants from Study 1 were invited to participate again, but only one participant was able to participate again in the current study (P8 in Study 1, P1 in Study 2). For the remainder of recruitment, a parallel sampling approach was taken (Onwuegbuzie & Dickinson, 2008). This means participants were drawn from the same population using the same recruitment techniques, but were not the same individuals.

University	P#	Gender	Age	Occupation
B	P1	M	52	Payroll Supervisor
	P2	M	31	Payroll and Pensions Apprentice
A	P3	M	-	Administrative Assistant
	P4	F	27	Postgraduate Research Manager
	P5	F	36	Personal Assistant/Research Administrator
	P6	F	28	Personal Assistant
	P7	M	-	Departmental Manager
	P8	M	40	Payroll and Pensions Assistant
	P9	F	-	Payroll Officer

Table 3.4: Participant information.

As in Study 1, they were employees from financial offices at public universities dealing with processing expense claims. Table 3.4 shows further demographic details of participants.

Participants had roles such as payroll officer, personal assistant, research manager, and administrator. Their experience in their current role ranged from one to 20 years. Participants were recruited through a combination of convenience and snowball sampling. They were invited to participate via emails sent to opt-in mailing lists of Finance departments, and emails forwarded by contact persons and people who had already participated.

Study setting

The study took place in the same type of office setting as in Study 1. Participants were recruited from the same two public universities, and one office from Study 1 was visited again in this study. The other participants were from three different offices within these universities. They worked in open plan offices with two or more colleagues working nearby.

Procedure

A single session with a participant lasted approximately 2 to 2.5 hours, and participants were reimbursed £15 for their participation. All sessions were audio and video recorded. A session followed the following four stages:

1. *Interview.* Participants were briefed about the study and asked questions about the type of tasks they are involved in and the type of information sources they used. The aim of this interview was to make the participant feel comfortable and become familiar with the

study and researcher, and for me to get an understanding of the participant's work and job role. The interview was shorter than the interviews from Study 1 and participants were specifically asked about information sources they used.

2. *Think-aloud part.* In this part, the participant demonstrated processing an expense claim while thinking out loud. The participant was asked to elaborate if something interesting or unusual happened, or if the participant fell quiet.
3. *Observation.* After demonstrating the task out loud, the participant continued to process expense claims as he/she would normally without explaining what he/she was doing, while I observed and took notes.
4. *Summary.* The session ended with a short interview and debriefing session. I summarised findings and confirmed with the participant if these assumptions were correct. If some parts of the observation needed clarification, segments of the video recording were played back to the participant, and he/she was asked to explain what was happening during these moments.

Pilot study

In order to test the suitability of the study set-up, a pilot study was conducted with a financial administrator at one of the two universities who dealt with processing expenses. The study took place at her workplace at the university, and notes were taken with pen and paper.

Initially, the intended method of this study was to conduct a contextual inquiry, followed by a week-long diary study where office workers would log diary entries of their expenses tasks. The aim of this diary study would be to get a further insight in additional information sources the workers may sometimes use, that were not covered in the contextual inquiry. Participants would submit a diary entry either by writing down a description or by taking a photograph of the task setting, showing the information sources. At the end of the day, they would have to answer the following questions about their short entries: what information did you need, where did you need to get it from, and when did you look up the information. This method built on a study by Sohn et al. (2008), where a diary study was a useful method to collect information about people's mobile information needs and how they addressed those needs.

The pilot participant explained that expense tasks usually are conducted in the same manner, and stated I was unlikely to find a lot of instances that differ from my observations of having office workers do the task. Furthermore, by observing them I am able to see the access people have to the resources and how much time it takes them to get the data, as well as when in the task they decide to look up information. This information would be more difficult to get insight to through diary entries.

It was therefore decided after this study to not conduct a diary study but instead only observe office workers, and ask them to explain their work.

Data collection

All sessions were audio recorded, and the think-aloud and observation stages of all sessions were video recorded. Every time the participant used an artefact, I asked them to show it to me. Examples of artefacts included paper documents, digital spreadsheets, computer programs, or calculators. I wrote down a brief description of the artefact, or if it was difficult to write a suitable description, a photo was taken of the artefact. Screenshots were made of the data entry system participants used to enter the expenses data. These screenshots did not include any data entries. In addition to video recordings, notes were made by hand whenever something interesting was observed. During the final stage of the session, participants were asked further questions regarding these observations.

Data analysis

To understand people's self-interruption strategies to get information, it was important to first get a thorough understanding of how this information was distributed in the environment. For this purpose, the data was first analysed using a Distributed Cognition (DC) perspective (Hutchins, 1995). Distributed Cognition is a theoretical framework that views cognition as distributed between people, internal and external sources and over time. As it takes the distributed nature of cognition as focus of analysis, it was considered to be a useful framework for this study to help make sense of the fragmentation of, and access to, information for data entry work.

For this first step of data analysis, I followed the guidelines of Furniss & Blandford (2006) on constructing the following descriptive DC models of the task environment:

- The physical model: this model describes the physical layout of the task environment
- The information flow model: this model describes how information flows through all

users involved in the task

- The artefact model: this model describes all artefacts involved in the task

The models are included in Appendix E. These models are based on the working models of contextual design to identify work activities (Holtzblatt & Beyer, 2014), but are more focused on how information is distributed in the environment. Though the models were originally developed to apply Distributed Cognition for teamwork, the models can also be useful with an individual as the focus of analysis. The methodology facilitated to better understand people's current strategies and workarounds.

Each model consists of a diagrammatic representation that visualises the data, and a narrative representation which verbally describes the data. Creating these models helped gain insight in the type of information sources and how information is distributed across people, the physical and the digital task environment. However, the models themselves do not directly answer the research question of this study, which was not how information sources are distributed, but rather how people self-interrupt to access these sources. The models are therefore not included here in the main text, but can be viewed in Appendix E to clarify how the access to information sources was observed and analysed.

After each study session, written notes taken during the think-aloud and observation stages were typed out and any initial thoughts or findings were added. After data collection from the first four participants, initial versions of the DC models were made. Making these initial versions helped identify any gaps in the data collected so far, and helped guide further data collection sessions.

The audio recordings of the interviews and think-aloud verbal protocols were transcribed verbatim. Video recordings were played back, and additional notes were made if anything new was observed by watching these video recordings. The written transcripts and notes were reviewed and categorised based on the model it related to. The groupings were reviewed and used to refine and expand the models. Video recordings were consulted to fill any gaps from the written data.

After an understanding of the distribution of information in the environment, the transcripts and notes were reviewed again to identify self-interruption strategies. Common types of task strategies and self-interruption strategies were grouped together and coded. There was no pre-existing coding scheme, and codes were created based on what emerged from reading over the

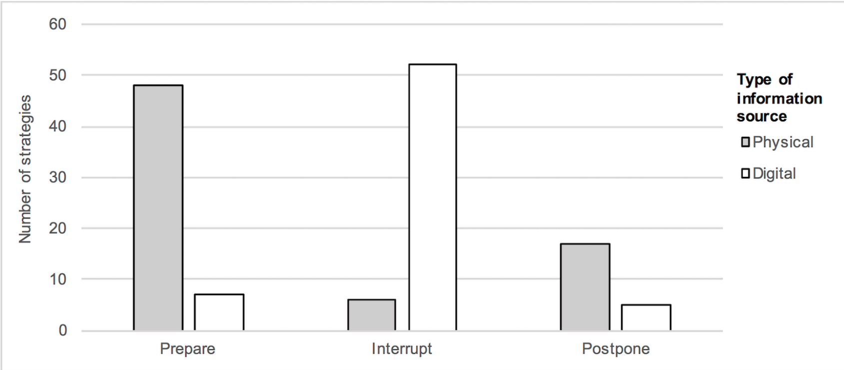


Figure 3.1: Bar chart showing the number of strategies grouped in each high-level category for physical and digital information sources. The most common strategy to collect information from physical sources was to prepare information before starting a data entry task. The most common strategy to collect information from digital sources was to interrupt and switch to the source during a data entry task.

transcripts and notes, but the analysis was focused on revealing strategies to address inquiries. Video recordings were played back and used to iterate and refine the codes. The identified self-interruption strategies are discussed in more detail below.

3.3.3 Findings

The analysis of the data revealed three high-level categories of self-interruption strategies to collect information: Prepare strategies involved collecting information before starting a data entry task, Interrupt strategies happened when people interrupted a data entry task to collect information, and Postpone strategies occurred when participants were aware they needed information, but deferred collecting it. The number of strategies grouped in each category for physical and digital sources is illustrated in Figure 3.1. The figure shows that strategies to collect information from physical sources were primarily grouped in either the Prepare or Postpone category, but barely in the Interrupt category. This means that most physical sources were prepared beforehand, or participants postponed collecting them. On the other hand, strategies to collect information from digital sources were predominantly grouped in the Interrupt category. This indicates that participants most often interrupted a data entry task when collecting information from digital sources.

Table 3.2 provides an overview of all strategies identified in the study. The three columns of the table indicate the high-level categories. Each column is filled with examples of observed behaviour that fall under this high-level category, and numbers in parentheses indicate for which participants this behaviour was observed. These examples are further split into rows, to indicate for which particular information source this behaviour was observed. For example, in the top row it can be seen that participants P1-P9 Prepared (column) collecting a Paper claim form (row) by Placing it on their desk (top-left cell). Each row indicates a different information source: the top seven rows are Physical sources, and the bottom eight rows are Digital sources. I next provide more detailed examples of some of the strategies, first for physical sources and then for digital sources.

Paper information sources

As in Study 1, all participants were aware of the disruptiveness of interruptions, and the importance to focus on their data entry work: *'Expenses claims, (...) they do require high detail to attention. So I like to make sure that's done before I do anything else.'* (P3 - interview). As a result, they avoided switching to unrelated tasks, and participants prepared most paper information sources before starting data entry work. As can be seen in the Prepare column of Table 3.2, several participants prepared sources such as claim forms, receipts, calculators and written instructions by placing them on their desk (P1-P9), personal files were retrieved from cabinets and drawers (P1, P2), or paper sheets were already taped on walls (P7). Participants inspected these sources, and sometimes retrieved additional information sources to check the reliability, *'especially with foreign receipts, you don't really know (...) what they are.'* (P6).

A common observation was that people often discovered that they needed additional information partway through working on a data entry task. If information was nearby, for example, if it was placed in a drawer (P6) or if it could be easily handed over by a colleague (P1, P4, P9), participants interrupted their data entry task and retrieved it straight away.

If colleagues were not available and the information was situated further away, participants postponed looking for the information, and tried to complete other parts of the main task first. In some cases, it was not possible to progress with the task until the required information had been found. This often stopped the task altogether, and people switched to working on a different task instead. As shown in the Postpone column of Table 3.2, strategies to postpone collecting information included sending the claim request back to the claimant (P4, P9), sending an email to the claimant (P2), and writing a note to a colleague who could provide the information (P4, P5). For example, *'I'm going to put this to one side. And come back to it. (...) What I do is just make a post-it note [writes post-it note], and just put it here [places it on a pile in left-hand corner of desk, and*

Type	Information source	Prepare	Interrupt	Postpone
Physical	**Paper claim form (9 participants)**	Place on desk (P1, P2, P3, P4, P5, P6, P7, P8, P9); correct (P7); check against other sources for reliability (P1, P2, P6, P7); interpret (P3, P4); process acceptable errors (P9)	-	Send request back to claimant (P4, P9); email claimant (P5, P9); delegate to colleague (P1, P9); place note on pile on desk (P2)
	Paper receipt (9 participants)	Place on desk (P1, P2, P3, P4, P5, P6, P7, P8, P9); photocopy (P4, P5, P6); check against other sources for reliability (P1); interpret (P6); annotate (P2)	-	Email claimant (P2); place note on pile on desk (P2); place in drawer (P5)
	Calculator (6 participants)	Place on desk (P2, P3, P4, P5)	Retrieve from drawer (P1, P6)	-
	Colleague (4 participants)	-	Ask colleague (P1, P4, P9)	Email/write note to colleague (P4, P5); delegate to colleague (P1)
	Written instructions (4 participants)	Place on desk (P3, P4, P6); interpret (P3, P4, P6); check against other sources for reliability (P3, P6)	-	Email claimant (P5, P6); place in drawer (P5)
	Paper personal file (2 participants)	Retrieve from shared cabinet (P1, P2)	-	Retrieve from shared cabinet (P2)
	Created paper cognitive aids (physical) (2 participants)	Tape next to desk (P7)	Retrieve from drawer (P6)	-

Type	Information source	Prepare	Interrupt	Postpone
Digital	**Search engine (9 participants)**	-	Look up information when needed (P1, P2, P3, P4, P5, P6, P7, P8, P9)	Stop task after not having found information (P4, P5, P7)
	Spreadsheet (digital) (9 participants)	Print out document (P6, P7); create own document (P5, P6); display on second screen (P4)	Open document when needed (P1, P2, P3, P4, P8); browse (P1, P2, P4); use search option (P3); create own document (P3); memorise information (P3, P4); interleave between expenses (P4, P7, P9)	-
	Currency converter application (6 participants)	-	Convert foreign currency (P2, P3, P4, P5, P6, P7)	-
	Email inbox (6 participants)	Open email on computer (P3, P6); print out email and attachments (P3)	Look up information when needed (P1, P2, P4, P5); use search option (P1, P4); browse (P1); attend to notifications (P3, P4); read non-relevant emails (P1)	Read non-relevant emails (P6)
	Intranet (5 participants)	Create own document (P4)	Look up information when needed (P2, P4, P5, P6, P7)	-
	Other external websites (2 participants)	-	Look up information when needed (P4, P6)	-
	Created digital cognitive aids (2 participants)	-	Look up information when needed (P4, P5)	-

Figure 3.2: Overview of observed strategies to collect information. The columns indicate the three high-level categories Prepare, Interrupt and Postpone. Each column is filled with examples of observed behaviour that was categorised under this high-level category. Numbers in parentheses indicate for which participants this behaviour was observed. The rows indicate for which particular information source this behaviour was observed.

goes to new claim].' (P2, think-aloud).

Digital information sources

Participants tried to prepare some digital information sources beforehand as well, as illustrated at the bottom of the Prepare column of Table 3.2. For example, participants prepared spreadsheets by printing them out (P6, P7), displaying them on a second screen (P4), and opened a relevant email on their computer (P3, P6) in advance of starting a data entry task.

As before, people often discovered that they needed additional information partway through working on a data entry task. However, when additional information was needed from a digital information source, rather than postpone looking it up, participants were far more likely to interrupt the task and retrieve it immediately. This can be seen by looking at the bottom rows of Table 3.2 and comparing the Interrupt and the Postpone column: most of the strategies for digital sources are grouped under the Interrupt column, while the Postpone column is mostly empty.

Participants explained they tried to retrieve it immediately because they assumed that digital sources were easy to access and retrieving these involved little time away from the task. Interruptions to look up digital information could however take far longer than intended, as illustrated by the following quote from P4: *'I go and make sure I've got the codes and stuff, ready to go. (. . .) I get halfway through and it goes, Oh, I don't know what that is. And I have to look it up. Then I'll get logged out, because it will take me longer than 5 minutes to do so.'* (P4, think-aloud). Three main underlying reasons for these unexpected time costs emerged from observations, and were supplemented by participants discussing past incidents. First, participants were observed going in and out of several documents to find what they were looking for (P1-P9), and sometimes could not find what they were looking for at all (P2, P4-P7).

Second, participants had to search through large documents with irrelevant information (e.g. spreadsheet tables with 1,000 rows and 20 columns). For example, for each expense claim, a project code had to be entered to specify for which research project the expense was made. Participants had to find this code from a large spreadsheet that contained all codes used within the organisation. During observations, participants used the search option, but also regularly did not know what specific terms to look for, and ended up scanning through the document (P2, P4, P5, P8).

Third, the irrelevant information provided potential distractions and participants were observed being diverted, for instance when they had to find information in email. Email was used by participants both as a communication tool and information source. In its role as communication tool, participants tried to ignore it during data entry work, as it was considered distracting (P1, P2, P4-P6). However, they often needed to access it to find information relevant to their work. During the think-aloud part, P1 tried to find a relevant email and opened several emails to see if it had the information he was looking for. After opening one email, he quickly knew it was not relevant but continued to read it anyway, as it reminded him of something else he had to do later on the day.

These digital interruptions had negative consequences. First, the data entry system logged out after a period of inactive use, which forced participants to restart the task from the beginning: *'You'd sit down to do something, and someone (. . .) or something distracts you, and by the time you go back, the system's frozen and locked you out.'* (P4 - interview). Five participants reported they had experienced these logouts in the past (P4-P8), and in most cases their information was lost. Participants said the added cost of logouts kept them focused on the data entry task, and they were less likely to attend to external interruptions or switch to other, unrelated tasks. Observations however showed that each participant did interrupt their data entry and switched computer windows to look up digital information, without saving their data. Two participants were observed being logged out during the sessions (P6, P8). It was not clear to participants how long the system would wait before logging them out, or how long it would take to look up information, making it difficult to plan for these logouts: *'It doesn't time out, that's why I call it a crash out. We tend to lose various amounts of information.'* (P8, think-aloud).

A second negative consequence was that participants switched back to the wrong window, or entered the wrong information: *'If you, by mistake, left that menu, and went into another linking menu that comes up with somebody else's payroll number, you would never know that you're inputting somebody else's calculation into another record. You have to be so careful.'* (P9, think-aloud).

All window switches during data entry work happened on the same screen, even though most participants had access to two screens (P3-P9). Digital information was only displayed on a second screen if it was prepared beforehand and needed for a longer period of time: *'If it was a credit card claim, (. . .), I would have the list of credit card expenditure on one screen, and then the claim on the other. But then I'd also have another tab where I can look up codes.'* (P4, think-aloud).

3.3.4 Discussion

The aim of this study was to investigate how people self-interrupt data entry work to access task information. The results show that while task-unrelated interruptions are avoided, as in Study 1, and people try to organise their data entry work so they can complete it uninterrupted, they regularly self-interrupt during the task to switch to digital information. They motivated this behaviour by saying they expected these switches to be short, but observations revealed that interruptions often took far longer than people intended, which suggests there is an unawareness of time spent on interruptions. I first discuss possible reasons for people's behaviour, and then discuss what the implications are for interruption management tools.

Paper versus digital interruptions

The first finding is that participants either carefully prepared paper information sources before starting a task or postponed retrieving it, but did interrupt themselves regularly during the task to switch to other computer windows and find additional digital information. One possible reason for this difference in behaviour is that these switches were not experienced as 'interruptions' from the activity, but rather just another part of the same activity. Participants stayed on the same monitor screen when switching windows, and explained they only used a second screen for different tasks. Though participants were observed switching between computer windows to find task information, interviews revealed that they do deliberately try to minimise interruptions to unrelated tasks as data entry work requires focused attention. While this may at first seem like a contradictory finding, it is important because it provides a nuanced understanding of how people think about this type of digital interruption.

Window switching behaviour is consistent with previous research that has shown people switch between application windows when working on a computer every few minutes (Gonzalez & Mark, 2004). This study extends these findings by making a distinction between the types of windows people switch to. The findings suggest that even when people in this context are fairly good at reducing switches to irrelevant windows, they switch immediately to windows needed to locate information for the current task.

Time spent on an interruption

Another possible reason for the different treatment of paper and digital sources is the time involved in retrieving them. Participants predominantly prepared or postponed physical sources, but some instances were observed where they interrupted their work to locate information necessary to complete the task. In these cases, the information was nearby in the physical environment and retrieved rather quickly. The findings suggest that people's decisions regarding whether or not to self-interrupt a task are influenced by the expected time involved in locating the information.

Distracted by other information

As described above, digital interruptions often took far longer to find than intended, as people had to spend effort finding what they needed and were distracted by other, task-irrelevant, information. A likely reason for this outcome was that people needed to access digital sources which are likely to be distracting, such as email. Arguably, participants were largely unaware of the time spent on these digital interruptions, as they adopted deferral strategies for non-digital interruptions when they perceived that they would take excessive amounts of time. This finding is important as it suggests that, even if an interruption is motivated by the goal to locate specific information and then return to the task, people can still get distracted by surrounding information. These distractions may make it difficult for people to be aware of the time that they actually spend on these interruptions – as a result, it is difficult for people to manage them effectively.

The tendency to attend to irrelevant information is similar to so-called chains of diversion, where the user diverts from the current task and forgets the original objective (Hanrahan & Pérez-Qu, 2015; Iqbal & Horvitz, 2007). Previous work has explored tools that aim to prevent these diversions during a task, for example by enabling users to group windows needed for the same task (Smith et al., 2003) and disable switches to distracting sources (Kim et al., 2017). This study illustrates that these types of interventions may not be appropriate in situations where people do not know they need certain sources until they have started the task, and often need to access the sources they find distracting for work.

The study provides further insight in interruptions at the workplace and how task-related interruptions, presumed to be quick and easy, can end up being time-consuming and disruptive to work. This means we not only need to consider blocking interruptions that may be distracting from work, but also what support people can be given to control interruptions which are needed for, and considered part of, the task they want to focus on.

Implications for interruption management tools

The results provide useful initial insights for interruption management tools in the workplace, as they demonstrate that current interruption management tools can provide insufficient support to manage inquiries. First, whilst there are many tools that aim to block interruptions, (Lyngs (2018) suggests that about 40% of the 112 tools he reviewed in 2018 had this functionality), workers might benefit from adopting tools that do not block interruptions, but instead adopt tools that support deferral of self-interruptions until a more convenient moment in the task. For example, upon switching windows, users may be presented with a message discouraging them to switch immediately and an option to set a reminder to collect the required information later. This approach fits with the model proposed by Lyngs (2018), which uses the underlying cognitive mechanisms of self-regulation to frame self-regulation difficulties in ICT use. From the perspective of the model, difficulties occur because at the time of action, people's usage goals are either not strongly represented in working memory, or the value of meeting these goals is too low to control behaviour. By making people more aware of the value of deferring interruptions, or rather the harm of not deferring them, they may be more able to self-regulate their interruptions.

Second, providing people with information about the length of their interruptions may help them self-regulate their interruptions better. One of the self-regulatory guidelines in Lyngs' model is to inform users about their behaviour. When applying this to the setting studied in this study, time information may in particular be appropriate, as participants in the study did already effectively manage some interruptions, when they presumed them to be time-consuming: they addressed them before starting a task or postponed them until later. People may be made more aware of time if they are given timed reminders to return to their task, or explicit feedback about the length of their self-interruptions. Reducing the length of interruptions can be very beneficial for data entry work, as the longer people interrupt, the more disruptive it is to their main task (Altmann et al., 2017).

Finally, though prior studies have shown that two screens can improve task immersion (Bi & Balakrishnan, 2009), in this study a second screen was largely unused for data entry work. In this case, an extra screen can have the contradicting effect of being distracting from work, if it is filled with task-irrelevant information. This finding highlights that the beneficial effect of multiple screens on productivity depends on the type of work: it loses its benefit if a task requires switching between various different documents, rather than one large document which can be displayed on one screen.

3.4 Summary of Chapter 3

The aim of this chapter was to get a better understanding of how people manage inquiries for data entry work in an office setting. Study 1 showed that people try to avoid task-irrelevant interruptions, but have to make inquiries as part of data entry tasks, and a critical component of data entry work is not just entering the data, but also retrieving data from multiple sources distributed in the environment. Study 2 showed that people try to prepare and collect most physical information before starting a data entry task. However, computer window switches during the task were commonly observed as workers often realised during the task that they needed additional information. Digital interruptions often took far longer than people intended, and this had a negative impact on work: software logged users out because of inactivity and they made errors on resumption, entering information in the wrong fields, or entering information incorrectly. These findings suggest that people do not always know how to effectively manage digital self-interruptions that are seen as part of the same activity.

Based on these findings, I make the hypothesis that presumed time costs involved to collect information play an important part in how people address self-interruptions, and that workers are unaware of the time they actually spend on digital interruptions. However, all findings reported in this chapter are qualitative, and are insufficient to make any concluding claims on the effect of time costs on interruption strategies. The use of interviews and observations were suitable to get a better understanding of how data entry work is situated in an office setting and people's self-interruption strategies during this work, but the extent to which these strategies were influenced by the observed time costs associated with inquiries is unclear. Therefore, the next chapter reports three lab experiments to study the effect of time costs on people's interruption behaviour in a controlled setting, and to measure the effect of these strategies on data entry performance.

CHAPTER 4

THE EFFECT OF TIME COSTS ON INQUIRY STRATEGIES

Chapter outline

This chapter describes three controlled experiments that investigate the extent to which time costs of inquiries influence the number, duration and timing of inquiries for a data entry task.

Together these studies show that if the time cost of inquiries can be learnt in a controlled setting, participants adapt their behaviour to try and minimise time by reducing the number of inquiries with a high time cost, and postponing these to be addressed later, rather than addressing them immediately.

4.1 Introduction

The previous chapter revealed that office workers have different interruption strategies to address inquiries, and that there are different time costs associated with these inquiries. If information had to be retrieved from another physical location, participants postponed to access it later. However, digital interruptions were addressed immediately, as they were presumed to be quick. These interruptions could often take longer than intended, which suggests that people are not aware of time spent on digital interruptions. In the Discussion section of Study 2, I considered whether making people more aware of time costs may help them in better managing their interruptions. However, it is not clear from the qualitative studies alone whether time costs actually influenced people's strategies, and what effect this has on task performance. This is important to understand, in order to know whether time information would be effective in managing inquiries and reducing its disruptiveness.

The aim of the studies reported in this chapter is to understand the effect of time costs on people's inquiry strategies for a data entry task. In particular, the first study of this chapter (Study 3) aimed to investigate the effect of time costs on the *number* and *duration* of inquiries and task performance. The second study (Study 4) aimed to investigate the effect of time costs on the *timing* of inquiries. The aim of the third study (Study 5) was to investigate the effect of time costs on timing of inquiries in a *multi-task setup*.

To address the aims of this chapter, a controlled experimental study design was chosen as the study method, which is a useful method to measure the effect of changes of one variable on another variable (Cairns & Cox, 2008). To be able to study the effect of time costs on inquiry strategies, it was necessary to simplify the complexity of various time costs observed in Study 2 into a single independent variable that could easily be manipulated in a controlled environment. In this chapter, time costs are manipulated as the time effort to access task information: in each experiment, participants were given task information and had to copy this information into a data entry interface. Time costs were manipulated by including a time delay to reveal and perceive the information to-be-copied. This manipulation enabled people to learn the time costs associated with inquiries through interaction with the interface. Furthermore, it is a type of time cost which has been used in prior research (e.g. Gray et al., 2006; Morgan et al., 2009), making it straightforward to compare the study findings with prior studies.

4.2 Study 3: Inquiries to a single source

A subset of the study sessions was run by Katherine Corneilson, an affiliate undergraduate student at UCL, as part of an undergraduate project. I designed the study, ran the majority of the sessions, and conducted the data analysis.

4.2.1 Introduction

Participants in Study 1 and 2 had to retrieve and copy data from multiple sources with varying time costs for their work. Overall, participants were motivated to complete these tasks in an efficient manner: participants in Study 1 said they held information in memory when between different computer windows, rather than writing it down. Furthermore, if participants in Study 2 knew they were going to need certain information, and knew there was a high time cost associated with getting this information, they prepared the information before starting a data entry task.

The finding that participants tried to complete tasks efficiently is in line with the soft constraints hypothesis, a cognitive theory which states that people adapt their cognitive strategies to the constraints of a task environment with the aim to optimise task completion time (Gray et al., 2006). This theory has been tested through a series of lab experiments that manipulated the cost to access task information. These studies have consistently shown that as the cost to access task information increases, people try to minimise interruptions to access task information, and instead rely on information they have memorised. This may be a good strategy: making fewer interruptions is less disruptive, and people are quicker to resume the task after being away, because task information is still in memory (Morgan et al., 2009). However, the effect of this strategy on task performance differed across studies: while in some studies it made people more efficient and accurate, in others it slowed people down and increased errors.

For example, Waldron et al. (2007) used a flight simulation task in their study, in which participants had to use flight information to navigate aircrafts to sites of interest. If there was a high time cost associated with accessing task information, participants increasingly relied on information in memory, and were more efficient to complete the task, as they did not have to interrupt their task as often to look back at the information.

Gray & Fu (2004) conducted an experiment where people had to copy over VCR programming information. Participants either had permanent access to the information, or they were explicitly instructed and trained before the trial to memorise the information. In the latter condition, the information during the trial was covered by a grey box which could be uncovered by hovering over it with the cursor. The people in the latter condition were more accurate than the other condition in entering the information. These studies showed that a memory-intensive strategy to save time can improve task performance, if the information is well-encoded in memory.

However, other studies instead found a decrease of task performance when the cost to access task information was increased (Gray et al., 2006; Morgan et al., 2009). In these studies, the Blocks World Task (BWT) was used as a task paradigm, which requires people to copy a 3x3 pattern of coloured blocks, by dragging blocks from a resource window to a target window and putting them in the correct order. The cost to access task information was manipulated across three conditions. In the Low Cost condition, the pattern was permanently visible on the screen. In the Medium Cost condition, the pattern was covered by a grey mask and participants had to hover over the mask with their mouse to reveal the pattern. In the High Cost condition, there was a 2.5 second delay before the pattern was revealed. As in prior studies, participants made fewer interruptions to access information if there was a time cost associated, and instead relied on information in memory. Looking at overall task performance however, they made more errors and took considerably longer to complete the task. In a later paper, the researchers reflected that the coloured blocks participants had to copy may have been too demanding to memorise (Waldron, Patrick, & Duggan, 2011). This abstract visuo-spatial information did not bear any meaning to the participant, in contrast with the VCR programming information used in Gray & Fu (2004) and the flight information used in Waldron et al. (2007) which resembles more familiar information used in a real-world task, and is easier to memorise.

The different effect of a memory-intensive strategy on task performance across studies suggests that the type of task information matters: if information is easy to remember, a memory-based strategy makes it better encoded in memory, with improved task performance as a result. However the studies not only differed in type of task information they used, but also in task paradigm, which makes it hard to compare their findings and say for certain the difference in results is due to the information: for instance, in Gray & Fu (2004) people were explicitly instructed to memorise the information and completed a test prior to a trial during which they had to fill in the information, and could not continue until they had stated everything correctly. This ensured that people had the information well-memorised before they started the experimental trial. In the Blocks World Task studies, participants were not trained or instructed to memorise information well.

It is important to not only understand to what extent people avoid time costs during a task, but also what effect this has on their task performance. Therefore Study 3 first replicates the BWT and only changes type of information, with the aim to see whether the effect of time costs on task performance depends on the type of task information. Study 3 is the only study in the thesis that uses this task paradigm, which was necessary to make a comparison with findings from previous work. The remainder of the thesis uses an expenses task which is based on the type of expenses work observed in Study 1 and 2.

4.2.2 Method

Participants

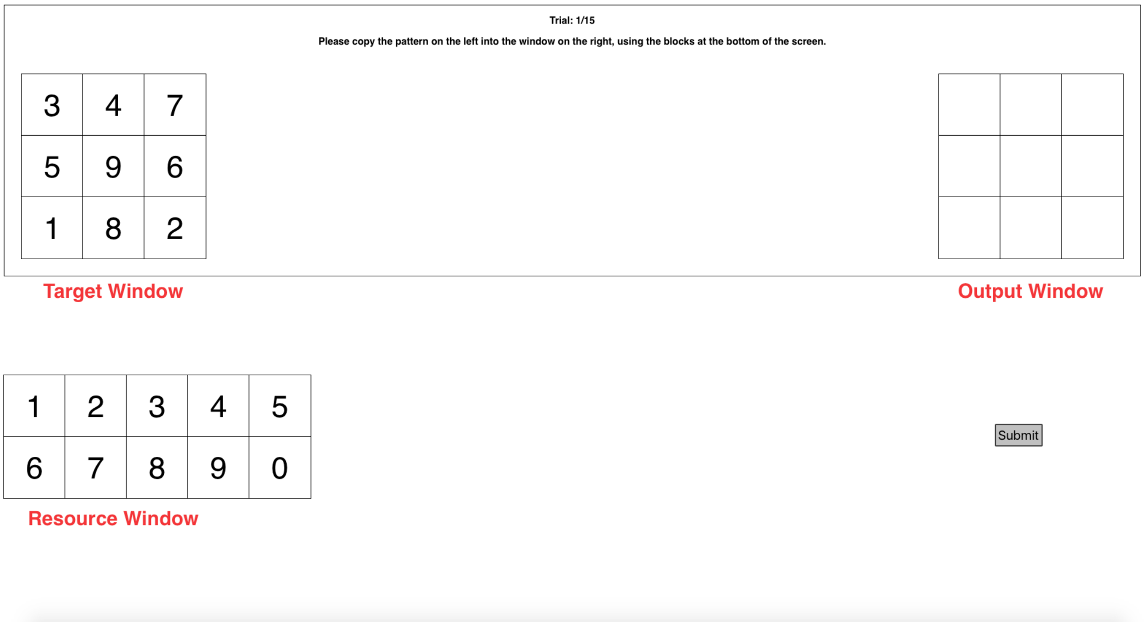
Fourty-two participants (eight male) were recruited from the UCL Psychology Subject Pool. Ages ranged from 18 to 52 with a mean age of 22.38 (SD = 7.45). Participants received course credit or £3.75 as a compensation for taking part in the study.

Materials

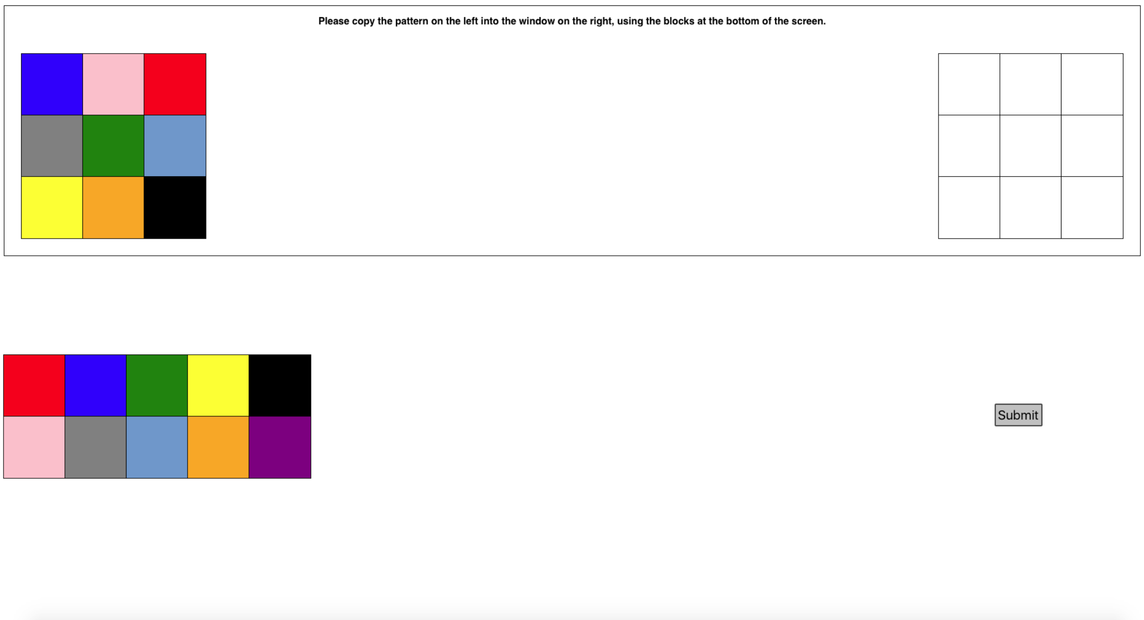
Figure 4.1 shows the task paradigm that was used. Each colour or number was only used once. The colours used were similar to the colours used in previous BWT studies (e.g. Gray et al., 2006; Morgan et al., 2009). Participants had to copy and complete fifteen patterns of each block type, and each participant had to copy over the same patterns. The target window showed a 3x3 grid with either coloured or numbered blocks. The output window showed an empty 3x3 grid, and was the same size as the target window. Participants had to copy the pattern shown in the target window by dragging blocks from the resource window and moving them into the output window.

The study was conducted on a desktop computer, using a 24-inch monitor with a resolution of 2048 x 1152 pixels. Participants used a computer mouse to drag and drop blocks. The experimental task was implemented using HTML, Javascript and PHP and run in a browser. All relevant browser events, such as mouse movements to (un)cover the grey mask, dragging and dropping the blocks and mouse clicks, were recorded and saved in a mySQL database. The browser window covered the whole screen to minimise distractions.

A Tobii T60 eyetracker was used for recording people's eye fixations. Eye movements were recorded at a rate of 60 gaze data points per second for each eye, with an accuracy of 0.5 degrees and timestamp accuracy of 4 ms. For the analysis, all consecutive eye fixations with no drag or drop actions in-between were added together and counted as one fixation.



(a) The number condition.



(b) The colour condition.

Figure 4.1: The task lay-out with the three different components.

Design

A mixed design was used with two independent variables: time cost and block type. The between-participants variable was the level of time cost which had three levels. If the Cost was Low, the target pattern was permanently visible. In the Medium and High Cost conditions, the target pattern was covered with a grey mask, and could only be uncovered by moving the mouse cursor over the window. The mask reappeared as soon as the cursor left the window. In the High Cost condition, there was an additional 1-second delay to uncover the mask. This delay time was used in previous BWT studies where it showed to have a significant effect on task strategies and performance (Gray et al., 2006; Morgan et al., 2009; Waldron et al., 2007). The within-participants variable was the block type to be copied, which was either coloured or numbered blocks. The order was counter-balanced across participants.

The dependent variables are listed in Table 4.1. For the Low Cost condition, the number and duration of visits to the target window were measured using eye fixations; consecutive eye movements within the target area were counted as one visit. Eye-tracking data was also obtained for the Medium and High Cost conditions. However, this eye-tracking data was not used for these conditions, as people were able to look at the target window area without uncovering and perceiving the target pattern. Therefore, in accordance with previous studies (Morgan et al., 2009; Patrick, Morgan, Tiley, Smy, & Seeby, 2014; Waldron et al., 2011, 2007), for the Medium and High Cost conditions the frequency with which the target pattern was uncovered was used as a measurement for visits to the target window. These uncoverings were measured by Javascript. Both the usefulness and limitations of using these measures are discussed in the Discussion.

The number of blocks copied after each visit were measured by Javascript, which recorded all drop events of blocks into the output window. The position where it was placed on the 3x3 grid was used to determine whether the block was placed correctly. The primary focus is on the measures of the first visit, as participants do not have any information yet on the target pattern. On subsequent visits, they may already have partial information in their head from previous visits. Therefore, the items copied after the first visit is believed to be the most 'sensitive measure of performance' (Janssen & Gray, 2012).

Two measures were used to assess accuracy. Incorrectly placed blocks measured instances where a participant initially placed a block in the incorrect place, but then moved this to the correct place prior to submitting the pattern. Incorrectly submitted trials measured instances where the participant had finished copying a pattern and clicked the Submit button, but the pattern was incorrect.

Number of visits
Duration of first visit (s)
Blocks placed correctly after first visit
Number of incorrectly placed blocks (per trial)
Number of incorrectly submitted trials (per experiment)
Trial completion time incl. lockout (s)
Trial completion time excl. lockout (s)

Table 4.1: Dependent variables used in the study.

Procedure

Participants were welcomed and briefed about the experiment. It was explained they would be shown nine blocks which were in a certain order, and had to copy this order by moving blocks around. Participants were instructed to complete the task as fast as possible, but it was explained that they were not able to continue until they had copied a pattern correctly.

After the briefing, participants were asked to read and sign a consent form and were given an information sheet with a summary of the study and the researcher's contact details. In addition to the verbal briefing, the explanation of the study was written out on the computer screen for the participant to read and they were shown an instruction video that showed how the experiment worked. The experiment was broken down in two parts, one where they had to copy colours, and one where they had to copy numbers. For each part, they were given two practice trials first to get familiar with the set-up, and to give them a chance to ask questions if anything was unclear. There was an opportunity for the participant to take a break between the two parts. The study took around 20-30 minutes to complete.

4.2.3 Results

The means and standard deviations of all dependent variables are shown in Table 4.2. Two-way mixed ANOVAs were used to analyse the effect of time cost and block type on the dependent variables. A p-value of 0.05 was used for assessing the significance of all statistical tests.

	Colours			Numbers		
	Low	Medium	High	Low	Medium	High
Number of visits	6.36 (2.28)	4.24 (1.62)	2.98 (0.90)	5.10 (2.48)	2.03 (0.63)	2.05 (0.67)
Duration of first visit (s)	0.39 (0.23)	0.04 (0.02)	2.18 (1.59)	0.51 (0.45)	0.04 (0.05)	1.49 (1.01)
Blocks placed correctly after first visit	1.86 (1.75)	3.22 (1.83)	4.07 (1.20)	2.36 (1.74)	5.96 (1.52)	5.98 (1.55)
Number of incorrectly placed blocks (per trial)	0.15 (0.18)	0.67 (0.40)	0.79 (0.44)	0.17 (0.19)	0.31 (0.18)	0.46 (0.16)
Number of incorrectly submitted trials (per experiment)	0.27 (0.65)	1.9 (2.51)	2 (2.13)	0.36 (1.21)	0.5 (1.08)	0.83 (1.03)
Trial completion time incl. lockout (s)	19.60 (2.98)	25.40 (5.16)	31.80 (6.08)	19.47 (3.03)	20.83 (3.08)	25.95 (4.21)
Trial completion time excl. lockout (s)	19.60 (2.98)	25.40 (5.16)	28.84 (6.34)	19.47 (3.03)	20.83 (3.08)	23.89 (4.06)

Table 4.2: The means (and standard deviations) of dependent measures for the different conditions.

Cleaning the data

Eight participants were removed from the analysis due to weak eye-tracking calibration. Furthermore, one participant misunderstood the experiment and did not know she was allowed to uncover the mask of the target window more than once. This participant had scores that were more than three times the interquartile range from the rest of the participants' scores on six different variables, so this participant was considered an outlier and removed from the analysis.

Number of visits to the target window

The number of visits was measured using eye fixations for the Low Cost condition, and the uncovering of the target pattern for the Medium and High Cost conditions. Participants made fewer visits to the target source when they had to copy numbers (M = 3.06, SD = 2.08) than when they had to copy colours (M = 4.49, SD = 2.18), $F(1,30) = 41.62$, $p<.001$, $\eta^2 = 0.58$. Participants also made fewer visits as IAC increased from Low (M = 5.73, SD = 2.41), to Med (M = 3.13, SD = 1.65), to High (M = 2.51, SD = 0.91), $F(2,30) = 15.16$, $p<0.001$, $\eta^2 = 0.50$. To investigate differences between conditions, post-hoc Tukey comparisons were performed. Results showed that participants made significantly fewer visits in the Medium-IAC condition than in the Low-IAC condition, $p < .01$. However, there was no difference in number of visits between the Medium-IAC and the High-IAC conditions, $p=.59$. Participants looked at the target window for colours more on every level of IAC, and so there was no significant interaction, $F(2,30) = 2.82$, $p=.08$, $\eta^2 = 0.16$.

Duration of first visit to target window

There was no significant main effect of block type on the duration of the first visit, $F(1,30) = 3.05$, $p=.09$, $\eta^2 = 0.09$. There was a significant effect of IAC, $F(2,30) = 16.64$, $p<0.001$, $\eta^2 = 0.53$. Participants looked longer at the target source as IAC increased from Low to High. Post-hoc comparisons showed that participants looked longer in the High-IAC condition ($M=1.84$, $SD = 1.35$) than in the Low/Medium-IAC conditions, ps $<.001$. However, there was no difference in duration between the Low-IAC ($M = 0.45$, $SD = 0.46$) and the Medium-IAC ($M = 0.05$, $SD = 0.04$) conditions, $p=.47$. There was a significant interaction effect between IAC and block type, $F(2,30) = 5.70$, $p<.01$, $\eta^2 = 0.28$. There were no difference between block types in the Low-IAC condition, $t(10) = -1.86$, $p = 0.09$, nor the Medium-IAC condition, $t(9) = -0.29$, $p = 0.7$. However, in the High-IAC condition, participants looked significantly longer for colours ($M = 2.18$, $SD = 1.59$) than numbers ($M = 1.49$, $SD = 1.01$), $t(11) = 2.76$, $p = 0.02$.

Blocks placed after first visit

People placed more blocks correctly after the first visit for numbers ($M = 4.77$, $SD = 2.33$) than colours ($M = 3.08$, $SD = 1.81$), $F(1,30) = 63.86$, $p<.001$, $\eta^2 = 0.68$. They also placed more blocks as IAC increased, $F(2,30) = 12.54$, $p<0.001$, $\eta^2 = 0.46$. Tukey post-hoc comparisons show there was a difference between the Low IAC and Medium/High IAC conditions (ps$<.01$), but not between Medium and High IAC conditions ($p=.77$). There was a significant interaction effect between IAC and block type, $F(2,30) = 8.96$, $p<.01$, $\eta^2 = 0.37$. When IAC was Low, the number of blocks that were copied correctly after the first visit did not differ significantly for colours or numbers.

Global task performance

The interactions between block type and IAC on global task performance measures all had the same trend: people performed the same for colours and numbers when IAC was Low, but differences appeared between the block types as IAC increased. As this trend was the same for each interaction, the statistical results of the interactions are reported but their specific trend will not be repeated.

Number of incorrectly placed blocks

Participants placed more blocks incorrectly for colours (M = 0.54, SD = 0.45) than numbers (M = 0.32, SD = 0.21), $F(1,30) = 10.72$, $p=.003$, $\eta^2 = 0.26$. As IAC increased and participants were keeping more items in memory, they increasingly placed more incorrect blocks, $F(2,30) = 14.71$, $p<.001$, $\eta^2 = 0.50$. Tukey post-hoc comparisons show there was a difference between the Low IAC condition (M = 0.16, SD = 0.18) and Medium/High IAC conditions (ps<.01), but not between the Medium (M = 0.49, SD = 0.35) and High IAC conditions (M = 0.63, SD = 0.36) ($p = .3$). There was a significant interaction effect between IAC and block type, $F(2,30) = 3.36$, $p<.05$, $\eta^2 = 0.18$. When IAC was Low, the number of blocks that were copied incorrectly did not differ significantly for colours or numbers, but as IAC increased, participants placed more blocks incorrectly for colours.

Number of incorrectly submitted trials

The number of trials that were submitted incorrectly was generally low, but participants submitted more incorrect trials for colours (M = 0.1, SD = 0.16) than numbers (M = 0.04, SD = 0.08), $F(1,30) = 5.28$, $p=.029$, $\eta^2 = 0.15$. There was no significant effect of IAC, $F(2,30) = 2.70$, $p=0.08$, $\eta^2 = 0.15$, nor any interaction, $F(2,30) = 1.65$, $p=.2$, $\eta^2 = 0.10$.

Trial time

Two trial completion times are considered here: total time and time excluding lockout. Looking at the actual completion time, participants took longer to complete a trial when they were copying colours (M = 25.80, SD = 7.06) compared to when copying numbers (M = 22.24, SD = 4.47), $F(1,30) = 44.09$, $p<.001$, $\eta^2 = 0.60$. As IAC increased from Low to Medium to High, participants took longer to complete a trial, IAC, $F(2,30) = 15.91$, $p<0.001$, $\eta^2 = 0.52$. Tukey post-hoc comparisons show there was a difference between Low/Medium and High (ps<.01), but not between Low and Medium ($p = .12$). There was a significant interaction effect between IAC and block type, $F(2,30) = 11.05$, $p<.001$, $\eta^2 = 0.42$.

With the lockout time in the High-IAC condition removed, the same effects were found for block type, $F(1,30) = 34.55$, $p<.001$, $\eta^2 = 0.54$, and IAC, $F(2,30) = 8.18$, $p=0.001$, $\eta^2 = 0.35$. Tukey post-hoc comparisons show there was still a difference between Low and High ($p=.001$), but no longer between the Medium IAC and Low IAC or High IAC conditions (ps >.1). There was a significant interaction effect between IAC and block type, $F(2,30) = 8.13$, $p=.002$, $\eta^2 = 0.35$.

Qualitative data

The screen recordings from the eye-tracker were played back to further investigate people's behaviour. Although this helped understand some behaviour which could not be determined from the quantitative data alone, these observations only serve to explain some of the quantitative measures and are not the main focus of the analysis.

The visit durations in the Medium IAC condition were suspiciously short. Upon replaying the screen recordings, it appeared that participants often accidentally moved their cursor over the grey mask of the target source. This was counted as a visit by the program, even though participants may have not intentionally moved their cursor to this part of the screen to look at the target source. They did not spend a long time looking at the target window, but also did not immediately move blocks either, and sometimes waited multiple seconds before they made a move.

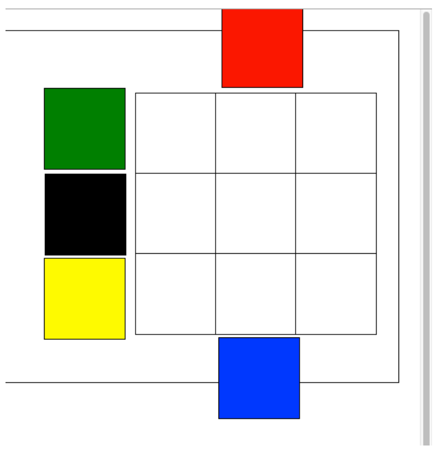


Figure 4.2: Participants placed blocks outside of the output window as 'placeholders'.

During the 1-s lockout in the High IAC condition, participants changed their minds about visiting the target window on numerous occasions. They placed their mouse cursor on the mask, but left this field before it was uncovered to move one or more blocks. It could be this decision also occurred in the Medium IAC condition, but as there was no lockout the mask was already uncovered before people made this decision, and would explain the very short visits.

People sometimes placed the blocks as ’placeholders’ as shown in Figure 4.2: they placed several blocks outside of the output window next to the position they thought it belonged to, but did not place it there yet. Only after viewing the target again, they placed the blocks in the output window. Looking at quantitative data alone, this type of strategy would be depicted as one long view at the target, after which all blocks were placed in one go. This is true to some extent, but as people could already place the blocks and offload their memory without this being recorded by the program, they only had to check if this position was correct on the subsequent visit, and is different from a strategy where people spent a long time trying to memorise the blocks after which all blocks were placed.

4.2.4 Discussion

The aim of this study was to investigate the effect of time costs on the number and duration of inquiries, and the effect on task performance. The main findings are:

- increases in time costs make people adopt memory-intensive strategies
- the effect of a memory-intensive strategy on task performance depends on the type of information
- the effect of a memory-intensive strategy on task performance depends on the type of task

The effect of time costs on people's strategies is consistent with prior work (Gray et al., 2006; Morgan et al., 2009; Waldron et al., 2007). People switched from a perceptual to a memory-based strategy by making fewer but longer visits to the target window and placing more blocks immediately after the first visit. This further confirms that in a controlled setting, people are sensitive to small increases in time costs and try to avoid time costs.

In the Colours condition, a memory-intensive strategy worsened participants' task performance, as they took longer to complete the task and placed more incorrect blocks throughout the trials. In the Numbers condition, this strategy did not increase errors, which shows that the type of information matters when considering whether a memory-intensive strategy is beneficial or not. Numbers were likely easier to memorise, which was demonstrated by the higher number of blocks that were copied after a first visit to the target window: on average, people placed six numbered blocks after a first visit, which is about the number of items people can hold in short-term memory (Miller, 1956). In comparison, people only placed on average three coloured blocks after a first visit. Numbers can be rehearsed, and therefore refreshed in working memory, whereas visuo-spatial information such as coloured blocks is more difficult to memorise (Baddeley & Hitch, 1974).

Contrary to prior work however (Gray & Fu, 2004; Soboczenski et al., 2013), a memory-based strategy when copying numbers did not make people more efficient or accurate, which could have been due to the nature of the task. The error rate was overall low and upon reflection the interaction of moving blocks may have made people sufficiently slow to hardly make any errors. In previous studies, people typed in data using a computer keyboard, in which it is more likely to make data entry errors due to slips (Oladimeji et al., 2011).

Limitations

The study used a similar manipulation of time costs as in previous BWT studies (Morgan et al., 2009; Patrick et al., 2014; Waldron et al., 2011, 2007): in the Low Cost condition, the information was permanently visible, and in the Medium and High Cost conditions, the information was covered by a grey mask. Using this manipulation, it was difficult to measure visits to the target window in the same manner for all conditions. For the Low Cost conditions, eye fixations were used, whereas for the Medium and High Cost conditions, uncoverings of the mask were used.

Measuring visits to the target pattern through eye fixations and mouse movements had a number of limitations. First, while eye-tracking measures show how long and how often people are looking at a particular part of the screen, it can not reveal if people are actually perceiving or processing the data that is displayed (Waldron et al., 2007). Second, the results showed unexpectedly short visits for the Medium Cost condition. Playing back the screen recordings revealed that participants often accidentally uncovered the target window when they were moving their computer mouse, and it is therefore unclear if these uncoverings are a reliable measure of actual visits. Lastly, because visits were measured differently across conditions, the results from the Low-IAC condition may not be directly comparable with the Medium-IAC and High-IAC conditions. Therefore, for the next two experiments in this chapter the experimental setup was adapted so that participants had to make a conscious decision to reveal the target information, and were less likely to accidentally access the source when they did not intend to. Furthermore, the same consistent measure was used across conditions (i.e. a mouse click to make a window switch) to study inquiries.

4.2.5 Conclusion

The aim of this study was to study the effect of time costs on number and duration of inquiries and task performance. The results show that if people retrieve all data from the same source, they will reduce switches between entering and looking up data if the access costs to this source increases. As it took more time to access, offloading behaviour was observed as well, and several participants prepared items they were going to need nearby, but did not use them yet.

The overall aim of this chapter is to investigate the effect of time costs on inquiry strategies. Study 3 showed that time costs reduce number of switches and increases duration of switches. The task however only involved one source, in contrast with the task studied in Study 1 and

2, where people had to deal with various sources, all with different time costs. While we now have a better understanding of the effect of time costs on number and duration of inquiries, we do not know the effect on the timing of inquiries yet. This will be investigated in the next two studies.

4.3 Study 4: Inquiries to multiple sources

4.3.1 Introduction

Study 3 showed people avoid time costs by making fewer inquiries to an information source. Participants tried to group and memorise as much information, in order to minimise the number of revisits to this source. In the experiment, all information was to be found on a single source. As discussed in Chapter 3, data entry in office workplaces is often more complex than switching between a task and a single source: information can be spread over various sources with different time costs associated with them. Information can be one click away, or time has to be spent accessing it. What we do not know from Study 3 is how time costs affect how people schedule inquiries to different sources with different time costs. Observational findings from Study 2 suggest that inquiries with a high time cost are postponed: participants prepared physical information sources either before starting work or postponed it to access later. Digital information sources however were often accessed during the task, as these were presumed to be quick to retrieve. Furthermore, in prior work (Sohn et al., 2008) information access cost was found to be a main factor that determined whether participants looked up information on their mobile phone as soon as they needed it, or whether they postponed to address it later. While these findings demonstrate people take time costs into account when accessing information on physical sources and mobile phones, it is unclear whether participants take time costs of switching windows on a desktop computer into account. Easy switching between windows on conventional desktop computers may give the false impression that information is easy to access (Sellen & Harper, 2003).

The aim of Study 4 is to understand the effect of time costs on the timing of inquiries for a data entry task. An experiment was conducted in which participants had to complete a data entry task, and look up the to-be-entered items by switching to two different computer windows. While prior work has demonstrated that various tasks can involve the use of multiple information sources (Cangiano & Hollan, 2009; Murphy, Chen, & Cossutta, 2016; Su, Brdiczka, & Begole, 2013), it has not been measured how people access these sources, and to what extent the time cost to access a source influences these decisions. Based on the postpone strategies observed in Study 2, the following hypothesis is made:

H1. As the experiment progresses and people become aware how costly it is to access certain sources, they will learn to postpone entering High-Cost items, and choose to enter the Low-Cost items first.

Prior work has shown that increased time costs encourage people to learn more efficient strategies, which they then transfer to use in other situations in which time costs are no longer high (O'Hara & Payne, 1998; Patrick et al., 2014; Waldron et al., 2007). For instance, Patrick et al. (2014) conducted an experiment where participants had to complete a Blocks World Task. Some participants had permanent access to the target pattern, whereas other participants had to complete a number of trials first, in which the target pattern was hard to access. People who were exposed to the interface with an increased access cost first adopted a memory-based strategy and retained this strategy, even when they then interacted with an interface with lower access costs. It is therefore expected that once participants learn it is more efficient to group High-Cost items, they may adopt this strategy for Low-Cost items as well:

H2. As the experiment progresses, participants in the High-Cost conditions will learn and choose to enter all Low-Cost items in a batch, and then the High-Cost items in a batch, rather than looking up each item one by one.

4.3.2 Method

Participants

Thirty-three participants (12 male) ranging from 18-52 years (M = 26, SD= 8) took part in the experiment. They were recruited from a university subject pool and received £4 for their participation.

Task

The aim of the study was to study how people address inquiries from multiple sources for a data entry task. There are currently no existing tasks available that are suitable for this purpose:

in existing task paradigms, all information is usually located on a single source. For the purpose of this experiment, I therefore created an experimental task. The experimental task was based on an expenses task, a routine data entry task observed in the studies in Chapter 3. For this task, the user has to complete a number of data entries regarding incurred expenses in order to get the expenses reimbursed. They enter this into a claim form, which looks similar to a spreadsheet.

For each trial, participants were presented with a data entry sheet consisting of two expense claims (see Figure 4.3). They had to complete each row by entering a financial amount to specify an expense that was made, and an account code to specify which account to use to reimburse the expense. They retrieved these data items by switching to two other windows. One window contained the amounts (Step 1 in Figure 4.3), and another window contained the account codes (Step 2 in Figure 4.3). The participant could go to a window by clicking on the corresponding name in the horizontal menu at the top of the screen. Only one window could be viewed at a time and covered the full screen.

Materials

The numbers to be entered were made to resemble values that are ecologically relevant to an expenses task. The account codes were similar to codes that are currently used by the universities studied in Chapter 3, and have a fixed length of six digits (e.g. 654273). The string of digits was random with no particular pattern. Amounts consisted of two digits on the integer part and two digits on the fraction part (e.g. 11.95).

The experiment was conducted in a maximised web browser on a desktop computer with a 24-inch monitor and a resolution of 2048x1152 pixels. Participants used a computer mouse and number keypad, and it was not possible to copy and paste information. If the participant switched from the data entry form to another window and back, the cursor stayed in the same data entry field. The task interface was developed in HTML, CSS, JavaScript and PHP. All mouse clicks, key presses and timestamps were recorded using JavaScript.

Design

A between-participants design was used with one independent variable, the presence or absence of a time cost when switching to one of the information windows. The time cost manip-

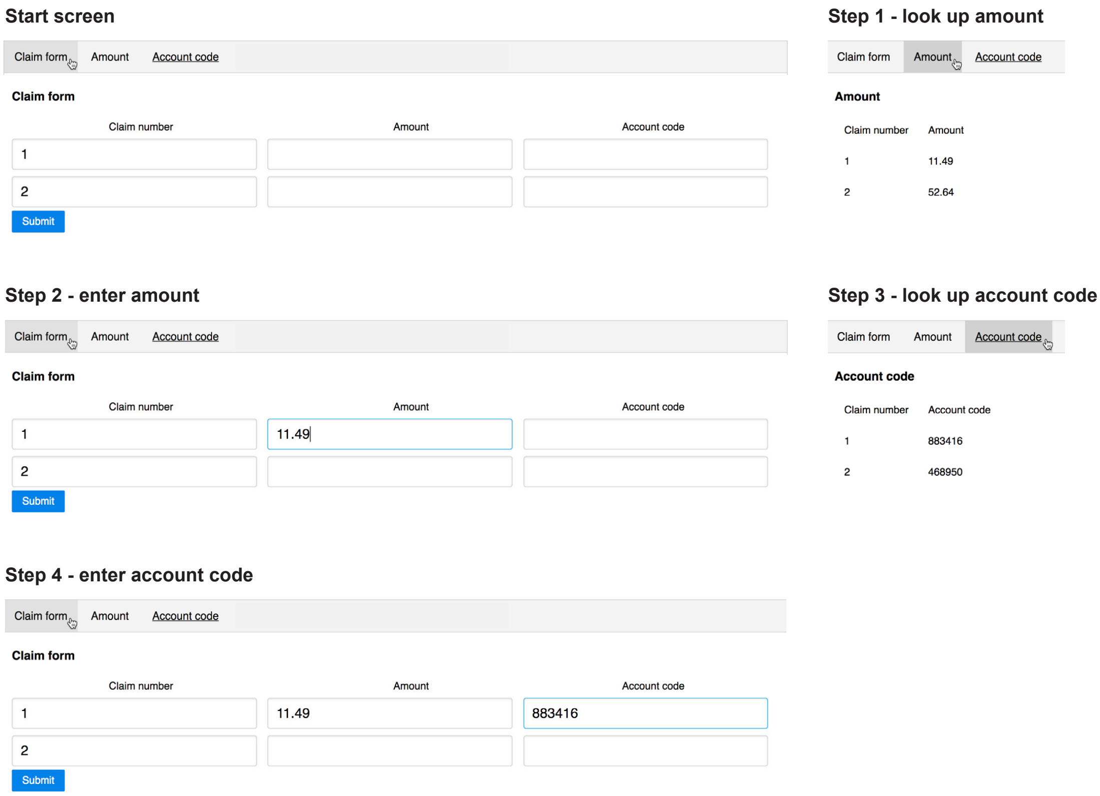


Figure 4.3: The data entry task. At the start of each trial, participants were presented with a data entry form with two expense claims, and had to enter four data items in a data entry form. The participant had to switch to an Amounts window (Step 1), and switch back and enter the correct amount in the correct place on the form (Step 2). The participant then had to switch to the Account code window (Step 3) , and switch back and enter the correct amount in the correct place on the form (Step 4). Step 1-4 were repeated for the other items, until all four items had been looked up and entered. The participant can switch back and forth between windows as often as is needed.

ulation in the current study differs from Study 4.2, as the main focus here is to see how people manage inquiries that each have a *different* time cost, as opposed to inquiries that all have the *same* time cost. The manipulation is explained in more detail below.

In the *Low-Amount, Low-Account* (Low) condition, there were no delays in opening any of the windows. In the *High-Amount, Low-Account* (High-AM) condition, there was a 2-s delay when opening the Amount window, and no delay when opening the Account window. In the *Low-Amount, High-Account* (High-AC) condition, there was no delay when opening the Amount window, and a 2-s delay when opening the Account window. There were no delays in opening the data entry form in any of the conditions.

The Low condition was added as a control condition to understand strategies in a situation where all inquiries had the same time costs, and compare whether these differed from strategies in a High-Cost condition. A High-Cost condition only had a delay on one of the two windows, so that people were presented with both a low and high time cost. This manipulation enabled me to test the hypothesis that people postpone entering High-Cost items in a situation when they can enter Low-Cost items first. Furthermore, there were two high cost conditions, because there were two types of data to be entered (an amount and account code). To know whether any measured differences in inquiry strategies were due to time costs associated with opening a window, rather than the type of data item or the order in which the items were presented, there was one high cost condition where there was a high time cost to access amounts (and a low time cost to access account codes), and another high cost condition where there was a high time cost to access account codes (and a low time cost to access amounts). To simplify notation, from this point onwards these two different High-Cost conditions will be referred to as the High-AM condition and High-AC condition, respectively.

To investigate the timing of inquiries, the order in which participants entered the data items was analysed. On a trial-by-trial basis, the main dependent variable was whether people interleaved between expenses or not: did participants enter the data items in sequential order (i.e. enter one expense first, and then the second expense), or did they interleave between the two expenses to enter items from the same source first (i.e. enter all amounts first, and then all account codes)? Two values had to be entered for each expense: an amount and an account code. If participants entered the amount and account code of one expense before entering the other expense, this was considered a sequential order. If participants entered amounts of each expense first, followed by entering the account codes or vice versa, this was considered interleaving. All window switches and key presses were recorded to determine in which order data was entered. Window switches were recorded to capture the number and duration of switches to information windows. Other dependent variables were trial completion time and data entry error rate. In addition, the type of errors was analysed.

Procedure

The experiment took place in a closed quiet room. It was explained to participants that the task involved entering expenses, and that for each trial they had to enter two expenses. They were not instructed to use a particular strategy, but it was explained it was important to complete all data entry fields before proceeding to the next trial, as they could not return as soon as they had pressed 'Submit'. There were no restrictions in the number or duration of times they could switch between windows, or the order in which they completed the trial. One trial consisted of two expenses, i.e. four data entries. Participants first completed two practice trials to familiarise themselves with the task, and were free to ask any questions; data from these trials were not included in the analysis. After that, the experimental session consisted of 50 trials, divided into 5 blocks of 10 trials. After each block, there was an opportunity for the participant to take a short break. A prompt appeared on the computer screen, and the recording time was paused. Participants could carry on with the experiment by pressing a button on the screen. For each block, a set of 20 different amounts and 20 different account codes were used. These sets were re-used for every block, so in total, each number was presented five times throughout a session. The experiment took approximately 30 minutes.

Pilot study

Because the task was newly created for the study and had not been used in previous studies before, two pilot studies were conducted with colleagues to test the task as well as experimental design. The pilot studies were also intended to see if the length of the experiment was long enough for participants to learn and develop strategies, but not too long and tiring to complete.

During the pilot studies, there was a scheduled break after every 5 trials. Both participants mentioned the break prompts happened too frequently, and experienced them as disruptive. They did not find the experiment too long. One participant could not remember which computer windows had an increased time cost. As a result, he did not adapt his strategies according to anticipated time costs and kept entering the data items row by row. The second participant mentioned that the increased time costs definitely made her more careful in checking the numbers were correct. The participants were aware some of the numbers occurred more than once, but the numbers did not occur often enough to be able to memorise them.

For the main experiments, the breaks were reduced to happen after every 10 trials. In addition, the names of information windows with an increased time cost were underlined in the horizontal menu. This visual feature was added to help users see more easily which windows had a delay.

Data analysis of task strategies

A bottom-up approach was taken to group and analyse people's data entry strategies. For the first iteration of analysis, each trial was grouped into one of two categories: a sequential or interleaving category. If participants first entered the amount and account code of one expense before entering the other expense, this trial was grouped in the sequential category. If participants entered amounts of each expense first, and then entered account codes, or the other way around, this trial was grouped in the interleaving category. On a small subset of trials (<1%) neither of these strategies was chosen: for example, participants first entered the amount of one expense, followed by the account code of the second expense. These trials were also grouped in the interleaving category, as participants switched to entering the second expense before completing the first expense.

Mouse clicks to switch between windows were used to code the order of people's actions, and get insight into the order in which people visited and entered data items. During the second iteration of analysis, for each trial the order of actions was considered and the trial was either grouped under a new strategy group for this order, or the trial was grouped under an existing strategy group. The most common order of actions is shown in Figure 4.6.

4.3.3 Results

Table 4.3 summarises the results of the dependent measures for the three conditions. The distribution of dependent measures were skewed, so non-parametric Kruskal-Wallis tests were used to analyse effects of time costs on the dependent variables. A p-value of 0.05 was used for assessing the significance of all statistical tests.

Condition	Interleaving rate	Number of visits	Duration of visits (s)	Error rate	Trial completion time (s)	Trial completion time (excl lockouts) (s)
Low	31.17% (42.24%)	5.06 (1.03)	1.83 (0.80)	8.68% (10.91%)	29.99 (10.94)	29.99 (10.94)
High-AM	34.18% (41.50%)	5.24 (1.81)	1.60 (0.34)	3.77% (2.80%)	28.88 (6.92)	24.42 (5.91)
High-AC	73.20% (41.10%)	4.29 (0.45)	1.78 (0.60)	5.18% (4.13%)	29.23 (7.44)	25.05 (7.32)

Table 4.3: The means (and standard deviations) of all dependent measures for each condition. The rates are calculated by dividing the number of occurrences to the number of opportunities, e.g. an interleaving rate of 50 percent means that on average, a participant interleaved on 50 percent of the trials.

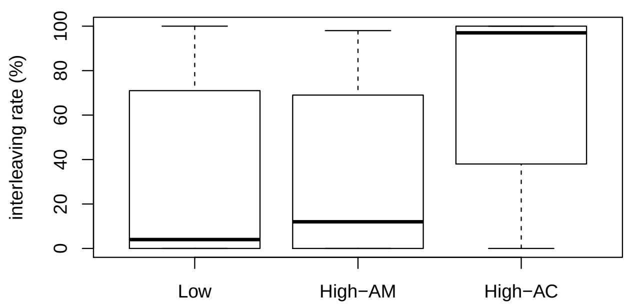


Figure 4.4: Boxplot of interleaving rates in each condition.

Interleaving strategies

A trial was labelled as 'interleaving' if the participant started entering one expense but interleaved to the other expense before completing the first one. The interleaving rate for each condition was calculated by dividing the number of trials where people interleaved by the number of total trials.

The boxplots in Figure 4.4 show the variability of interleaving rates across conditions. The Low condition had a median interleaving rate of 6%, the High-AM condition had a median interleaving rate of 12%, and the High-AC condition had a median interleaving rate of 96%.

Participants interleaved most often between expenses in the High-AC condition (M = 73.2%, SD = 41.1%), compared to the Low (M = 31.17%, SD = 42.24%) and High-AM (M = 34.18%, SD =

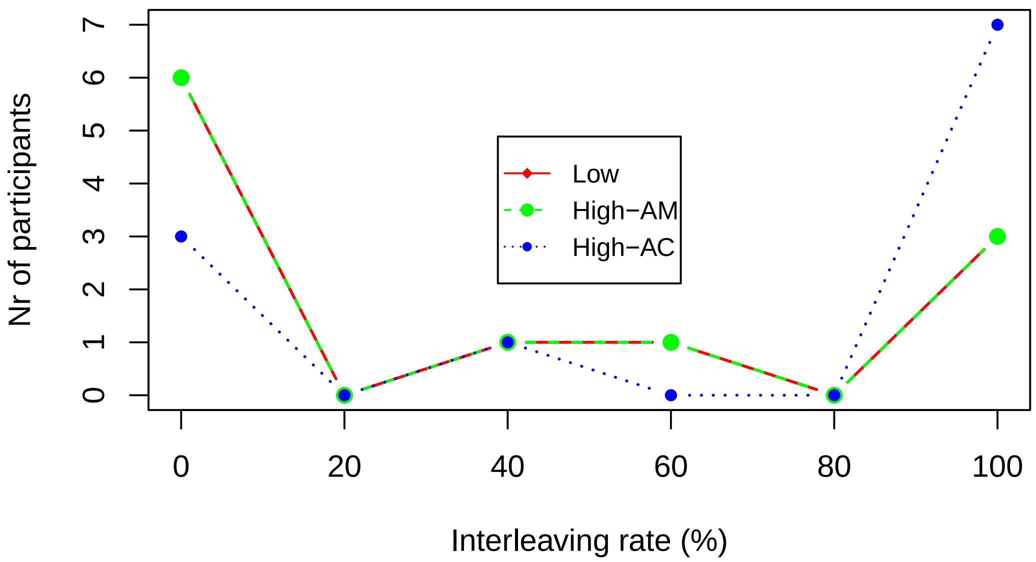


Figure 4.5: Line graph showing the frequency of interleaving rates for each condition; the lines of the Low and High-AM condition overlap and follow the same trend. As can be seen, all three lines have two peaks at 0 and 100, which means that most participants interleaved on 0% or 100% of all trials.

41.5%) conditions, $\chi^2(2) = 6.81$, $p = 0.03$. A post-hoc Dunn's test showed there was a difference between the High-AC condition and the Low (p = 0.02) and the High-AM ($p = 0.03$) conditions, but not between the Low and High-AM conditions (p=0.9).

Across conditions, most participants were consistent in their strategy choice, and either interleaved between expenses on almost no (0%) or all (100%) trials. This is illustrated in Figure 4.5, which shows the distribution of interleaving rates for each condition. The lines all have peaks at the left and right end, indicating the interleaving rate was predominantly 0% or 100% in each condition. Graphs of each individual participant are included in Appendix F, which shows per trial whether a participant interleaved or not. These graphs further illustrate that participants used the same strategy throughout the experiment.

Number and duration of visits

There was no difference in the number of visits, $\chi^2(2) = 2.90$, $p = 0.23$. On average, participants made 4 visits per trial (i.e. one visit per data entry). Participants visited an information page for 1.8 seconds on average, and there was no significant difference in duration of visits between conditions, $\chi^2(2) = 0.30$, p= 0.8.

Most common order of actions

To get a better insight in the specific order in which participants viewed and entered items, the trials were grouped based on the order of actions. There were six different possible actions: viewing the amounts window (V-Am), viewing the account codes window (V-Acc), entering the first amount (E-Am1), entering the second amount (E-Am2), entering the first account code (E-Acc1), and entering the second account code (E-Acc2). This iteration of grouping the trials resulted in 12 different strategy groups in total, with the majority of trials (92%) grouped in the same four groups, which are shown in Figure 4.6. In the Low and High-AM conditions, the most common strategy was strategy (a): participants first viewed the Amounts window and entered the amount of the first expense, and then viewed the Account code window and entered the account code of hte first expense, before they visited the Amounts window again to enter the amount of the second expense, and view the Account code window to enter the account code of the second expense. Strategy a is illustrated in Figure 4.3.

In the High-AC condition, participants predominantly used Strategy (c): they first switched to the Amount window, which had no delay, and entered the amount of the first expense in the data entry form (xb), after which they switched to the Amount window again to view and enter the amount of the second expense. After entering the amounts, they viewed and entered the account codes one-by-one. The strategy is shown step-by-step in Figure x.

Table 4.4 shows the frequency with which these strategies were chosen per condition. The table highlights that even though Strategy (a) and (c) were the most commonly observed strategies, these only accounted for about half of the trials: on the other trials, participants tried out other strategies.

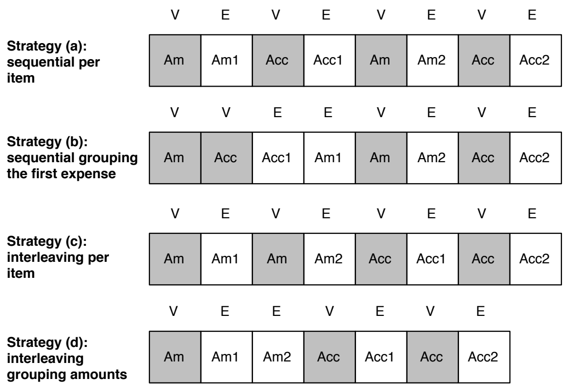


Figure 4.6: The sequence of the most common order of actions. V = visit to an information window, E = entry of a data item. For example, in Strategy (a) a participant first visited the Amounts window, and entered the Amount of the first expense, then visited the Account code window and entered the Account code of the first expense. He/she then viewed the Amounts window again and entered the Amount of the second expense, and then viewed the Accounts window and entered the second expense.

	Sequential			**Interleaving**		
Condition	**Strategy (a): sequential, per item**	**Strategy (b): sequential, grouping the first expense**	**Other**	**Strategy (c): interleaving, per item**	**Strategy (d): interleaving, grouping amounts**	**Other**
Low	**52.12%**	9.13%	5.57%	17.82%	15.14%	0.22%
High-AM	**46.76%**	15.74%	6.48%	9.95%	20.83%	0.23%
High-AC	15.00%	10.00%	2.63%	**48.42%**	15.00%	8.95%

Table 4.4: The occurrence of the most common strategies per condition; the most common strategy per condition is highlighted in bold. The rates are calculated by dividing the number of occurrences to the number of opportunities, e.g. a rate of 50 percent means participants used this strategy on 50 percent of the trials. The strategies are shown graphically in Figure 4.6.

Task performance

There were 200 data entries, so in total there were 200 opportunities for a participant to make a data entry error. The error rates were calculated as the number of errors divided by the number of entries. Though the mean error rate was higher in the Low condition (M=8.68%, SD=10.90%) compared to the High-AM (*M*=3.77%, *SD*=2.79%) and High-AC (*M*=5.18%, *SD*=4.13%) conditions, this difference was not statistically significant, $\chi^2(2) = 0.41$, $p = 0.8$. The High-Cost conditions had an extra time cost to overall completion time, due to the delay when switching one of the windows. Therefore, two completion times were calculated: one measure considered the actual completion time with the delay times included, and another measure considered the completion time with the delay times removed. Considering these two times, there was no difference in the time it took to complete a trial using the actual completion time, $\chi^2(2) = 0.15$, $p = 0.9$, or with the delay times removed, $\chi^2(2) = 2.92$, $p = 0.2$. On average, with the delay times included, participants took about 29 seconds per trial across conditions.

While the above analysis shows no significant difference in the number of data entry errors, it does not give any indication of the type of errors that were made. Having insight into the type of errors can inform how to design better data entry interfaces to prevent these errors(Wiseman, Cairns, & Cox, 2011). For example, it has been shown that interleaving between tasks increases the likelihood of omitting task steps (Back, Cox, & Brumby, 2012). I was therefore interested to see whether interleaving between expenses increased the number of omission errors. To study the type of errors that were made,Wiseman et al.'s (2011) taxonomy of number entry errors was used to categorise data entry errors. This taxonomy was originally created by grouping and coding 350 number entry errors gathered during a number entry experiment. As can be seen in Figure 4.7, the most prominent error types were when participants had a digit(s) wrong (60 times), when a data entry was skipped (75 times) or when they entered a 'wrong' number, which was supposed to be entered in another data entry field (57 times): these types of errors make up for 61% of all errors. The 'digit(s) wrong' and 'skipped' errors happened more frequently in the Low condition, but there was otherwise no remarkable difference in type of errors between conditions.

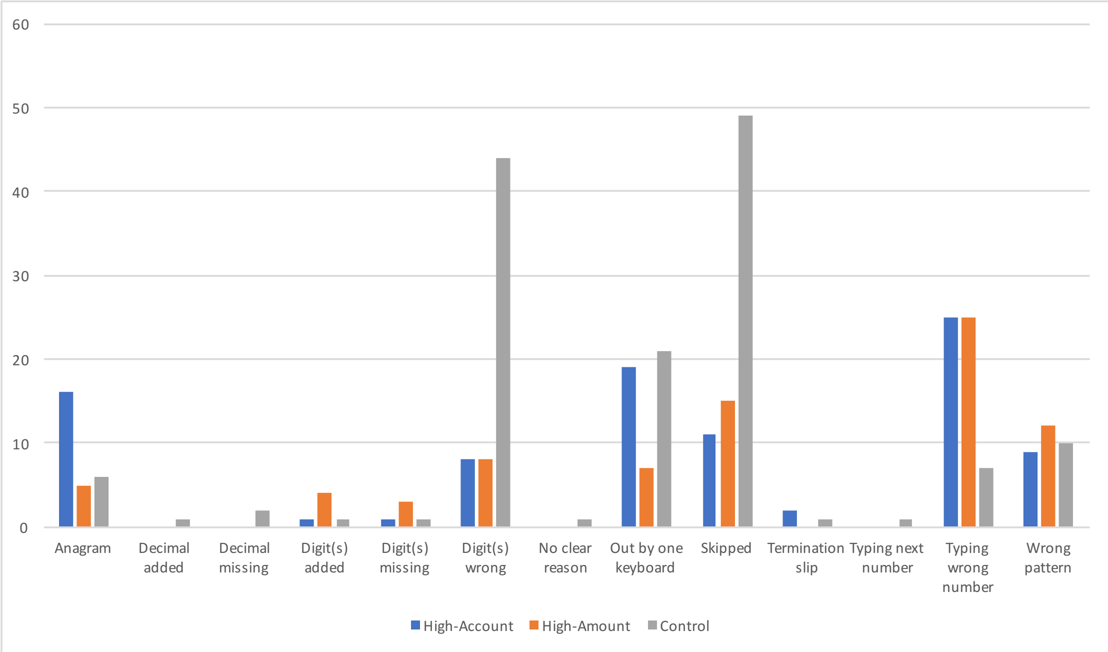


Figure 4.7: The type of data entry errors made in each condition. The most common error types were when participants had a digit wrong, when a data entry was skipped, or when the wrong number was entered in an input field.

Qualitative findings

After the experiment had ended, participants were debriefed and the purpose of the study was explained. Some participants reflected on their strategies and gave additional explanations behind them. While these explanations are not the main focus of analysis and only serve to complement the quantitative measures, it helps understand people's motivation behind some of the measured strategies.

Participants mentioned they adapted their strategy several times throughout the experiment, in order to find the quickest way to complete the task. Because amounts were shorter and easier to remember, five participants mentioned they tried to first view all amounts before entering them. They tried this strategy with account codes as well, but these were longer and therefore it was more difficult to memorise two items at a time. As a result, most participants ended up viewing and entering each account code one by one. This kind of behaviour is illustrated as Strategy (d) in Figure 4.6.

Four participants noticed that numbers re-occurred throughout the experiment. They felt it was easier to memorise a number that had already occurred earlier in the experiment, so when a trial contained a number they recognised, they would memorise this item as well as another item, before returning to the entry form. If they did not recognise the number, they would memorise one item. Furthermore, as data items had a fixed length, some participants started a trial by entering placeholders: in the amount data entry field, they placed a number consisting of four digits and a decimal point, and in the account code data entry field a number of six digits was entered. They would then visit the information windows to check which of the digits of the items they needed to change.

4.3.4 Discussion

The aim of this study was to understand the effect of time costs on the timing of inquiries during a data entry task. To address this aim, I created a new experimental task that involved switching between three different windows to look up and enter data. The main findings of Study 4 are:

- if there were no differences in time costs, participants completed a data entry sheet in sequential order, and completed one expense before moving to the next one.
- in the High-Cost conditions, people interleaved significantly more between expenses in the High-AC but not High-AM condition.
- participants grouped items in all conditions, and there was no difference in number of visits between conditions.

Timing of inquiries

The findings partly support the hypothesis that people postpone inquiries with a high time cost, but it does not explain why participants entered the data entry sheet in sequential order in the High-AM condition. These results can be explained when considering the order in which the data was presented, and the order in which items were entered. Across conditions, participants predominantly started each trial by entering the first cell of the data entry sheet, the amount of the first expense, regardless of whether the Amounts window had a 2-s time cost or not. However, the second item they entered was dependent upon which window had a time cost: if the Amounts window had a time cost, participants would enter an account code next. If there was a time cost when switching to the Account codes window, they would enter the second amount next. This behaviour suggests that time costs do not influence the first visit, but do affect subsequent visits. Even though the time cost was consistent throughout the experiment, potentially the experiment was too short for participants to learn which of the windows had a delay and only adapted their strategy after they had already entered the first item. Furthermore, participants tended to stick to the same strategy they had started with throughout the experiment.

The finding that participants postpone inquiries with a high time cost is consistent with findings from Study 2 and suggests people schedule their inquiries more efficiently and effectively. Though there was no measured difference in task performance in the study, long interruptions have been shown to be disruptive (Altmann et al., 2017; Monk et al., 2008), and leaving these until a natural breakpoint can reduce errors, as it is easier to resume a task (Gould et al., 2013; Iqbal & Bailey, 2005).

Chunking of data items

In contrast with Study 3 and prior work (Gray et al., 2006), an increase in time costs did not reduce the number of visits. However, in prior studies there was no interaction involved to view information in the Low-Cost condition: information was permanently visible in the task interface. In the current study, participants always had to move their mouse and click in order to view the information pages, which may have encouraged them to try and reduce visits and chunk items even in the Low condition. The time cost affected which items participants chunked together, but not whether they chunked items or not.

Transfer of strategies

People adapted their strategies even if only some, but not all, of the information was hard to access. Exposure to time costs may have made people adapt their strategies for all inquiries. This transfer of strategies is consistent with previous research, that has shown a more memory-based strategy can be trained and transferred to other situations where the cost to access information is no longer high (Patrick et al., 2014). This study extends these findings by showing that inquiry strategies can also transfer within a task, when the user has to access multiple information sources with both a low and high time cost.

Limitations

The results suggest that the order in which data was presented may have influenced the order in which people entered data: across all conditions, participants mostly started a task by entering the first data item. An increased time cost affected subsequent items that were entered after the first item. Future studies could be done to investigate whether changing the order has an effect on people's inquiry strategies.

The experimental task was modelled on an expenses task, and the numbers to be entered were similar to data items used for that type of task: financial amounts and account codes. To ensure that any measured differences in inquiry strategies were due to time costs associated with an inquiry, rather than the type of data item that was being accessed, there were two conditions with an increased time cost: in the High-AM condition, there was a time cost to access the Amounts window, and in the High-AC condition, there was a time cost to access the Account window. To simplify the task, the experimental design in Study 6 was adapted so participants only had to enter one type of data item.

4.3.5 Conclusion

Taken together, we learn from this study that people avoid time costs by postponing some inquiries with an increased time cost, and addressing inquiries with a low time cost first. In contrast, if all inquiries have the same time cost, participants predominantly filled in a data entry sheet in sequential order. They completed in one expense first, before moving to the next one.

In the current study, both expenses were shown in the same window. Even though switching between expenses was labelled as an 'interleaving' strategy in the study, the expenses were part of one form, and could be seen as part of the same task. What we do not know from Study 4 is whether participants will also avoid time costs of inquiries by interleaving between two data entry tasks separated over two different windows. Based on the results of the current study, the hypothesis is made that a difference in time costs makes people more likely to interleave between different data entry tasks to enter items with a low time cost first. This hypothesis is tested in Study 5.

4.4 Study 5: Inquiries for multiple tasks

4.4.1 Introduction

So far, Study 3 has shown that people avoid time costs by reducing the number of inquiries, and Study 4 suggested that people avoid time costs by postponing inquiries with a high time cost. In these studies, people were only presented with one task at a time. Workers in Study 2 often dealt with several data entry tasks and windows at a time, and had to be careful not to enter information in the wrong windows. How would people deal with time costs when they have to coordinate multiple tasks? Participants in Study 1 and 2 avoided switches to tasks that were completely unrelated to their data entry work, but people may switch between similar data entry tasks if it makes them faster. For instance, upon opening a spreadsheet that takes time to retrieve, it may be more efficient to enter the account codes from that spreadsheet for multiple tasks. However, multitasking can also be prone to errors (Carrier et al., 2015).

Prior research, studying the effect of time costs on multitasking in a hospital setting, found that increased time costs reduces multitasking. Back et al. (2012) conducted a lab experiment where participants had to enter information from a prescription form into two simulated infusion pumps. For each pump, they had to enter two types of information: the medication dose and the time duration. If the form was physically further away from the pumps, participants more often completed one pump before starting another and as a result made fewer errors in omitting a task step. The higher access cost had the effect that participants memorised and chunked information on the form according to the pump rather than type of information, which reduced multitasking.

However, in Back et al.'s study, all information was located on one information source, and participants incurred a single cost to access it. People therefore chunked information to memorise as much information per visit as possible, so that they did not have to revisit the source too often. It is unclear what the effect of time costs is in scenarios where people do not have to get multiple information items from one source, but rather information from multiple sources. How do people prioritise which information to look up first? Do they still complete looking up information for one task first, before starting another task?

The aim of Study 5 is to test whether the effect of time costs, as found in Study 4, extend to a multi-task setup. Participants were asked to complete an experiment similar to the task in

Study 4, but had to complete two (of the same) data entry tasks per trial. The hypothesis is:

H1. Participants in the Low condition will enter items from one data entry task first, before entering the second task. Participants in the High-Cost conditions will enter the Low-Cost items first, and postpone entering the High-Cost items.

4.4.2 Method

Participants

Thirty-nine participants (32 female), ranging from 18-46 years (M = 25, SD= 8) took part in the experiment. They were recruited from a university subject pool and received £4 for their participation.

Materials

The task was similar to the one used in Study 4 but differed in one aspect. Instead of filling in one data entry form per trial, participants had to complete two sheets per trial, which were shown on two different windows (see Figure 4.8). Each data entry sheet contained one expense, and participants completed the trial by entering the amount and account code for each sheet. The aim of this follow-up study was to investigate if differences in time costs of the two information windows makes people more likely to interleave between two separate data entry tasks.

Design

The experiment was a between-participants design with the presence of a delay as the independent variable. As in Study 4, in the *Low-Amount, Low-Account* (Low) condition, there were no delays in opening any of the windows. In the *High-Amount, Low-Account* (High-AM) condition, there was a 2-s delay when opening the Amount window, and no delay when opening the Account window. In the *Low-Amount, High-Account* (High-AC) condition, there was no delay when opening the Amount window, and a 2-s delay when opening the Account window. There were no delays in opening the data entry form in any of the conditions The main dependent variable was whether participants interleaved between sheets or not: did participants enter the data items in sequential order, or did they interleave between the two sheets? If participants

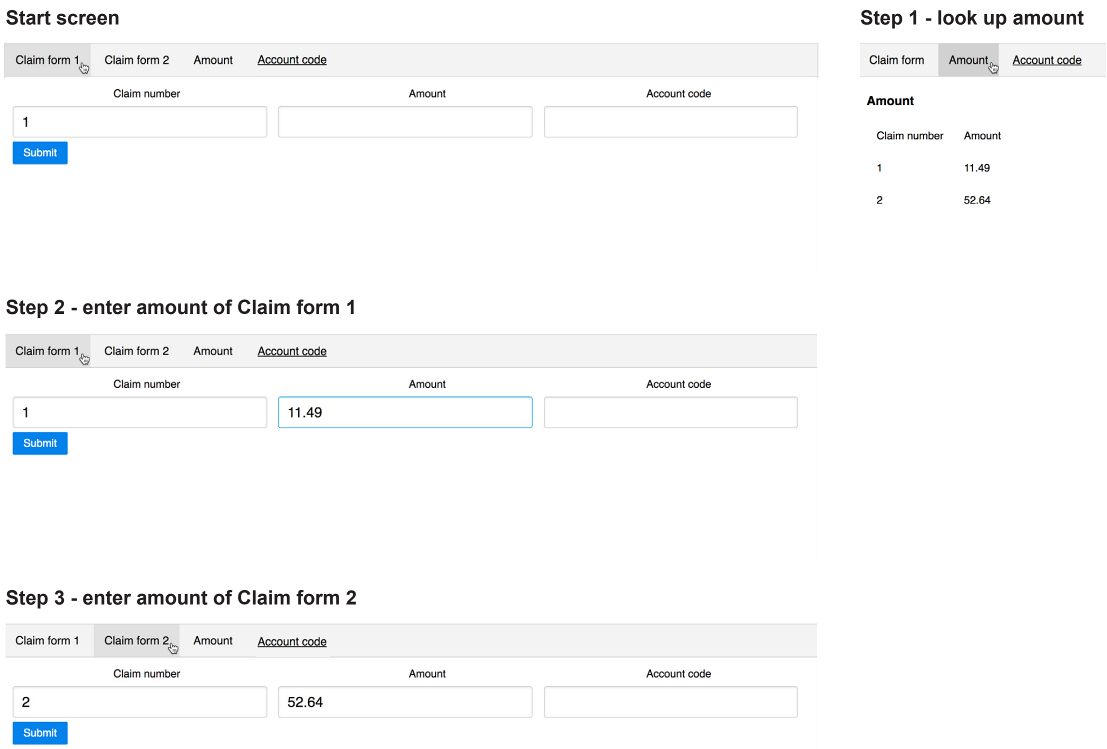


Figure 4.8: Participants had to enter two data entry tasks per trial, each containing two items. Each trial started by showing the first data entry form. As in Study 4, the data items for both tasks were retrieved from a separate Amounts window (Step 1) and entering the items for the first expense (Step 2), and the second expense (Step 3). Participants had to repeat Steps 2 and 3 for the account codes, before submitting the data entries and moving on to the next trial. The participant could switch back and forth between windows as often as is needed

entered the amount and account code of one sheet before entering the other sheet, this was considered a sequential order. If participants entered amounts of each sheet first, followed by entering the account codes or vice versa, this was considered interleaving. All window switches and key presses were recorded to determine in which order data was entered. Window switches were recorded to capture when and how often a participant looked up the data items. Other dependent variables were trial completion time, data entry error rate, and type of errors.

Procedure

The experimental setup was similar to Study 4. For each experimental trial, participants had to enter four data items: they had to complete two forms with two entries each, an account code and an amount. For each experimental trial, participants had to enter four data items, two for each sheet. It was explained that they could use any strategy they wanted, but that it was important to complete both sheets before continuing to the next trial. Participants first completed two practice trials to familiarise themselves with the task, and data from the practice trials were excluded from the analysis. The experiment took approximately 30 minutes.

Data analysis

The main interest of Study 5 was to see whether people interleaved between tasks or not, and not the specific order of individual actions (e.g. did participants enter multiple items after a single visit to a data window). Strategies are therefore not presented here in the same detail as in Study 4 (see section 4.3.3). Trials were categorised into an interleaving or sequential category. On a trial-by-trial basis, it was also considered whether people started the trial by visiting and entering a High-Cost or Low-Cost data item.

4.4.3 Results

Table 4.5 shows a summary of the results of all three conditions for the dependent variables. Kruskal-Wallis tests were carried out to test if there were significant differences between the conditions. A p-value of 0.05 was used for assessing the significance of all statistical tests.

Condition	Interleaving rate	Number of visits	Duration of visits (s)	Error rate	Trial completion time (s)	Trial completion time (excl lockouts) (s)
Low	30.48% (37.71%)	4.5 (0.58)	2.00 (0.68)	3.88% (4.13%)	27.39 (3.49)	27.39 (3.49)
High-AM	83.77% (21.59%)	4.80 (1.85)	2.61 (0.85)	7.54% (4.33%)	33.11 (8.16)	28.71 (6.22)
High-AC	73.41% (32.13%)	4.35 (1.28)	2.25 (0.67)	8.42% (9.07%)	33.83 (6.08)	29.47 (5.94)

Table 4.5: The means (and standard deviations) of all dependent measures for each condition. The rates are calculated by dividing the number of occurrences to the number of opportunities, e.g. an interleaving rate of 50 percent means participants interleaved on 50 percent of trials.

Cleaning up the data

Three participants were removed from the data due to extreme values on performance measures. P28 and P23 made at least one error on every trial. They made 118 and 153 errors out of 200 error opportunities, respectively. P26's session was terminated before the end had been reached, as 45 minutes had passed. This participant spent on average 65 seconds per trial, which is twice as long as the mean trial time of other participants.

These three participants were considered outliers and did not seem to engage with the study, and their data was removed from the dataset. Data of the remaining 39 participants was taken into the data analysis.

Interleaving strategies

A trial was labelled as 'interleaving' if the participant started entering one data entry sheet, but interleaved to entering items on the other sheet before completing the first one. The interleaving rate for each condition was calculated by dividing the number of trials where people interleaved by the number of total trials.

The boxplots in Figure 4.9 show the variability of interleaving rates across conditions. Participants interleaved most often between data entry sheets in the High-AC (M = 73.4%, SD = 32.1%) and High-AM (M = 83.8%, SD = 21.6%) conditions compared to the Low (M = 30.5%, SD = 37.7%) condition, $\chi^2(2) = 11.13$, $p < 0.01$. A post-hoc comparison showed there was a difference between the Low and the High-AM ($p < 0.01$) and High-AC ($p = 0.01$) conditions, and

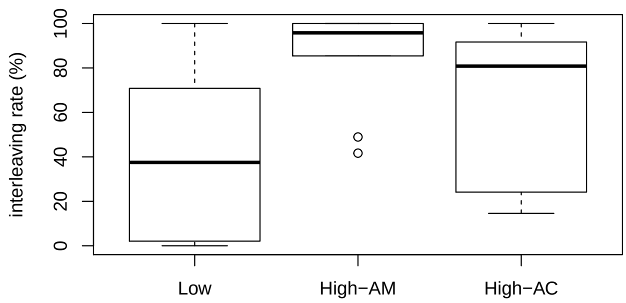


Figure 4.9: Boxplot of interleaving rates in each condition.

no difference between the High-AC condition and the High-AM (p = 0.4) conditions.

On the majority of trials (81.95%) where participants interleaved, they visited and entered Low-Cost items first. On a small subset of trials, participants interleaved by entering High-Cost items first: 2.77% of these trials occurred in the High-AC condition, meaning participants entered account codes first, and 15.28% of these trials happened in the High-AM condition, which means participants entered amounts first.

As can be seen in Figure 4.10, which shows the distribution of interleaving rates, all participants in the High-Cost conditions interleaved on at least a part of the trials: the lines of the High-Cost conditions have a frequency of 0 (participants) at an interleaving rate of 0%. The Low condition has a flat line with no peaks, indicating that interleaving rates in this condition were evenly distributed: participants interleaved on zero, a portion, as well as all of the trials.

As in Study 4, participants made on average four visits per trial, i.e. one visit per data entry. There was no difference in the number of visits, $\chi^2(2)$ = 1.59, p = 0.5. Participants made significantly shorter visits in the Low (M = 2.00s, SD = 0.68s) condition compared to the High-AC condition (M = 2.25s, SD = 0.67s) compared to the High-AM (M = 2.61s, SD = 0.85s) and $\chi^2(2)$ = 6.14, p= 0.04. Post-hoc comparisons found a significant difference between the High-AM and the Low (p=.02) conditions, but not between High-AC and Low conditions (p = 0.2) or the High-AC and the High-AM (p = 0.2).

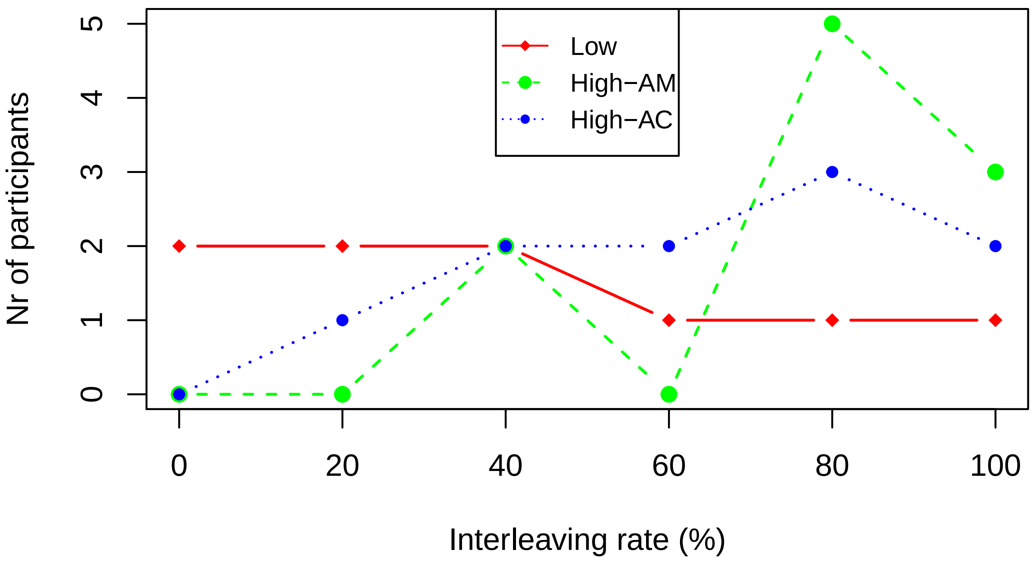


Figure 4.10: Line graph showing the frequency of interleaving rates for each condition. It can be seen that in the Low condition, the line is even, which means that an even distribution of participants interleaved on all, a portion, or all trials. The lines of the High-Cost conditions peak at the right end, which means most participants in these conditions interleaved on at least a portion if not 100% of all trials.

Errors and trial completion time

There were 200 data entries, so in total there were 200 opportunities for a participant to make a data entry error. The error rates were calculated as the number of errors divided by the number of entries. There was a marginal though not significant effect of time cost on error rate, $\chi^2(2) = 5.37$, $p = 0.06$. The mean error rate was marginally higher in the High-AC condition (*M*= 8.42%, *SD* = 9.08%) compared with the High-AM (*M*=7.54%, *SD*=4.33%) and Low (*M* = 3.88%, *SD* = 4.13%) conditions. When comparing the actual completion time including lockouts, participants were significantly faster in the Low condition (*M* = 27.39, *SD* = 3.49s) than the High-AC (*M* = 33.83s, *SD* = 6.08s) or High-AM (*M* = 33.11s, *SD* = 8.16s) conditions, $\chi^2(2) = 8.52$, $p = 0.01$. With the lockout times removed, the difference is no longer significant, $\chi^2(2) = 1.61$, $p = 0.4$.

The type of errors can be seen in Figure 4.11. The most common error type was when a data entry was skipped: this happened 243 times. In the Low condition this type of error occurred 16 times. The error happened more frequently in the High-Cost conditions: in the High-AC condition it happened 114 times, and in the High-AM condition it happened 116 times. Typing the correct number but in the wrong field happened 78 times. This happened 18 times in the Low condition, 14 times in the High-AC and 46 times in the High-AM condition. When comparing across conditions, these types of errors happened on a significantly higher proportion of data entries in the High-AC (*M* = 4.58%, *SD* = 3.6%) and High-AM (*M* = 6.54%, *SD* = 5.01%) compared with the Low condition (*M* = 1.23%, *SD* = 1.82%), $\chi^2(2) = 11.29$, $p < 0.01$. A post-hoc comparison showed there was a difference between the Low and the High-AM ($p < 0.01$) and High-AC ($p = 0.01$) conditions, and no difference between the High-AC condition and the High-AM ($p = 0.4$) conditions.

4.4.4 Discussion

The results show that participants in High-Cost conditions interleaved more between expenses, and made more omission errors, which means a data entry was skipped. The study further supports the notion that people avoid time costs and try to minimise time by postponing some inquiries with an increased time cost, and it demonstrates that this effect of time costs extends beyond a single task setup.

The results provide support for the hypothesis that participants in High-Cost conditions enter Low-Cost items of each expense first, and postpone entering the High-Cost items. On a small

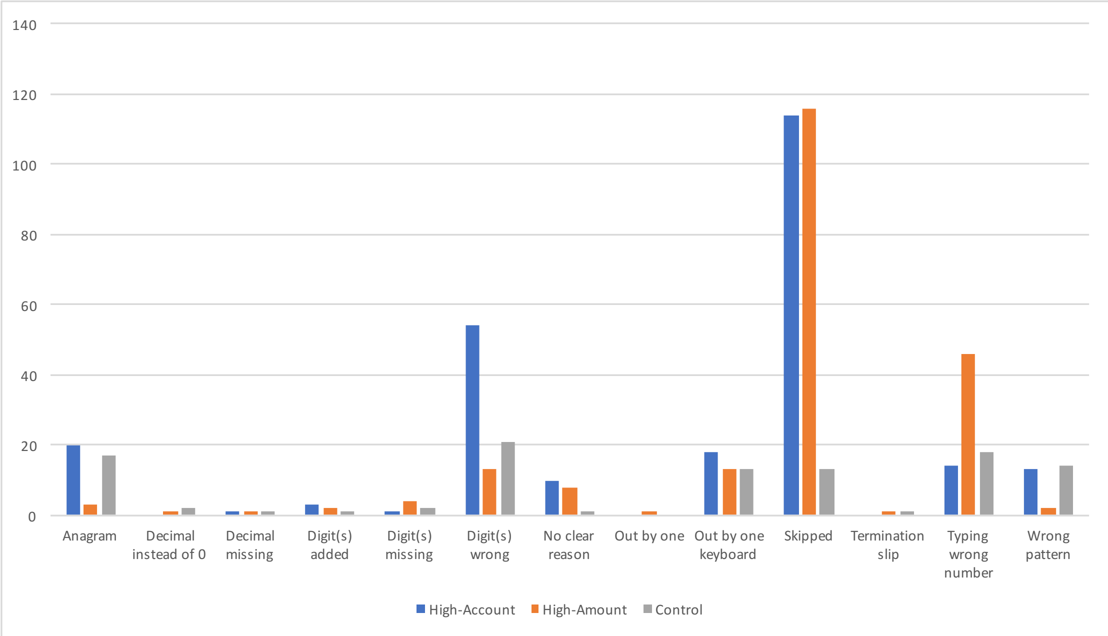


Figure 4.11: The type of data entry errors made in each condition. It can be seen that skipped errors occurred more often in the High-Cost conditions, in which people interleaved more often.

subset of trials, participants interleaved but instead of entering Low-Cost items, chose to enter High-Cost items first. One explanation for this behaviour is the order in which data was presented, which was briefly discussed earlier in Study 4. In the High-AM condition, participants may at times have chosen to start a trial by viewing and entering the first data item, which was the amount of the first expense, even though there was a time cost of accessing this item. Another possible explanation is that participants were trying out different strategies, to learn the most efficient strategy for them.

The finding that people increase their interleaving behaviour to avoid time costs is consistent with the soft constraints hypothesis that people adapt their strategies to millisecond changes in an interface (Charman & Howes, 2003; Gray & Fu, 2004). In previous studies, it was shown how time costs affect the number of steps taken to complete a task (Gray et al., 2006). Study 4 and 5 contribute to this line of work by showing time costs also affect the order of steps in a routine task.

The finding that participants interleaved more as time costs increased, contrasts with Back et al. (2012), who found that an increase in time costs made people less likely to interleave between two data entry tasks. This may be due to the presentation of the information. In Back et al. (2012)'s study, people had to retrieve all information for both data entry tasks from one sheet. If the sheet was nearby, participants read one item at a time, and interleaved between tasks on 59% of the trials. As the cost to access this source increased, they chunked the data items associated with one task, and then after completing this task, returned to the source to chunk data items for the second task.

This study contributes to our understanding of how time costs affect task switching behaviour, and can have implications for tools aimed to minimise task switches. In the current study, there was only a time cost when switching to one of the information sources, but not when switching between tasks. The results showed that people try to avoid switching to something with a high cost. Therefore, adding a cost when switching between tasks may encourage people to complete one task, before switching to another task.

Limitations

Even though the two data entry tasks were separated on different windows, participants may still have felt it was part of the same activity, as the two data entry tasks were shown in the same interface and browser window. Study 2 suggests that even though people tried to avoid task switches, they made frequent switches that were seen as part of the activity. For Study 6, participants will have to switch between different browser windows, to investigate window switching behaviour.

4.5 Summary of Chapter 4

The aim of this chapter was to investigate the effect of time costs on the number, duration and timing of inquiries for a data entry task. Study 3 showed that if people retrieve all data from the same source, they will reduce switches between entering and looking up data if the access costs to this source increases. As it took more time to access, offloading behaviour was observed as well, and several participants prepared items they were going to need nearby, but did not use them yet. Study 4 further demonstrates that when people have to retrieve data from multiple sources, they collect and group items that are quick to access first, and leave items that take longer to access until the end. Study 5 demonstrated the robustness of the effect of time costs in a multi-task setup: when dealing with two data entry tasks, people still entered items with a low time cost first, and interleaved between tasks to enter items with low costs first. As a result, participants made more skipping errors and submitted tasks before they had completed entering all the items. These studies contribute to our understanding of time costs on self-interruption behaviour to collect information: if people know the expected time duration of an interruption, they make fewer interruptions that are long and postpone these switches.

Office workers in Study 2 took time costs into account in a similar manner as shown in this chapter when managing physical interruptions, but did not adopt the same strategy for digital interruptions: these were addressed immediately as participants presumed these to be quick. Comparing those results with the results reported in this chapter, it suggests that people may not be aware of time costs of digital interruptions in a naturalistic setting. In the experiments in this chapter, time costs were manipulated in a specific way: a time delay was added to the interface to reveal information. In practice, the time spent on an inquiry may be because of the time to access it, but also because of time spent searching for information, or people get distracted, and further self-interrupt to other activities. These factors may further make it difficult for participants to learn the time costs associated with interruptions and adapt their self-interruption behaviour.

The studies presented in this chapter only looked at digital inquiries. Future work could be done to compare the use of physical and digital sources in a controlled setting. For example, a future study could investigate how people prioritise digital and physical inquiries with different time costs, and whether people address physical or digital inquiries first.

The studies so far have shown that people try to minimise time to complete their data entry work in an efficient way, but are not aware of the time they spend away from their task looking up information required for the task. The next chapter explores whether a design intervention showing people how long they go away for can make people more aware of these digital interruptions, and whether this has an effect on interruption strategies and task performance.

CHAPTER 5

USING TIME FEEDBACK TO MANAGE INTERRUPTIONS

Chapter outline

This chapter describes two studies that evaluate whether giving people feedback on the duration of inquiries can influence people's switching strategies and data entry performance. A browser notification was developed which showed people how long they switch away for on average. Study 6 evaluated the notification with an experimental task to measure switching behaviour and task performance. Study 7 evaluated the notification in the office setting with workers doing data entry work, to ascertain how appropriate the proposed recommendations are for a naturalistic task.

Together these studies show that time feedback can reduce the duration of inquiries, making participants faster and more accurate in completing data entry tasks.

5.1 Introduction

The studies reported in this thesis so far have shown that people adopt different strategies to manage inquiries, given the time costs associated with these different inquiries. Office workers in Study 2 postponed physical interruptions if they took time, or prepared physical information sources beforehand. Digital interruptions however were addressed immediately as participants presumed them to be quick, which suggests that people are not aware of the time these interruptions actually take: participants were commonly observed being distracted and getting logged out of the data entry system for being away for too long. The experimental studies in Chapter 4 showed that if participants learn the time it takes to make a digital interruption, they postpone addressing inquiries with a high time cost and enter other information first. The problem is that outside of a controlled setting, it is difficult to know how long a digital interruption may actually take. People do not always know where to get information from, may get distracted, or can further self-interrupt to other tasks. How can people be better supported in managing these inquiries with unknown time costs?

Based on the thesis findings so far and prior literature on interruption and information management tools, this chapter presents a design intervention which shows how long people go away from their data entry work. The intervention was evaluated in an online experiment and field study, to investigate whether time feedback can not only help postpone long inquiries, but also reduce the number and duration of inquiries and improve people's focus during data entry work. Before presenting the intervention, prior work on information and interruption management tools are reviewed.

5.1.1 Delaying the intention to interrupt

The timing of an interruption matters, and it is better to defer an interruption at a more convenient moment in the task. For example, it is less disruptive to interrupt a task at a low-workload than high-workload moment (Gould et al., 2013; Iqbal & Bailey, 2005). Prior studies have shown that people choose to defer interruptions until low workload moments if given the option to do so, and if they do not have to hold the intention to interrupt in memory (Gilbert, 2015; Salvucci & Bogunovich, 2010). Gilbert (2015) looked at people's off-loading behaviour of future tasks in both an experimental and naturalistic setting. Participants had to remember to perform an action later, and had the option to offload this intention or to keep it in memory. In both settings, a majority of participants offloaded these intentions when they had the option, and this significantly improved their task performance. Additionally, in Study 3 of this thesis, where participants had to copy a block pattern and remember which blocks to drag to which location, a selection of participants placed blocks nearby what they though the correct location was, to not have to remember its location, and as a reminder to place them there later.

These findings suggest that if people have to memorise which information to retrieve, they may benefit from options to offload these information needs, and are able to effectively defer inquiries until a convenient moment in the main data entry task. However, there is also time effort involved in offloading information, and the time it takes to offload is only worthwhile if it is outweighed by the time it takes to address the interruption. If people are not aware how long an inquiry will take, as was observed in Study 2, there may be little incentive to offload and delay these inquiries: it is presumed to be faster to address the interruption immediately.

5.1.2 Improving information search

The duration of an interruption matters as well: the longer it is, the more disruptive it is (Altmann et al., 2017; Monk et al., 2008). Inquiries may take time if information is scattered across documents and applications, and users have to go in and out of these separately to search and find what they are looking for. To shorten the time it takes to find information across applications, Dumais et al. (2003) developed Stuff I've Seen, a unified search interface which allows users to search through information they had already seen before across applications, such as emails, documents and web pages. A user study found that participants rather used the tool than individual search tools of each application. However, searching for information was not the only type of time cost found in Chapter 3: sometimes participants were quick to find the information source they were looking for, but were distracted by other information. Information search tools are therefore insufficient to reduce the duration of these inquiries.

5.1.3 Preparing task information

Lastly, while some interruptions can be beneficial, all interruptions may become disruptive if they happen too often.The number of inquiries may be decreased if people organise information to have it nearby during the task. People already displayed this behaviour somewhat in Study 2 by collecting physical information they knew they were going to need nearby, to make it easier to access. Some tools have looked at making digital information easier to access during a task as well. For example, GroupBar Smith et al. (2003) allows users to group windows needed for a task in the task bar. This can be particularly useful when resuming an interrupted task: the user can see which documents were used before leaving the task. Similarly, Microsoft Office's new feature TAP [1] allows users to place relevant documents in a task pane next to their working document. The aim of the feature is to keep focus on document creation, rather than looking up information. The feature is presented as a task pane within a document, such as a text document or email, and contains an overview of documents that may be relevant to the current document. These tools are mainly focused on re-using content from archived documents, and assumes the user knows which documents to re-use. The tools provide less support for a new activity in which new sources need to be accessed.

5.1.4 Feedback to improve task performance

An alternative approach is to give people information about their task strategies, as giving people feedback can help them in improving their performance on a task (Farmer, Janssen, Nguyen, & Brumby, 2017; Maior, Wilson, & Sharples, 2018). The findings in this thesis so far suggest that people can be good at managing interruptions, if they are aware of the time costs. Could people therefore improve how they manage digital interruptions if they are shown the time costs associated with digital interruptions?

Giving users feedback on time spent on digital activities has been utilised by a series of time and interruption management tools before (Lyngs, 2018). The primary aim of these tools is to support users in self-regulating their ICT use and making more effective use of their time. Commercial applications such as RescueTime and ManicTime provide users an overview of their computer activities, to reflect how much they spend in total on certain sources. These applications show people's entire computer usage, and interview studies revealed it is often not clear to users what to do with this data (Collins et al., 2014). Furthermore, a problem with retrospective information is that it lacks context, and users have to remind themselves to look at it (Whittaker et al., 2016). On the other hand, feedback which is given during a task allows users to apply the information immediately on the current task they are working on (Gould, Cox, & Brumby, 2016; Maior et al., 2018). Gould, Cox, & Brumby (2016) looked at switching behaviour during online crowdsourcing work, and found that an intervention during work that encouraged people to stay focused after they had self-interrupted reduced the number of switches to unrelated tasks. Whittaker et al. (2016) interviewed office workers and students to establish user requirements for a time awareness application, and found users were primar-

[1] https://support.office.com/en-gb/article/Find-and-use-the-content-you-need-when-you-need-without-leaving-Word-860118fc-1f61-41f6-922f-40084a284658

ily interested in their current activities rather than long-term behaviour. They developed and evaluated an application which presented users with a visualisation of the last 30 minutes of computer activity. The application reduced the time spent on non work-related activities such as social media, but it did not increase time spent on work. While these tools have shown how feedback during a task can reduce task-irrelevant interruptions, the effect of time feedback on managing work-related interruptions has been unexplored.

In summary, prior research has shown that people adapt to feedback given in the moment to improve performance on their task. My studies found that people adapt to time costs of inquiries in a controlled setting (Chapter 4), but are not aware of time costs of inquiries in an office setting, due to distractions, switching documents, and task-irrelevant information (Chapter 3). This chapter explores whether a design intervention which gives feedback on the duration of interruptions can help people manage these interruptions.

5.2 Development of design intervention

The first three ideas revolved around reducing the number or duration of people's interruptions. However, these ideas would not necessarily make people more aware of the time they spend on interruptions, which was found to be a potential issue in Study 2. Upon reflection on these ideas, I was uncertain whether people would be willing to invest time in using a tool, if they would still be under the impression that their interruptions were short and not disruptive.

I therefore focused on developing a design intervention that would also make people more aware of time costs of interruptions, and evaluating if this increased awareness would encourage people to reduce the number and duration of interruptions themselves. To increase awareness, the idea of a notification was developed that would give people feedback on their interruption behaviour.

5.2.1 Task sequence

The feedback was given through a notification. Figure 5.1 shows what the interaction with the design intervention would look like. Step 1 shows the window of an expenses task the user has to complete. To complete this type of task, information needs to be retrieved by switching to another window (Step 2). Upon every switch, a notification appears telling the user how long these switches are on average. If the user switches for the first time, the notification says that no data is available yet to give an average time. Once the right information has been found, the user then has to switch back and enter this information in the correct field (Step 3).

5.2.2 Implementation

The intervention was implemented as a browser notification. As was found in Chapter 3, participants' data entry work was conducted in a web browser and revolved around a main data entry web page. Every switch between the task page and another computer window, such as another browser window or a different application, were recorded to calculate the number and duration of switches. The browser notification used this data to show users the average time of their interruptions.

The notification was evaluated in an online experiment in Study 6, and in an office workplace in Study 7. The presentation of the intervention to participants was the same in both studies, but because of the different study environments, the implementation differed slightly between studies. In Study 6, the browser notification was implemented in the experimental data entry interface. In Study 7, the browser notification was implemented as a browser extension the participants could install and use on any website they wanted. Further implementation details are further discussed in the separate study sections. The notification was evaluated in a pilot study with a number of colleagues before using it in a study.

5.3 Study 6: Looking up information in email during an online experiment

This study and early results have been published in Borghouts, Brumby, & Cox (2018) and were presented at the CHI conference in 2018.

5.3.1 Introduction

This study aimed to investigate whether an intervention showing people how long they switch on average reduces the duration and number of switches during a data entry task. An online experiment was conducted where participants had to complete a data entry task. Participants

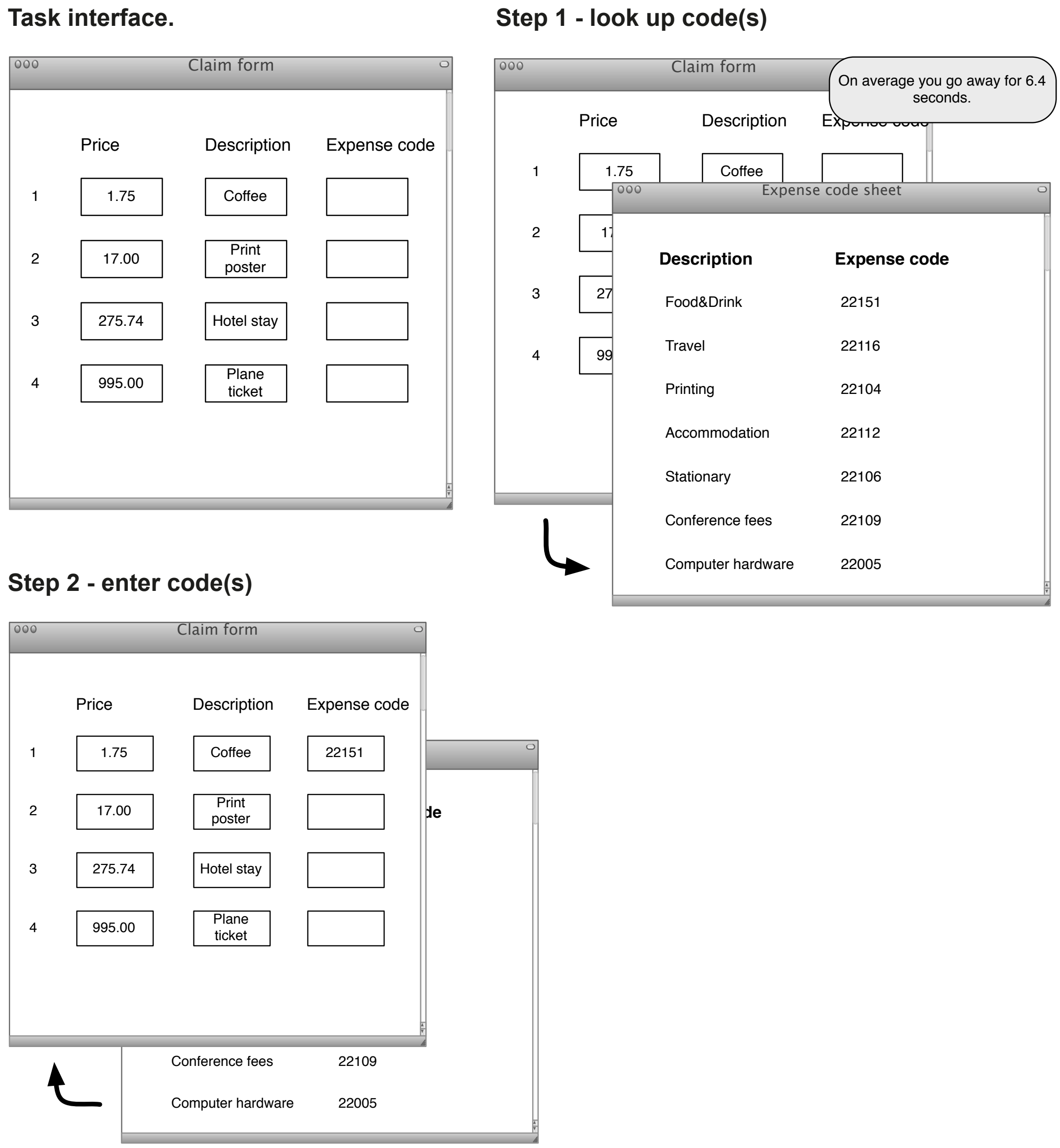


Figure 5.1: A wireframe to indicate what the interaction with the intervention will look like. The upperleft corner shows the window of an expenses task the user has to complete. Information for this task is retrieved by switching to another window (Step 1). Upon every switch, a notification appears telling the user how long these switches are on average. The user then has to switch back and enter this information in the correct field (Step 2). Step 1 and 2 are repeated until all codes have been entered; the user can switch back and forth as often as is needed.

had to enter numeric codes into a form, which they had to retrieve from a message sent to their personal email. The information was presented as a message in participants' email inboxes, as email is an integral part of data entry work but known to be a source of distraction, and people often spend more time on it than originally intended (Hanrahan & Pérez-Qu, 2015; Mark, Iqbal, Czerwinski, Johns, & Sano, 2016a). It was therefore expected to have a distracting effect during the switches to look up information. Half of the participants received feedback on the average length of their switches through a browser notification. This study aims to investigate whether time feedback on interruption length reduces the number of switches, reduces the duration of switches, makes people faster in data entry, and makes people more accurate in data entry.

5.3.2 Method

Participants

Fourty-seven participants (30 women) took part in the online experiment. Ages ranged from 20 to 63 (M = 29.3 years, SD = 9.1 years). The participants were recruited via university email lists, social media and online platforms to advertise academic studies, and participation was voluntary. Participants were alternately allocated to the control or experimental condition.

Design

The study used a between-participants design with one independent variable, a notification. In the control condition, participants did not receive a notification, but switches away from the data entry window were recorded. In the notification condition, participants were shown a notification every time they completed a trial. This notification showed how long on average they were away for when switching away from the window, before returning to the task. The purpose of this notification was to see if the number and duration of switches could be reduced by giving participants feedback on the time spent of on switches. Dependent variables were number and duration of switches away from the data entry interface, trial completion time, and data entry errors. Switching behaviour was recorded using JavaScript's blur and focus events. These were triggered whenever a participant switched away from the data entry window, whether to their email inbox or to a different window or application.

Materials

The task used was based on a common routine data entry task from Study 1 and 2 involving processing expenses. Participants were presented with an online sheet containing a set of ten 'expenses' (see Figure 5.2). They had to complete each row by entering the correct expense code for the expense. They retrieved this code by looking it up in a table of 25 expense categories which each had a corresponding 5-digit expense code, shown in Step 2 of Figure 5.2. Participants had to determine which category an expense belonged to, look up the code of this category and enter it in the row of the expense. Expense categories and codes were used that are currently used by one of the universities from Chapter 3 to process expenses.

In the example in Figure 5.2, the expense in the top row belongs to the category 'Postage' and the participant would have to copy the code 22104 from the expense table into the empty cell of the top row. A code did not occur more than once in a trial. The codes within a trial could be entered in any order.

Once the codes of the ten expenses had been entered, participants clicked the Next button to go to the next trial and the sheet was filled with ten new expenses. In the notification condition, a browser notification appeared at the end of each trial at the right-hand corner of the screen that told participants the average duration of window switches away from the primary data entry task. The notification stayed visible for several seconds (a default set by the browser), or until dismissed by participants (by clicking on it).

Participants were not alerted to any mistakes and once they had pressed 'Next', they could not return to the previous trial to correct any errors. Participants had to complete one practice trial, and five experimental trials. The purpose of the practice trial was for the participant to get familiar with the task, and the recorded data from this trial was excluded from the analysis.

Procedure

The study was advertised online with a brief description and a website link to sign up. Participants signed up for the experiment by entering their email address, and were sent an email with the table of expense categories and expense codes. The email also included instructions with a new link where the study was available. Participants were asked to complete the task on a desktop or laptop computer and open the experiment in Google Chrome, Firefox or Safari. Participants were not informed beforehand which condition they had been allocated to, and

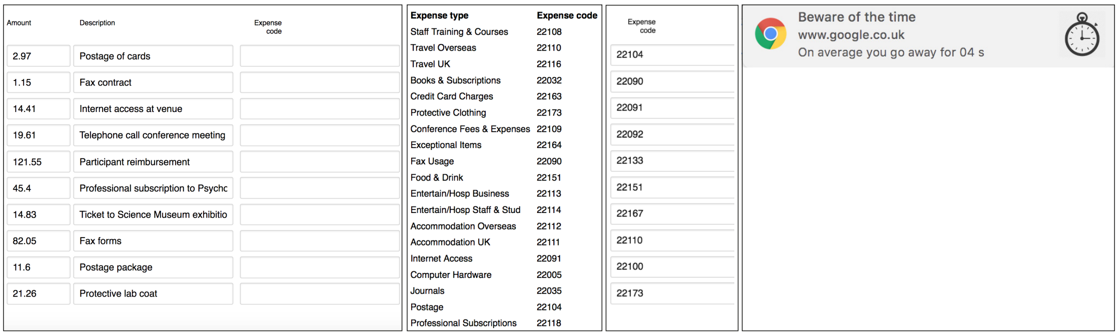


Figure 5.2: The data entry task as shown in the browser. Participants had to look up codes from their email (Step 1) and enter this into a sheet (Step 2). After every trial, the notification condition received time information (Step 3)

were told the purpose of the study was to understand how people perform data entry tasks. Participants in the notification condition were informed that they would receive notifications during the experiment.

Participants first read an online consent form on the website, and were not able to continue to the experiment until they had given explicit consent to participate. Participants in the notification condition received an additional dialog box to enable notifications in their browser, and had to click 'OK' to continue. Participants were instructed to have both their email and data entry window open on the same device, and to keep both windows maximised at all time, to ensure they had to switch back and forth between the two windows.

After completing all experimental trials, participants were shown a page of debriefing information, explaining the purpose of the study. An email address was included as a point of contact if participants had any further questions. Participants took between 10 and 20 minutes to complete the experiment.

Pilot study

A pilot study was conducted with four colleagues to test the study set-up. Initially, the notification was set to appear on every switch, and showed the duration of the last interruption, instead of the average interruption time. Participant 1 looked at the notification at the start of the study, but experienced that the notification interfered with information she was holding in working memory. Upon switching, she had to memorise what expense category code she had to look up in the email, and looking at the notification made her forget what she was looking for, forcing her to go back to the data entry interface to look it up again. She therefore tried to ignore the

Condition	Number of switches	Duration of switches	Error rate	Trial completion time (s)
Control	10.65 (1.67)	7.13 (3.05)	5% (5%)	126.27 (32.61)
Notification	10.61 (1.83)	4.76 (1.65)	2% (2%)	107.61 (31.15)

Table 5.1: Means and standard deviations of dependent variables for each condition.

notification for the rest of the study. For the remaining three pilot studies, the notification was adapted to only appear once after every trial, and show the average interruption duration. Two participants piloted the notification condition, and one participant piloted the control condition. Participant 3 used the information from the notification to try and find the codes in the email quicker, and consulted the notification after every trial to see whether his switches were shorter than the previous trials. The expenses task was experienced by participants as a realistic task, and all participants glanced at new incoming emails during the study.

5.3.3 Results

Table 5.1 summarises the results of the conditions in terms of the four dependent variables. The number of switches, length of switches and the error rate were not normally distributed, so non-parametric Mann-Whitney tests were used to analyse effects of a notification on these dependent variables. A Shapiro-Wilk test suggested that the trial completion times were normally distributed, $W = 0.97$, $p = 0.22$, so an independent t-test was used to analyse the effect on trial times. A p-value of 0.05 was used for assessing the significance of all statistical tests.

Cleaning the data

In total, 87 participants signed up for the study. Thirty-four participants did not complete the task and their data was excluded from the dataset. Furthermore, six participants made no recorded switches and were excluded from the dataset as well. Data from the remaining 47 participants was used in the data analysis.

Task performance

Participants with a notification were faster in completing trials ($M = 107.61$s, $SD = 31.15$s) compared to participants without a notification ($M = 126.27$s, $SD = 32.61$), $t(45) = 1.98$, $p < .05$, $d =$

Duration (s)	Control	Notification
20-25	39	11
25-30	7	6
30-35	4	0
35-40	1	5
>40	8	3

Table 5.2: Total number of switches longer than 20 seconds for each condition.

.59. Error rates were calculated by dividing the number of data entry errors divided by error opportunities. The error rates were significantly lower for participants with a notification (M = 2%, SD = 2%) compared to participants who had no notification (M = 5%, SD = 5%), $U(24, 23) = 403$, $p < .01$, $r = .44$.

Number and duration of switches

As can be seen in Table 5.1, participants who received notifications made significantly shorter switches (M=4.76s, SD=1.65s) than those in the control condition (M=7.13s, SD=3.05), $U(24, 23) = 406$, $p < 0.01$, $r = .44$. The number of switches per trial was on average 10.6 in both conditions, and there was no significant difference in number of switches between conditions, $U(24,23)=243$, $p = 0.6$. As there were ten codes to be entered per trial, this suggests participants switched once for every piece of data entered.

Figure 5.3 shows the distribution of switching durations for the Control and Notification condition, the red line marks the mean duration. For both conditions, the distribution was positively skewed with a long tail: 97% of the switches were under 20 seconds, but the longest switch was greater than seven minutes. To scale the distribution in one histogram, switches longer than 20 seconds are grouped as one bar. Table 5.2 shows the count of these long switches for each condition, and highlights that there were some occurrences of very long switches, where participants were likely distracted.

Interkey intervals

The primary measure to analyse switching behaviour were focus and blur events. These measures include any switch from the task window to another computer window. While this pro-

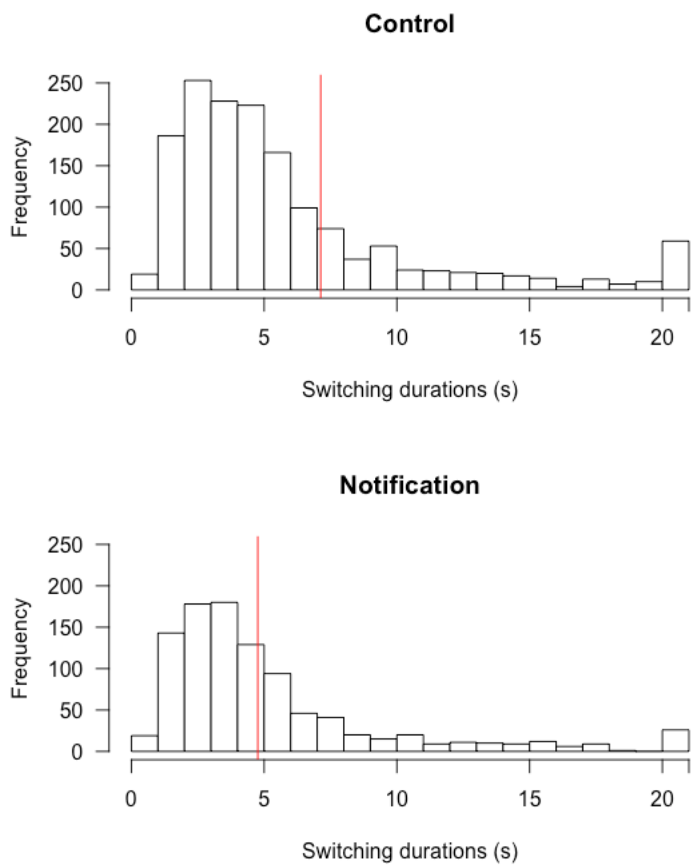


Figure 5.3: Histograms showing the distribution of switching durations for the two conditions; switches longer than 20 seconds are grouped in one bar at the right side of the histograms. The red line marks the mean switching duration.

vides a good measure of digital switching behaviour, it cannot capture task switches outside the device because the task window remains in focus during these task switches (e.g., a user might pause to fetch a paper document or make a cup of coffee). To help capture this broader range of instances of possible task switching behaviour, I analysed longer pauses in task activity captured by an analysis of inter-keystroke interval (IKI) data. Though these intervals may have also been moments where participants had briefly paused for thought, extremely long intervals between two keystrokes may point to moments where a participant switched to doing something else. The IKI data presented here excludes intervals where a window switch was recorded, as these moments have already been analysed in the previous section.

There was no significant difference in duration of IKIs between the Control (M = 1.70s, SD = 0.91s) and Notification (M = 2.02s, SD = 1.60s), $U(24, 23) = 261.5$, $p = 0.9$. The mean IKIs are considerably longer than number typing speeds found in prior studies, which ranged from 180ms to 500ms (Gould, Cox, & Brumby, 2016). The higher mean IKIs can be explained by looking at the whole distribution of all IKIs, as shown in Figure 5.4. As can be seen from these histograms, the distribution was very skewed with a long tail which affected the mean IKIs. When considering the median rather than the mean, the median IKI was 192 ms in the Control condition and 305 ms in the Notification condition, which is more similar to a typical typing speed. Figure 5.4 shows that the majority of IKIs were under one second. There were however some instances when there were long delays between keypresses: the longest measured IKI is four minutes. In the histograms, IKIs longer than five seconds are grouped in one bar. To give a closer view of longer IKIs, Table 5.3 shows the frequencies of long IKIs. These long IKIs were more than two deviations from the mean, and may have been additional task switches. However, we do not know for certain what people were doing during these instances, and what an appropriate IKI threshold would be to safely assume people had made a task switch. Therefore, we mainly focus our conclusions on our analysis of explicit window switches, and merely present the long IKIs to indicate that in addition to window switches, there may have been additional moments where people switched tasks.

Linear regression model

Participants who made longer switches took longer to complete trials. To see whether long switches also made people slower to resume a trial, we also consider trial completion times where the time spent on switches has been subtracted. For example, if a trial took 100 seconds,

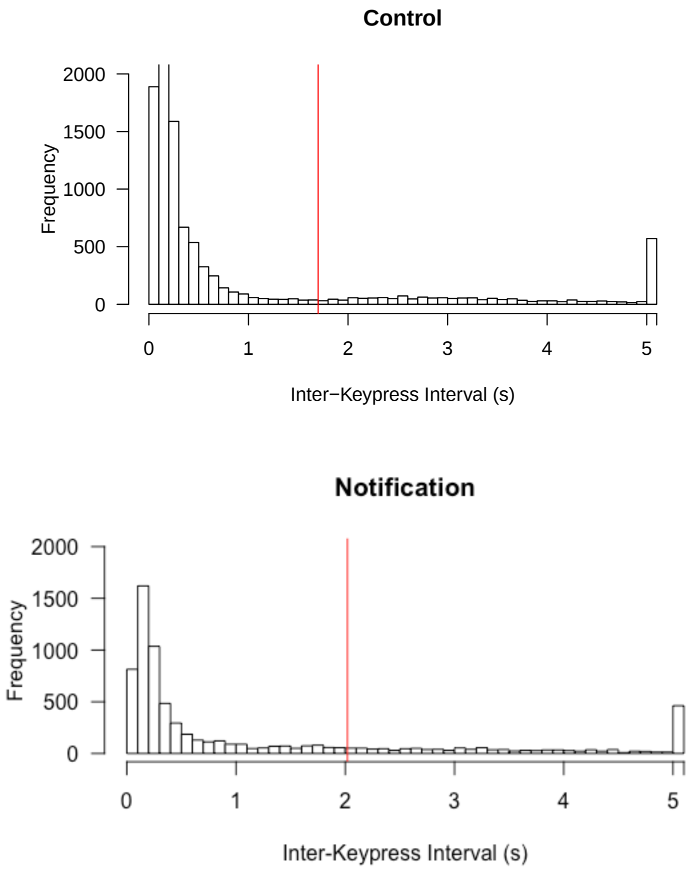


Figure 5.4: The distribution of inter-keypress intervals for the two conditions; switches longer than five seconds are grouped in one bar at the right side of the histograms. The red line marks the mean IKI.

Duration (s)	Control	Notification
5-10	399	141
10-15	96	36
15-20	33	15
20-25	18	11
>25	25	28

Table 5.3: Total number of IKIs longer than 5 seconds for each condition.

	β	SE Error	*t*	*p*
Number of switches	4.7002	2.14	2.20	0.03*
Duration of switches	4.2992	1.33	3.24	0.002**
Number of IKIs	0.8694	0.31	2.85	0.007**
Duration of IKIs	3.8001	3.15	1.21	0.24
Error rate	51.06	85.79	0.60	0.55

Table 5.4: Predictors of regression model that predicts task completion time. An asterisk (*) indicates a predictor is significant at $p < 0.05$; a double asterisk (**) indicates a significance at $p < 0.01$ level.

but the participant spent 20 seconds outside the task window, the adjusted trial completion time is 80 seconds.

With the adjusted trial times that has switching durations subtracted, we built a regression model to see if there was a potential relationship between trial completion times and number and duration of switches, as well as number and duration of IKIs. The linear model explained a significant amount of variation in trial times, ($R^2 = 0.41$, $F(5, 41) = 5.68$, $p < 0.01$). Table 5.4 shows there were three predictors that explained the variation: the number and duration of switches, and the number of IKIs.

This model provides further insight into the effect of longer switches on task performance: the longer people switched for, the slower they were to finish a trial, even when switching durations are subtracted from the total trial time. The number of switches explained a smaller portion of the variation, but still had a significant effect. Furthermore, the more people typed, the slower they were to complete a trial.

5.3.4 Discussion

The aim of this study was to see whether showing people time feedback on how long they switched away from a task on average reduced the number and length of their inquiries. The results show that participants made shorter switches, were faster to complete the task, and made fewer errors. These results are important, as shortening interruptions may provide substantial benefits in work: it can make people less distracted and more concentrated on their current task, which can lead to higher productivity (Mark et al., 2018; **?**). The improvements in data entry accuracy and task completion time are also in line with previous experimental studies showing shorter interruptions improve task performance (Altmann et al., 2017; Monk et al., 2008). In these experimental studies, the length of the interruptions was controlled by the experimenters. The current study contributes to this body of work by exploring ways to shorten self-interruptions, where the length of the interruptions is controlled by users themselves.

Long switches significantly increased task completion time, even after subtracting the switching times. This can be explained by the resumption lag (**?**): the longer people are away from a task, the longer it takes them to resume the primary task, as it takes longer to remember where people were in a task. In addition to window switches, there were also some long pauses between keypresses, which suggests that people further self-interrupted themselves outside of the computer, with the task window still in focus. However, it is difficult to know for certain what was happening during these moments in a remote study. Without an accurate estimate of how long participants should take to complete the task, it is difficult to determine moments at which participants were away from their computer (Rzeszotarski, Chi, Paritosh, & Dai, 2013). Future studies could use additional metrics to explore what people are doing during long pauses: for example, the data entry interface may prompt the user to confirm they are still working on the task, after a certain amount of inactivity.

Most experimental studies on self-interruptions have used an artificial distraction, such as chat messages, to measure how people self-interrupt to attend to this distracting task (Katidioti & Taatgen, 2013; Salvucci & Bogunovich, 2010). The current study makes a methodological contribution by using participants' own personal email inbox, based on the assumption that email provides a natural source of distraction (Hanrahan & Pérez-Qu, 2015; Mark et al., 2016a). The benefit of using an experimental task is that it enabled me to log data entries and measure a direct effect on task performance. In the current study, participants only needed to find and open an email once. Once they had this email opened, they did not have to re-find it in their inbox for the remainder of the experiment, and may have had this email maximised on their screen, hiding incoming messages. In practice however, people have to first find the email in their inbox, which can partly contribute to the distraction. This study has already shown an effect on behaviour by switching to an email inbox. It is expected there might be a higher potential for distraction if people have to also find the correct email in their inbox.

Giving participants feedback on the duration of their switches did not reduce the overall occurrence of switches as in prior work (Gould, Cox, & Brumby, 2016). A reason for this difference in results may be that in prior work, feedback was given after every switch, whereas in this study feedback was only given after every trial. As participants had to switch regularly as part of the task, giving notifications at every switch would have had the risk of overexposing participants to notifications and limiting its usefulness (Cutrell, Czerwinski, & Horvitz, 2001; Whittaker et al., 2016). When piloting the study, a notification after every switch was found to be too disruptive, and interfered with memorising what people were interrupting themselves for. Therefore, feedback was only given after every trial. A future study could explore showing people the number of times they switch away from a task, in addition to the duration of switches. As people work to maximise their task performance based on the explicit feedback that they are given (Farmer et al., 2017), showing people the number of switches away from a task may encourage them to reduce the number of switches.

Conclusion

The results of this experiment indicate that showing people how long they switch on average reduces the duration of switches and can improve people's task performance. The work makes a contribution to our understanding of switching behaviour for routine data entry tasks to distracting, but task-relevant, applications such as email. The results also suggest ways in which tendencies to attend to distractions might be mitigated, and can provide a useful pointer for the design of productivity interventions to improve focus.

One of the factors that influence the costs of a work interruption is its length: while short interruptions can be beneficial to productivity, too many interruptions and for too long take valuable time away from work (Mark et al., 2018). Based on this, together with the finding that time feedback reduces the duration of interruptions, it is expected that time feedback can make people more focused and productive in their work. In the current study, an experimental task was used in order to measure task performance. The focus of the study was on the effect of time feedback on duration of interruptions, and did not explore people's underlying motivations for their behaviour, and whether people indeed felt more focused or productive. Perhaps people were motivated to be focused on the task as they knew it was part of an experiment.

To investigate whether the effects of time feedback generalise beyond an experiment and are of real benefit to people working in office settings, Study 7 tested the notification with office workers doing their own data entry work.They were asked to install and use the notification on their work computer for two weeks, and were interviewed afterwards on their experience of using the tool.

5.4 Study 7: Looking up information for expenses in an office setting

5.4.1 Introduction

In order to understand whether showing people how long they go away from a task would have the same effect on a naturalistic data entry task, Study 6 was followed up with a field study testing the notification with data entry workers doing expenses work. To measure self-interruption behaviour during their work, participants were asked to install a free trial version of ManicTime [2] for two weeks. ManicTime is a time tracking software, which tracks application and web page usage. In addition, participants were asked to install a browser extension, and use it when they are processing expenses. Every time participants switched away from the browser window in which they did their expenses work, the extension showed a notification similar to the notification used in Study 6. The purpose of the study was to see whether a notification had an effect on self-interruption behaviour. To get a quantitative measure of self-interruption behaviour, ManicTime data was used to derive number and duration of window switches during expenses work. In addition, participants were interviewed to explore whether and how the use of both the extension and ManicTime led to any conscious changes in their behaviour. The study aimed to address the following question: how does time feedback on interruption length have an effect on people's self-interruption behaviour during expenses work in a finance office setting?

5.4.2 Method

To study the effect of a notification on people's self-interruption behaviour, I originally planned to use a between-groups design: participants would be divided into a control group and experimental group. The control group would be asked to install ManicTime for two weeks. They would be told the purpose of the study was to understand how people in offices manage tasks, windows and applications. The experimental group would also be asked to install ManicTime, but would additionally be asked in the second week to install the browser extension. They would be told that this extension would give additional information on the current task they were working on. Other than this distinction, all instructions would be identical between the

[2] https://www.manictime.com

two groups. However, upon conducting the study, five participants were unable to install ManicTime on their work computer due to security restrictions. Therefore, there was insufficient quantitative data collected to make a comparison in switching behaviour between conditions. Instead, the study's focus became the qualitative data collected during the interviews, and all participants used the browser extension in the second week. Participants who were able to install ManicTime did so at the start of the first week and kept it running for two weeks. In the second week, all participants were asked to install the browser extension.

Participants

Nine participants (six female) took part in the study. They were office workers at finance administration offices at one of the public universities from Chapter 3, and were invited to participate via emails sent to departmental mailing lists and snowballing. Participants worked in an open plan office, and seven participants occasionally worked from home. Participants' work included administrative and supportive tasks, such as processing payments, expenses, managing budgets, and responding to queries by university staff and students. The majority of participants' work was carried out in a web browser, and revolved around a number of web-based data entry systems. None of the participants had used a time or task management tool before. Participants were reimbursed with a £20 Amazon voucher after completing the study.

Materials

The notification was implemented as a Google Chrome extension using HTML, JavaScript and CSS. After installing the extension, an icon was permanently visible in participants' browser (see Figure 5.5). To use the extension, participants had to navigate to a web page in their web browser that they wanted to focus on, and click on the icon of the extension. Upon clicking on the icon, a pop-up appeared saying that the current web page was now the main task page, which indicated the start of a task session. Every time participants switched away during the session from this web page to another computer window, such as a different browser window, a document or an application, they received a notification indicating how long on average they go away for when switching away from the main task page. If participants switched away from a page for the first time, the notification showed a message that no switching data was available yet. To calculate the average switching duration, the extension recorded and saved the number and duration of switches away from the main task page for the whole session. Participants

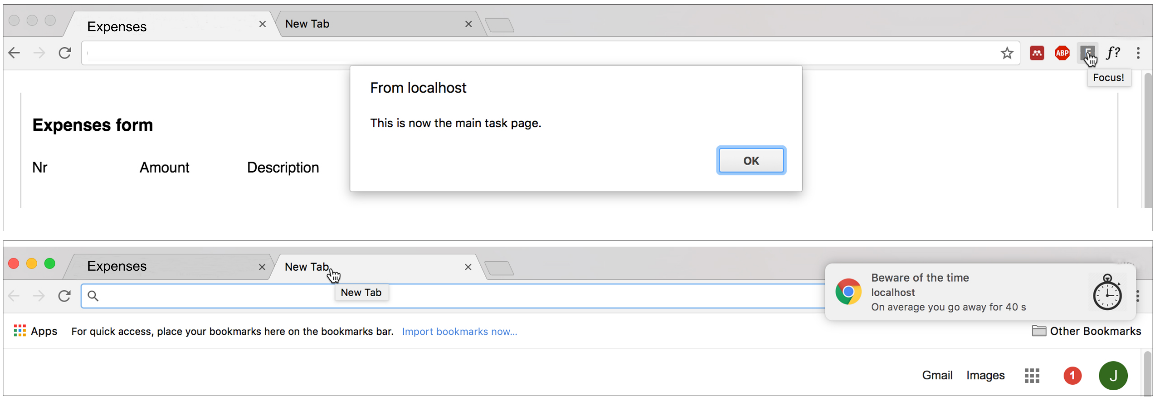


Figure 5.5: Participants selected a web page to focus on by clicking on the extension icon in their browser, after which a popup appeared to confirm this was now the main task page. Every time the participant switched to another window, application or document, a browser notification showed how long on average they switch away from this page.

ended a session by closing the page. Due to security restrictions of browser extensions, the extension was unable to save any session data after a session had ended.

The presentation of the notification was similar to Study 6 but differed in one important aspect. Whereas the notification in Study 6 appeared once after every trial, in this study it appeared upon every switch away from the task. Based on the observations and interviews reported in Chapter 3, I anticipated participants switched less frequently for their main work compared with the experimental task, and therefore a notification at every switch was not considered to be too disruptive.

To get an understanding of people's interruption and window switching behaviour, participants were also asked to install ManicTime, a computer logging software which records and stores the time spent in all application windows. I initially intended to give the extension to only half of the participants to see if there was a notable difference in interruption behaviour between people who used the extension compared to people who did not. However, five participants were unable to install ManicTime on their work computer, and could only use the extension. Due to this lack of quantitative data to make a fair comparison, I therefore distributed the extension to all participants and mainly focused on the interviews and people's experience of using the tool. A summary of ManicTime data of the remaining four participants (P3, P4, P5 and P9) is included in this chapter and used to complement the qualitative interview data and give an insight in the fragmented nature of people's work.

Procedure

Participants who expressed interest were sent an information sheet and consent form to read and sign. They were sent an overview of the study, instructions to install the tools, and a post-study interview was scheduled. The study was divided into three stages:

Week 1: Install ManicTime

In the first week, participants were sent instructions to install ManicTime on their work computer. They were given the option to pause or stop the application from running at any time. They were told that they were free to choose if, when and how often to look at the information, but that it was important to complete at least one expenses task with the application running.

Week 2: Install the browser extension

In the second week, participants were asked to install the extension. Again, they were instructed that they were free to choose when and how often to use the extension, but that they had to use it for at least one expenses task.

Week 3: Interview

After two weeks, participants were interviewed about how they currently manage documents, applications and tasks for their work, and asked questions on their experience of using the tools. In particular, it was discussed whether and how they used or would use the information that the tools provided, and whether they made any changes on how they went about their work. They were asked to share their ManicTime database for further analysis. Participants were offered guidance and assistance on deleting or adapting any data in their database, such as removing application and website names. Participants were still eligible to participate, if they did not wish to share their database.

The interviews were structured around the following themes: how participants currently manage interruptions, tasks, time and information, the context of using the extension, the usefulness of the information provided by the extension and ManicTime, and whether they made any changes on how they managed their work. Participants who did not install ManicTime were presented with screenshots during the interview, and discussed the usefulness of this type of information compared to the time information of the extension. An interview lasted about 60 minutes and was audio recorded.

Pilot study

A pilot study was conducted with two participants. One participant was a colleague and was sent instructions to install the browser extension on his computer. The purpose of this pilot session was to see whether the installation instructions were clear and to test the extension on other people's computers. The participant had the tendency to hover over the notification to stop it from disappearing, which placed extra buttons over the end of the notification message. Therefore, the message was shortened and the most important information, namely the duration of switches, was placed at the front of the message to ensure it was still visible upon hovering.

A second pilot session was then conducted with an administrator working at the same university as the study participants, who was asked to install ManicTime and the extension on her work computer. The purpose of this session was to ensure the tools could be installed on the work computers of the university, and that the extension worked with the university finance system.

Data analysis

Though ManicTime was piloted on a work computer of the university before starting the study, five participants were unable to install ManicTime on their work computer due to firewall restrictions. It appeared that different participants had different computers, operating systems and firewall settings. Therefore, ManicTime data of the remaining four participants was used to complement qualitative explanations of their task switching behaviour, but it was not analysed quantitatively as previously planned to compare switching behaviour with and without the extension. Instead, the primary focus of data analysis was on the post-study interviews and participants' subjective experience of using the tools. The interviews were transcribed verbatim, and analysed using thematic analysis.

5.4.3 Findings

Participants gained some insights to change their behaviour based on the information they received from the extension. People's switching behaviour as shown by the ManicTime data is reported first. I then discuss the usefulness of time feedback to manage interruptions around the following themes: awareness and change of behaviour, the type of interruptions, the effort

Duration (s)	Mean	Median	Min	Max
Window focus duration (s)	33.88 (80.74)	11.00	1.00	2893
Daily switches between windows	829.5 (422.85)	843	9	1741
Non-digital interruption durations (s)	1741.09 (1886.17)	992.5	47.00	10457
Daily non-digital interruptions	2.74 (1.96)	2	1	7

Table 5.5: Average window focus durations (s) and number of daily switches.

to record and use data, setting goals, and the work environment.

Switching behaviour

Participants' working hours differed slightly, but all participants worked at least ten hours per day during the study. To make the data comparable between participants, I only considered data between 9am and 7pm, during which all participants were at work. As I was only able to gather logging data from five participants, there is insufficient data to make a fair comparison between switching behaviour without and with the extension. Summary data is presented here to give insight into people's switching behaviour during the study, but these are not analysed using any statistical models.

Table 5.5 summarises the average number and duration of focus on a computer window screen. The mean duration of focus is about 34 seconds, with the longest focus being 48 minutes (2893 seconds). On average, participants made 830 computer window switches per working day. Figure shows the distribution of all window focus durations. Almost 90% of the window focus durations were less than a minute. To scale the histogram, window focus durations longer than 150 seconds are grouped in one bar; the frequency of these longer durations are shown in Table 5.6. Together with the interview findings, the data shows that participants' work was characterised by short durations of focus and frequent window switches. Figure 5.7 and Figure 5.8 show the average number of daily window switches and focus durations over the ten days of the study.

In addition to computer window switches, participants also made a smaller number of non-digital interruptions, for example when taking a break or attending a meeting (see Table 5.5). On average participants made three daily non-digital interruptions which lasted about 29 minutes (1741 seconds).

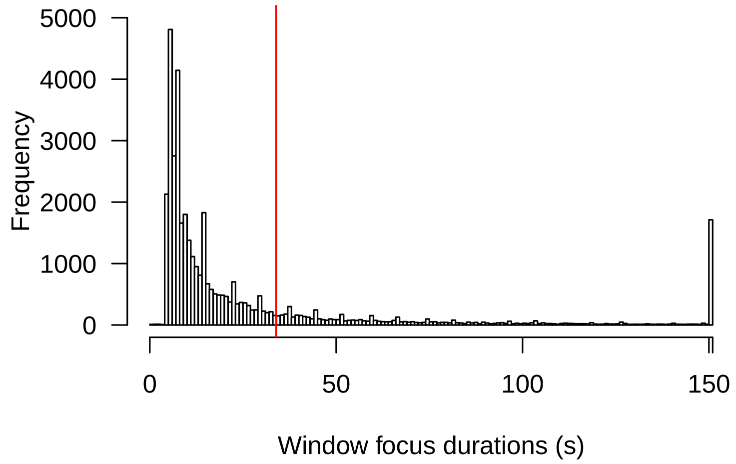


Figure 5.6: The distribution of window focus durations during the study, measured by ManicTime. Durations longer than 150 seconds are grouped in one bar at the right side of the histogram. The red line marks the mean window focus duration.

Duration (s)	
150-250	874
251-350	388
351-450	202
451-550	98
>550	156

Table 5.6: Total number of occurrences that a window was in focus for longer than 150 seconds.

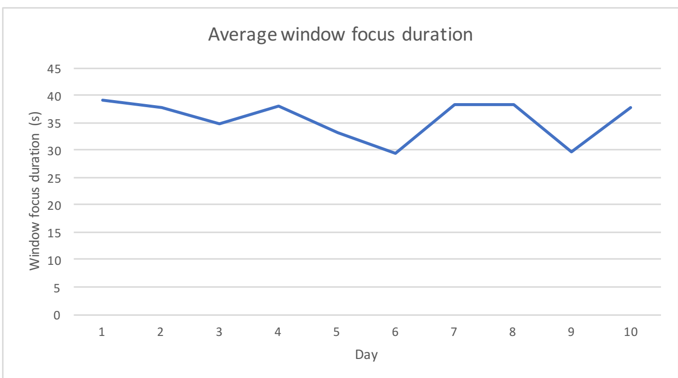


Figure 5.7: Average window focus durations during the study.

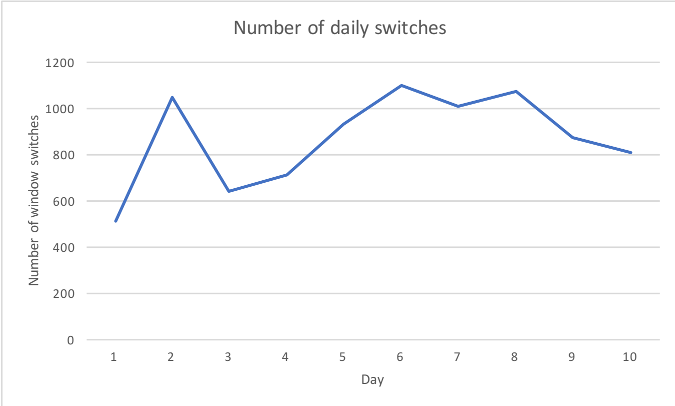


Figure 5.8: Average number of daily window switches during the study.

Awareness of interruption behaviour

Participants were largely aware they interrupted their work frequently and considered it the nature of their job: they regularly had to stop their work to look up task-related information, and had to address ad-hoc queries and requests from their department. The extension made participants realise however that they were unaware of the length of some of these interruptions. The average interruption time was considered much longer than they thought. Interview results suggest that common reasons for interruptions being longer than they thought were distractions and chains of diversion (Hanrahan & Pérez-Qu, 2015; Iqbal & Horvitz, 2007), where the user further self-interrupts for other tasks. Participants tried to avoid interruptions during work that were completely unrelated, but after they had interrupted themselves for work purposes, there were opportunities to further self-interrupt for other off-task activities. The notification made people more aware of the effect this had on the duration of their interruptions:

"It's a shock, because I knew it was bad, I didn't think it was that bad. (...) So it's reflecting on, actually, a two-minute task is turning into a 15-20 minute task - why is that? (...) Why? But again, it's distractions." (P9)

"It made me realise how long I was spending, spending/wasting, doing other stuff. (...) It's just going to do something and then ending up chatting with someone." (P3)

Reflecting on actions during an interruption

The increased awareness of time spent on interruptions caused people to reflect on what they were doing during past interruptions. Some interruptions were urgent, important, or necessary to progress with work, and therefore hard to avoid altogether. However, reflecting on the exact actions during the interruptions made people realise that some interruptions could be shortened, as participants often ended up getting diverted from the original goal of the interruption. For example, upon switching to their email inbox to retrieve information, participants would get diverted by reading and responding to other unread messages instead. To help remember what was happening during an interruption, P9 combined the extension with the data of ManicTime: *"[The extension] popped up and it said: "You go away for 7 minutes and 33 seconds. I would then have a browse [in ManicTime] And then I think: oh my gosh, I've been on emails for an hour! I haven't got anything done. So yeah, I checked it quite a lot. More so because I was so shocked. And so, I'm so interested to know, actually, what I'm doing at work."* (P9)

Having this insight into their actions, participants tried to be more focused on the goal of subsequent interruptions and be more wary of potential distractions. It also happened that the duration of an interruption was considered long, but justified. P7 was the only participant who, upon viewing the time information, was not surprised by the time she spent on work-related interruptions. She considered the amount of time necessary to complete her work and did not see any room to improve on this: *"To me, it doesn't kind of make me think: 'Oeh, I've been away too long'. I just think: OK, well I'm roughly aware that I've been away for an hour (. . .), I don't see how it kind of links with being more productive. Unless I suppose, you're really easily distracted."* (P7)

These findings show that awareness of time spent on interruptions can improve focus on the intention of interruption, and make people pay more attention to when they might get distracted. It also indicates that people may in particular benefit of feedback during interruptions where they are likely to get distracted, such as when switching to communication tools.

Reflecting on the relevance of interruptions

Some work-related interruptions were not urgent, but participants were used to addressing them anyway if they were presumed to be 'quick and easy'. These were addressed immediately so participants did not have to remind themselves to attend to it later, and it made them feel more productive if they completed more tasks. The notification made them more aware of the occurrence and actual length of these interruptions, and that they weren't as quick as they thought. For future interruptions, people tried to consider whether they needed to address the interruption immediately: *"I need to work on time management and (. . .) not spending my whole day answering irrelevant queries."* (P9)

Participants mentioned several sources they knew to be distracting, such as email, their phone, and colleagues. Some sources of distraction were not essential for work, and generally participants did not self-interrupt their work often with the intention to access these sources: for example, several participants said they did not check social media at work. However, other distracting sources were used both for work and non-work purposes, such as search engines, instant messaging tools and email. It was therefore difficult to simply eliminate these distracting sources from the work environment:

"As everyone says, 'we'll just switch email off' (...). But you can bet your life that there will come a moment in whatever task you're doing you think: Oh! I have to open up email. And the moment you open up your email, that's it." (P2)

"My phone is a distraction for me. (...) I put my phone in a tray under a load of documents. But then (...) I converse a lot with a professor via text." (P9)

This finding highlights the importance of giving people control over whether to address an interruption or not, as an interruption may be considered distracting but also essential for work. It demonstrates that an increased awareness of interruptions makes people reflect on why they interrupt and judge the relevance of the interruption, which may help in reducing unnecessary interruptions.

Time information for task management and perceived productivity

Completing tasks was an important component of people's work: they had an increased feeling of productivity if they explicitly ticked tasks off a list, and were driven by to-do lists and deadlines. While this could motivate people to focus on finishing a task before switching to another, it also had the contradicting effect that they interrupted their work often, if a task appeared that was considered easier to complete:

"It kind of contradicts what I told you before about (...) how I jump on them [incoming tasks] and finish them. But at the same time, it's because I don't want to have three things at once going. I want to finish, finish, finish." (P3)

The extension was started by the user deliberately selecting one window as a main task page, which forced participants to choose one task to focus on, and switching to other tasks was now considered an 'interruption' away from this task. As discussed earlier, the notification made people reflect more on whether other tasks they were switching to were relevant to address at that moment.

A clear interest among participants was to not only see how much time they spent on interruptions away from the main task, but also how much time they spent on that main task overall. Currently, participants planned tasks they wanted to complete on either a daily or weekly basis, and implicitly took the time each task would take into consideration. However, given the fragmented nature of their role and the frequency of interruptions, it was difficult to estimate how long they actually spent on these tasks:

"I think that might take me 3 hours, and I'd want to get that done in one day. But yeah, obviously, things quite often take longer than I think I will, because then when I'm doing them, I might get interrupted." (P5)

In the same way that they used time information to reflect on whether interruptions were as long as they thought they were, they wanted to reflect on whether tasks took as long as expected. They would use this insight to be more realistic when planning tasks over time:

"Down the line, I'd think it would be extremely useful to know how much time I'm actually spending [on tasks]. Because it would help me be more productive, or be more realistic in the amount of time I need for these things to happen." (P3)

These findings shows that both time away from a task, and time on a specific task, are important insights for workers to manage their work.

Setting goals for time limits

The notification was also used by participants to set goals on how much time they were willing to spend on interruptions. The way the notification was set up, participants did not know how much time they were spending on a specific interruption, and only saw the average interruption time. Therefore, several participants wanted to set time limits on each interruption. Similar to the relevance of an interruption, the appropriate length of an interruption was context-dependent as well: participants sometimes had to spend a relatively long time away from a task, for example if they had to find information in another window. It may therefore be more appropriate to give participants discretionary reminders to return to a task after they have reached a certain time limit, but then still give them control whether to adhere to that limit or not. Reaching the time limit could mean they were getting distracted, but it could also be the case that they were working on something relevant for work that needed more time:

"Say you have to work on that specific document, and then you end up spending half an hour on Slack, chatting to your colleagues, it would be good if something's like: mate, work. Stop doing other things. But it's really hard to know what people are actually doing on these things." (P3)

These findings show that people need not only be in control of whether to address an interruption or not, but also for how long. They may however benefit from time information during an interruption to help them learn how to self-adjust their behaviour.

Context of information

We asked participants about their use of the information provided by the extension versus ManicTime. Participants reported that the information provided by the extension was easy to read and interpret during a task. It was also clear what action to take, and participants used the information to decide whether they should reflect on past interruptions, and whether they could shorten the time away from their task. Participants looked at ManicTime at the start of the study out of curiosity and to make sure it was recording their activity correctly. However, in line with prior work (Collins et al., 2014), the extensiveness of the ManicTime data made it unclear to participants what action to take from the data, and most of them did not engage much with it for the rest of the study. It was considered too effortful and time-consuming to interpret and use the data:

"I didn't go into too much detail with it. One of the reasons is that, it would take me a lot of time and effort to use this information, to help me work better or quicker, or more efficiently. And this is either something that I don't have time to do, or I can't be bothered, depending on the day." (P3)

P9 did use ManicTime, in particular to help aid her reflection on what she was doing during past interruptions, a reflection which was triggered by the extension.

Participants commented that they would have liked to be able to see additional information of their interruption behaviour over time, in addition to interruptions during the current task they were working on, to place into context whether their interruption length was higher or lower than their usual behaviour. They would use this information to set realistic goals on interruption lengths, and to see how often they were meeting these goals. These findings show that short and actionable time information during a task can relieve users of the burden to interpret large amounts of information and easily adjust their behaviour. It also indicates that access to more information if needed is useful to get more insight into how people can change their behaviour, and whether they are achieving certain user goals.

Different work environments

Seven participants worked from home on occasion, and saved up tasks that required focused attention to complete at home as the office was seen as a more distracting environment. There was an implicit understanding within their department that working from home meant they needed to concentrate, and as a result, participants received fewer interruptions caused by queries from colleagues:

"You're working from home for a specific purpose, and therefore you don't really want to be disturbed. Unless it's absolutely urgent." (P2)

At the office, participants dealt with more interruptions taking place outside of the computer: for example, participants were interrupted by their colleagues or phone calls. Because the extension only provided time information about digital interruptions, some participants felt it provided an incomplete picture of their interruption behaviour. ManicTime provided participants with information on their non-digital interruptions, and participants considered this a good complement to the information that the extension provided. If the PC was inactive, and participants came back from inactivity, ManicTime presented users with a window on the screen saying how long they had been away for (see Figure 5.9), and gave participants the option to write down what they had been doing while they were away.

The office environment not only exposed participants to more external interruptions, but participants also self-interrupted more in the office: *"From home it's a bit different, I normally look at the emails but I generally try not to respond, unless it's too urgent. But at work, when I'm here, (...) if it is not too urgent, but still I can find that is nice and straightforward, I just straight reply back. But at home it's more focused, definitely."* (P8)

"When I'm at home, I generally don't look at my phone for some weird reason. (...) When I'm in the office I find that I'm easily distracted, and I don't get things done." (P9)

All seven participants reported there were more sources to get distracted in the office. For example, most participants had multiple computer screens and kept the majority of documents, browse windows and applications open on their work computer, even after they had finished with them. These windows were a further source of distraction if participants were trying to find task-related information in one of the windows:

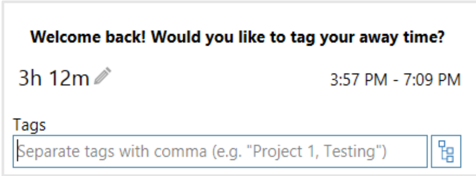


Figure 5.9: If participants had been away from their computer, upon returning ManicTime presented them with a window showing how long they had been away for, and an option to write down what they had been doing while away.

"It's like 15 tabs, and I need to go somewhere. And I end up clicking all of them. And if there is one that is personal stuff, I end up reading it. And then five minutes after, I'm like: what was I doing? (...) So it's distracting in the way that it makes me not solely focused on one thing." (P3)

When working in the office, participants tried to complete tasks that required focused attention in the morning. They were more easily distracted in the afternoon, as they received more external interruptions, such as email, phone calls, and colleagues. These differences in work environments indicate that people get more easily distracted in some environments and at certain times. Though participants only used the extension in the office, their descriptions of their office and home environments indicate that participants may in particular benefit from time information during afternoon work in the office, when participants were more prone to interrupt themselves and get distracted.

5.4.4 Discussion

The aim of this study was to investigate whether showing people how long they go away from a task has an effect on their self-interruption behaviour during data entry work in an office sestting. The main finding is that time feedback made people reflect on what they were doing during interruptions. As a result, participants said they were more focused on the intention of an interruption, and wary of potential distractions and diversions from this intention. Having insight into the time they spent on interruptions tried to avoid interruptions that were not relevant, and set goals for how much time they were willing to spend on interruptions that were relevant. Some participants indicated that they would like to see time spent on a task as well, to help them schedule their tasks and improve their productivity.

In line with previous work, the study indicates that an increased awareness of how people use time can improve focus on work (Mark et al., 2018; Whittaker et al., 2016). While previous work mainly found a reduction of time spent on non-work applications, a novel finding from the current study is that it can also benefit work-relevant interruptions: people shortened task-related interruptions to look up relevant information.

The finding that people reflected on what they were doing during an interruption, as well as on the relevance of an interruption, is important as it suggests that people become more aware and in control of the intention of their interruptions. Prior work showed that self-interruptions can be beneficial, such as a break from work, as long as the interruption is triggered by a clear intention from the user, and not an unintended distraction (**?**). The results of the current study indicate that increased awareness of interruption behaviour may make interruptions more intentional, and reduce unintended diversions during these interruptions.

In addition to time away from a task, participants wanted to see information of time spent on a task. Whittaker et al. (2016), who presented users with the last 30 minutes of their computer activity, found that giving people feedback on their use of time decreased the time spent away from work, but did not increase time on work, and speculated that people may have time limits to spend on work. Rather than increasing the amount of time spent on work, time feedback on task may be useful for people to better plan when and how they spend time on tasks. It may also make it even more apparent how interruptions slows down their work, as it is not only the time during an interruption, but also the resumption time after an interruption, that slows down work (**?**). This effect was also demonstrated in Study 6, where people who made longer switches were slower to complete a task, even after removing the duration times.

Participants were more distracted in the office in the afternoon, which is in line with **?** who found that boredom in the workplace is highest in the afternoon, and people are more likely to self-interrupt when they are bored. A further likely reason found in the current study was that it was busier in the office in the afternoon, and participants were exposed to more external interruptions: prior research found that an increase in external interruptions leads to more self-interruptions (**?**). In contrast, participants who worked from home said they were more focused at home and experienced fewer interruptions. This means that people may in particular benefit from time information at certain moments and settings, when people are more likely to get distracted. Participants only used the extension at the office and it is therefore unclear whether they would use it differently when working from home. Future research could further investigate people's self-interruption behaviour in different work settings.

Interruptions can have both benefits and costs, and it is important to consider the properties that make an interruption disruptive. The study highlights that whether or not to address an interruption, and for how long, not only depends on interruption properties but also the context in which an interruption occurs. For instance, email may be considered distracting and is best avoided in some situations, but in another situation the user may need to find information in email which is essential for the current task. Furthermore, the longer people interrupt a task, the more disruptive it is (Altmann et al., 2017; Monk et al., 2008), but in some situations it is important to find information and spend this time away from a task. Prior approaches to block self-interruptions to distracting sources (Kim et al., 2017; Mark et al., 2018) or impose time limits (**?**) are too restrictive in this situation. A more appropriate approach is to give users control to decide when and how long to address interruptions, and give them useful information to help them learn how to best control and self-adjust their behaviour. This echoes earlier conclusions by Mark et al. (2018) who found that blocking distractions at the workplace was experienced by several participants as too controlling, and participants rather wanted to learn to gain control of their work. The studies make an important contribution by showing how showing people feedback on interruption length can help in gaining control over time spent on interruptions.

Implications for design

Previous work has highlighted several problems with existing commercial time tracking and management applications: these often are time-consuming to use, they can restrict user activities too much, and it is not immediately clear to users what action to take based on the data (Collins et al., 2014; Whittaker et al., 2016). The findings from the current study partly corroborate these issues, and demonstrate several pointers that can inform the design of time applications.

Complement time information. First, when providing users with a data log of their computer activities, they need to have a specific starting point of what it is they want to find out for them to be able to use it and act on it. Participants were not interested in their overall computer activity, but were mostly interested in the time they spent on, or away from, a specific task. By presenting a simple and precise measure, namely the length of an interruption, participants were provided with a specific target of what to reflect on and change, and did not need to go through the effort of having to interpret information of all their activity. As some participants did want to have access to more detailed information about their activity during a specific interruption, a simple presentation in the moment can be complemented by a more complete log running in the background. It would also be interesting to give users control over what information they are interested in to see in the notification. For example, most participants were not only interested in the length of interruptions during a task, but also on the length of their task overall. This could help participants to better manage their tasks.

Time feedback during a task. Second, by showing information during the task, participants can react and change their behaviour immediately and do not have to remind themselves to look at information later (Gould, Cox, Brumby, & Wiseman, 2016; Maior et al., 2018). Participants were prompted by the notification to reflect on what they were doing during an interruption, but often forgot to look back at their computer activities on other occasions. Furthermore, participants in Study 6 were able to act on the explicit information they were given in the short time space of an experiment, which had a positive effect on their task performance.

Setting time limits and time goals. Several participants wanted to set time limits on interruptions, which has also been found in earlier work on interruptions (Mark et al., 2018; Whittaker et al., 2016). Based on the finding however that interruptions are context-dependent, imposing a strict time limit may be too restrictive. Rather, giving timed reminders to return to a task may make people more aware of the length of their interruptions, while still giving them control whether to actually return to a task or not.

Behaviour over time. Lastly, a promising area to investigate would be to record the interruptions and give participants insight in how their changes have an effect over time. Although it was clear to participants what action they had to take based on the data presented by the extension, some felt they did not have sufficient information as to whether their changes had any effect over time.

Limitations

While the results are promising, the study also has a number of limitations which would be worthwhile to address in future work. The notification only provides feedback on digital interruptions, but as was apparent both in Study 2 and Study 7, people also deal with interruptions and distractions beyond the computer. Future work could look at also collecting and showing data from these interruptions. For example, ManicTime uses PC inactivity to indicate when participants were away. Other sensitive measures to detect moments where the user has likely interrupted their work could be inter-key intervals or mouse clicks. Furthermore, due to the limited logging data, it is difficult to make any concluding claims as to whether time feedback had any significant effect on participants' window switching and task focus behaviour over time. In addition, though participants indicated they modified their behaviour after using the extension, it is not certain whether they based their behaviour on the specific information provided by the extension, or whether the notification simply made them reflect and become more aware of their time.

5.5 Summary of Chapter 5

The aim of this chapter was to investigate whether showing people how long they go away from a task has an effect on their self-interruption behaviour. The findings demonstrate that time feedback can reduce distractions and improve task performance. Study 6 showed that it reduced the duration of switches, made people faster in completing a data entry task and reduced data entry errors. Study 7 showed that time feedback made people reflect on what they were doing during an interruption. They avoided interruptions that were not relevant, and set goals for how much time they were willing to spend on interruptions that were relevant. They were more focused on the intention of an interruption, and wary of potential distractions and diversions from this intention.

The studies presented in this chapter contribute to our understanding of switching behaviour for routine data entry work to distracting but task-relevant applications such as email. The results suggest that a simple presentation of time information during a task can mitigate distractions but still keep users in control over their interruptions, and can inform the design of productivity interventions to improve focus. Showing users how long they go away from a task can increase awareness of interruption behaviour, which can reduce the duration of interruptions, shorten the completion time of tasks and reduce errors.

The next chapter brings together the findings from all studies, and further discusses the practical implications and the contributions that the findings make to knowledge. It also discusses opportunities for future work.

Chapter 6

General Discussion

Chapter outline

This chapter summarises the research findings of this thesis. It discusses the contribution that the findings make to knowledge, the practical implications, discusses limitations and suggestions for future work.

The aim of this thesis is to understand how people can better manage the time spent on inquiries needed for a data entry task, with the goal to reduce the disruptiveness of inquiries for the task. The results show that time costs can be used to encourage users to keep inquiries short, reduce the number of inquiries, and postpone them until a more convenient moment in the task. In this chapter, I first summarise the findings of each study. I then discuss the contribution of this thesis and the practical implications for tools aimed to help workers manage their interruptions in the workplace. Lastly, I discuss any limitations, outstanding questions, and how these could be addressed in future work.

6.1 Summary of findings

Chapter 3 reported an interview study and an contextual inquiry study looking at the context in which office workers in finance offices conduct data entry work, and how they self-interrupt to look up task-related information during this work. The studies had two contributions. First, there had previously not been many studies that tried to understand data entry behaviour in the uncontrolled setting of an office workplace. Interview findings from Study 1 revealed that a critical component of this type of work is not just entering the data, but collecting this information from various sources distributed in the physical and digital work environment. Second, observations in Study 2 revealed that there was a difference in how physical and digital information is collected. Physical sources take time to access and participants therefore prepared it beforehand, or postponed retrieving it until a more convenient moment in the task. However, computer window switches during the task were commonly observed as workers often realised during the task that they needed additional information, and they presumed these switches to be quick. These switches often took longer than intended, and participants were observed being logged out of the entry system, resuming the wrong data entry task, and reported it took time to resume their work after these longer switches. The hypothesis was made that presumed time involved to collect information played an important part in how people address self-interruptions, and that workers were unaware of the time they actually spend on digital interruptions. Though these qualitative studies gave a better understanding of how data entry work is situated in an office setting and people's self-interruption strategies during this work, the extent to which these strategies were influenced by time costs is unclear. Therefore, a series of lab experiments was carried out to study the effect of time costs on people's switching behaviour in a controlled setting.

Chapter 4 reported three controlled experiments to test the hypothesis that people prioritise collecting information according to time costs: they first switch to information with low time costs, and postpone collecting information with higher time costs. These studies showed that, in a controlled setting where participants can learn the time costs involved in accessing information, they first switch to information sources that are fast to access, and switch more frequently to these sources. On the other hand, people either prepare or postpone looking up information which takes time. Study 3 showed that if people retrieve all data from the same source, they will reduce switches between entering and looking up data if the access costs to this source increases. As it took more time to access, offloading behaviour was observed as well, and several participants prepared items they were going to need nearby, but did not use them yet. Study 4 further demonstrates that when people have to retrieve data from multiple sources, they collect and group items that are quick to access first, and leave items that take longer to access until the end. Study 5 demonstrated the robustness of the effect of time costs in a multi-task setup: when dealing with two data entry tasks, people still entered items with a low time cost first, and interleaved between tasks to enter items with low costs first. As a result, participants made more omission errors and submitted tasks before they had completed entering all the items. These studies contribute to our understanding of time costs on self-interruption behaviour to collect information: if people know the expected time duration of an interruption, they make fewer interruptions that are long and postpone these switches. However, what remained unclear after these studies was how people can learn time costs in a naturalistic setting. An issue with self-interruptions to collect information is that time costs are not always predictable: there are various opportunities to get distracted, and people may spend a longer time than they think or expect looking for information.

Chapter 5 reported an online experiment and a field study looking at whether making people more aware of time costs can be effective in managing self-interruptions outside a laboratory setting. This part of the thesis involved developing and evaluating a browser notification that showed people the average time they spend away from data entry work. These studies contribute to our understanding of how increased awareness of time spent on digital interruptions can help people be more focused on their work, and manage their time away from work. Study 6 found that using an experimental data entry task, people who were shown how long they were away for made shorter window switches, were faster to complete the task and made fewer data entry errors. Study 7 evaluated the intervention with office workers processing expenses. Data from post-study interviews indicated that time feedback made participants reflect on what they were doing during interruptions, and whether they could cut out unnecessary time.

6.2 Contributions

The findings contribute to our understanding of how time costs influence self-interruptions, and how information about time costs can help users self-regulate their interruptions. I first discuss the contributions this thesis makes to knowledge. I then discuss the practical implications these findings may have to inform the design of future studies as well as interruption management tools.

6.2.1 Contribution to knowledge

The first contribution is showing that **the time costs of an interruption influences people's self-interruption behaviour**. Prior work has shown that when people have to access information and the time cost to access this information increases, they reduce the number of switches, and increase the length of each switch (Gray et al., 2006). Chapter 4 extends this work and shows that if people deal with different time costs, they postpone switching to interruptions with the highest time costs.

The second contribution is to demonstrate how **feedback on length of interruptions helps people reflect on actions during, and reduce the length of, their interruptions**. Prior work has shown that people like to be in control of their own interruptions, and do not simply want to have distractions blocked (Mark et al., 2018), but that they often do not know what action to take based on reflective data (Collins et al., 2014; Whittaker et al., 2016). Chapter 5 showed that people are able to take action when they are given feedback on the length of their interruptions. Increasingly more applications are showing people how they spend their time in their applications, to better manage time on work-unrelated purposes (Constine, 2018a,b; Lynley, 2018). This thesis shows that time feedback can not only help in managing work unrelated interruptions, but also work-necessary interruptions to look up information.

The third contribution is showing that **the type of self-interruption influences people's decision-making in interrupting work**. Prior work has shown that task-irrelevant interruptions are more disruptive than relevant ones (Iqbal & Bailey, 2008). The studies in this thesis indicate individual differences between the ability to self-regulate task-irrelevant interruptions: whereas office workers in Study 1 and 2 were fairly good at avoiding task-irrelevant interruptions during data entry work, office workers in Study 7 did address some interruptions if they were presumed to be quick to complete. However, it is often the relevant interruptions that are needed for work that can be problematic for all workers: these cannot be avoided, are predominantly expected to be quick and easy, but can nevertheless be disruptive, as they can end up being time-consuming and there are various opportunities to get distracted.

6.2.2 Practical implications

This thesis makes a practical contribution by demonstrating that giving people feedback on the length of their interruptions influences their interruption behaviour: through a better awareness of the length of their interruptions, users reflect on what they were doing during these interruptions, and where possible tried to make them shorter. In this thesis, I focused on a particular type of work and setting: data entry work in financial administration offices. Managing self-interruptions however is not just important for this particular setting, as interruptions and distractions are common in many kinds of computer-based work (Gonzalez & Mark, 2004). Below I discuss the practical implications of my findings which can inform the design of future studies as well as the design and evaluation of interruption management technologies.

Regulating distracting but work-relevant interruptions. A common approach to improve task focus is to block any sources that are considered distracting from the primary task (Kim et al., 2017; Mark et al., 2018). However, this blocking approach centralises on the type of information source, but not the type of interruption: an information source may be distracting, but the interruption to access this source may be necessary for a task. Study 2 showed that people often need information sources they consider distracting, such as email, and cannot block these. This means we not only need to consider blocking interruptions that may be distracting from work, but also what support people can be given to control interruptions which are needed for, and considered part of, the task they want to focus on. Because it is difficult for a tool to determine why interruptions are being made (and whether they are necessary or not), it is better to give users tools to help self-regulate interruptions themselves.

Making time spent more visible. The studies in this thesis have shown that people try to avoid long interruptions, which further extends prior research showing how people try to minimise time, and are sensitive to milliseconds (Charman & Howes, 2003; Gray & Fu, 2004). This thesis has shown however that outside of a controlled setting, these milliseconds may not be that visible. Given the thesis finding that people do adapt to time if they are made aware of time costs, there is therefore a need to make it explicit to people how long certain actions take for them to be able to adapt their behaviour.

Differentiate between different types of self-interruptions. This thesis focused on inquiries, a particular type of interruption, and Study 2 and 7 showed that people address this interruption differently than interruptions that are completely unrelated to their current task. Participants in these studies tried to avoid work-irrelevant interruptions, but work-related interruptions were addressed immediately, if they were presumed to be quick and easy. When discussing and making conclusions about people's self-interruption behaviour, it is important to make a distinction between different types of interruptions.

Data retrieval as a part of data entry research. Prior data entry research has primarily focused on improving entry interfaces. For the type of data entry work studied in this thesis, a major part is collecting data in the first place from various locations, and the entry part is actually only a small part of the task. This has implications for future data entry research, as it highlights that more attention needs to be given to the retrieval aspect of a data entry task, which impacts how data is entered. Study 1 showed that office workers tried to batch as much data entry tasks together as possible. A consequence of this was that they had to enter large amounts of data, and did not check their data entries as carefully as when they were checking their own work. Furthermore, when switching between documents, people held items in memory which increases the likelihood of error when returning to the data entry interface. Lastly, Study 2 and 6 highlighted that people often spent long times away from the data entry interface, before they returned. Data entry interfaces should take this into consideration and make it easier to resume a task. If data entry interfaces are intended to be used in situations where information is not readily available, they should be evaluated by requiring participants to first collect data from the environment, to see how usable they are in this context.

6.3 Future work

My work contributes to our knowledge of the effect of time costs and feedback on self-interruptions, and introduces new opportunities that, building on its findings, further investigates how time costs can be utilised to effectively support self-interruption management.

6.3.1 Complementing with other solutions

One area to explore further is how the tool presented in this thesis can be combined and complemented with other approaches. For example, prior work has looked at different approaches to reduce disruptiveness by blocking interruptions and giving reflective information. In Chapter 5 the limitations of these interventions were discussed. The browser notification in my study addressed some of these limitations: The browser notification in my study did not block anything but specifically gave information on the duration of interruptions, as this was found to be an important deciding factor in people's self-interruption behaviour. It would be interesting to explore how different interventions could be combined and complement each other. For example, in Study 7 one effect of the browser notification was that it made people reflect on their actions during interruptions. The browser notification may work as an initial trigger, but could be complemented with a more extensive activity log such as those provided by RescueTime and ManicTime, so people are able to investigate what they were doing during an interruption, and why some interruptions may be longer than others. Prior work has shown that a barrier for people to currently engage with it is that it is not clear what to do with the extensive data and that it lacks context (Collins et al., 2014). Future work would need to investigate whether giving people a trigger may help as an entry point to better explore, understand and use reflective data.

6.3.2 Predicting the type of interruption to advise length of interruption

The longer people interrupt, the more disruptive it can be, and a longer duration can indicate that people are getting distracted. However, as became apparent in Study 7, sometimes people need to make a long interruption. In these cases, the duration may not be informative about whether people are getting distracted or not, and may not be useful as people are limited in their ability to shorten the duration of these interruptions. Future work could investigate people's actions during an interruption, to see whether people are making a necessary and useful interruption, or whether they are getting distracted. To this end, people's interactions during an interruption could be explored as a measure. For example, the browser extension DataSelfie[1] tracks users' interactions when switching to Facebook to try and predict what the user is doing: are they actively looking something up to return to their work, or are they browsing through feeds? This information may get used to advise people on an appropriate time length, and give them timely reminders to return to a task.

6.3.3 Tracking behaviour over time

Lastly, further work is needed to determine how people engage with time feedback over time. This thesis ended with two formative studies exploring the use of a browser notification for about 15 minutes in Study 6, and one working week during Study 7. Though the results are promising, a remaining question is whether people would continue to engage with the notification over time, as many personal informatics tools get abandoned (Lazar, Koehler, Tanenbaum, & Nguyen, 2015).

[1] https://dataselfie.it/

Furthermore, the notification only gave feedback on people's current interruption behaviour. Some participants in Study 7 commented that they would like to see an overview of all interruptions, to see whether the actions they are taking has any considerable effect, and not just on their interruptions at that moment.

6.3.4 Cross-device time feedback

In its current implementation, the notification was evaluated using browser-based work, and intended for managing switching windows on this same device. People increasingly use multiple devices and may switch between devices to look up information on their phone or tablet (Dearman & Pierce, 2008; Murphy et al., 2016; **?**). It would be easy to imagine to extend the notification across devices and take into account cross-device actions, which can show people a notification on a new device they switch to. Furthermore, it may be implemented as a standalone application to provide interruption information when doing a task in a non-browser computer window, such as when writing a document in a word processor.

6.3.5 Extending research findings to other settings

The type of task studied throughout this thesis revolved around a main data entry interface and was characterised by switching frequently and, usually, for short amounts of time. As such, information seeking was considered as a subtask to support the primary task and not analysed as a task in itself. I expect the results of this thesis to generalise to similar types of desktop-based work, where the user does not know beforehand which information is needed and has to make frequent and short switches to many different information sources.

In other domains and types of work, the balance between information seeking and information use may be different: for example, information workers in law offices have to spend a large proportion of their time seeking information (Cangiano & Hollan, 2009), and may spend approximately equal amounts of time in several different computer windows. Bondarenko & Janssen (2005) define these as 'research tasks', and states that in contrast to administrative tasks, people make fewer interruptions during these tasks, often have to re-find previously used information, and spend a longer time in documents. The results of this thesis may not extend to these types of tasks, and in this context it may be of less importance how long and how often people look for information, but rather how they can be supported where to find information.

A research area briefly discussed in Chapter 5 was to make it easier to collect and keep information nearby, reducing the duration of interruptions. As data entry work often requires new information, this idea was not developed further in this thesis, but it may be worthwhile to study for research tasks. It would also be interesting to explore how people address inquiries in these different domains.

6.3.6 Limitations

Limitations which are particular to a study have been discussed in previous chapters, but here I discuss a number of limitations that concern the methodology of the thesis overall.

Manipulating a qualitative phenomenon in a quantitative way

This thesis used a mixed method approach to both understand an underlying mechanism of observed real-world behaviour in a controlled setting, as well as understand the generalisability of experimental findings in a naturalistic setting. Based on observational findings of Study 2, I made the hypothesis that people either address or postpone inquiries because of time costs, which I tested through a series of controlled experiments. However, a limitation of this approach is that by converting a qualitative finding into a quantitative variable and controlling for other possible confounds, there is a disadvantage in it not accurately reflecting the actual phenomenon observed (Driscoll, Appiah-Yeboah, Salib, & Rupert, 2007). While the manipulation of time costs in Chapter 4 is a simplified version of all the time costs involved to retrieve information in data entry work (e.g. physical effort to retrieve information from a location, time spent searching information, time spent opening and loading information sources), the findings in Chapter 4 are supported by existing literature and findings from Chapter 3. Overall, the findings were consistent across qualitative and quantitative studies, and the manipulation was deemed appropriate as a first step to expand our understanding of how time costs contribute to self-interruption behaviour. Other time costs, such as physical effort (Potts, Pastel, & Rosenbaum, 2017), may affect task strategies differently but were beyond the scope of this thesis.

Different participant populations

In addition, participants of different studies were drawn from different populations: the participants in the qualitative Studies 1, 2, and 7 were office workers, while participants in the quantitative Studies 3-6 were from a range of backgrounds. The reason for not exclusively recruiting office workers for the experimental studies as well was that to make significant claims on the effect of time costs and time feedback on people's strategies, it was important to have a sufficient sample size. Office workers have busy work schedules and were hard to recruit, which is also reflected in the small sample size for the qualitative studies. Participation for the experimental studies was therefore opened up to other participants as well. The benefit of the mixed approach was that the research findings of this thesis can be generalised beyond a specific office setting: self-interruptions are common in many computer-based tasks, which makes the research findings not just relevant for office workers. Nevertheless, user expertise and job experience may influence people's strategies (Weir et al., 2007), and future studies could look more into the extent of which expertise contributes to the development of specific strategies.

6.4 Conclusion

Many computer-based tasks require users to interrupt themselves to collect information, which can lead to distractions, and it is challenging to remain focused on the primary task. Prior research has shown that the longer an interruption, the slower people are to resume and the higher likelihood of errors being made. The work in this thesis shows that presumed time effort affects how people address these types of interruptions: they address interruptions immediately if they except them to be short. The implication of this is that people may make many interruptions they think they can return from quickly - even if these end up being far longer than intended.

This thesis has also shown how making it explicit to people how long they actually go away from a task can help in managing self-interruptions: making people aware triggers them to reflect on what they are doing during these interruptions, and makes them reduce the length of interruptions. This is an important finding which has implications for the design of interruption management and productivity tools.

My work makes an important contribution to interruptions literature. Prior work has shown that interruptions become more disruptive the longer they are. By demonstrating in this thesis that time costs influences people's self-interruption behaviour, it highlights the need to make people more aware of when they are making interruptions and for how long. My findings extend our understanding of the factors impacting how people manage self-interruptions.

References

Adamczyk, P. D., & Bailey, B. P. (2004). If Not Now, When?: The Effects of Interruption at Different Moments Within Task Execution. In *Chi 2004* (Vol. 6, pp. 271–278).

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model. *Cognitive Science*, *26*, 39–83. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog2601{_}2

Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General*, *143*(1), 215–226. doi: 10.1037/a0030986

Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2017). Effects of Interruption Length on Procedural Errors. *Journal of Experimental Psychology: Applied*, *23*(2), 216–229. doi: 10.1037/xap0000117

Back, J., Cox, A. L., & Brumby, D. P. (2012). Choosing to Interleave : Human Error and Information Access Cost. In *Proceedings of the 2012 acm annual conference on human factors in computing systems* (pp. 1651–1654).

Baddeley, A. D., & Hitch, G. (1974). *Working memory* (Vol. 8). Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0079742108604521 doi: 10.1016/S0079-7421(08)60452-1

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2001). The Effects of Interruptions on Task Performance, Annoyance, and Anxiety in the User Interface. In M. Hirose (Ed.), *Interact* (pp. 593–601). Amsterdam: iOS Press. Retrieved from https://www.cse.unr.edu/{~}sushil/class/ps/papers/EffectInterruptions-interact-2001.pdf

BBC. (2018). *Junior doctors' job offers withdrawn after blunder.* Retrieved from http://www.bbc.co.uk/news/uk-44020235

Bi, X., & Balakrishnan, R. (2009). Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. In *Proceedings of the sigchi conference on human factors in computing systems (chi '09)* (pp. 1005–1014). Retrieved from http://dl.acm.org/citation.cfm?doid=1518701.1518855{%}5Cnpapers2://publication/doi/10.1145/1518701.1518855 doi: 10.1145/1518701.1518855

Blandford, A., Furniss, D., & Makri, S. (2016). Qualitative HCI Research: Going Behind the Scenes. *Synthesis Lectures on Human-Centered Informatics*, *9*(1), 1–115.

Bondarenko, O., & Janssen, R. (2005). Documents at hand: Learning from paper to improve digital technologies. In *Proceedings of the sigchi conference on human factors in computing systems (chi '05)* (pp. 121–130). Retrieved from http://portal.acm.org/citation.cfm?id=1054990 doi: 10.1.1.137.7259

Borghouts, J., Brumby, D. P., & Cox, A. L. (n.d.). Task Interruptions in the Office: Digital Self-Interruptions are Harder to Manage than Physical Ones. *International Journal of Human-Computer Studies*.

Borghouts, J., Brumby, D. P., & Cox, A. L. (2017). Batching, Error Checking and Data Collecting: Understanding Data Entry in a Financial Office. In *Proceedings of 15th european conference on computer-supported cooperative work.* Sheffield, UK. doi: 10.18420/ecscw2017-4

Borghouts, J., Brumby, D. P., & Cox, A. L. (2018). Looking Up Information in Email: Feedback on Visit Durations Discourages Distractions. In *Chi'18 extended abstracts on human factors in computing systems.* Montreal, Canada. doi: 10.1145/3170427.3188607

Borghouts, J., Soboczenski, F., Cairns, P., & Brumby, D. P. (2015). Visualizing Magnitude : Graphical Number Representations Help Users Detect Large Number Entry Errors. *59th Annual Meeting of the Human Factors and Ergonomics Society*, 591–595. doi: 10.1177/1541931215591130

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. Retrieved from http://eprints.uwe.ac.uk/11735/1/thematic{_}analysis{_}revised{_}-{_}final.doc doi: 10.1191/1478088706qp063oa

Cairns, P., & Cox, A. L. (2008). *Research Methods for Human-Computer Interaction* (P. Cairns & A. L. Cox, Eds.). New York, NY, USA: Cambridge University Press.

Cangiano, G. R., & Hollan, J. D. (2009). Capturing and restoring the context of everyday work: A case study at a law office. In M. Kurosu (Ed.), *Human centered design* (pp. 945–954). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-02806-9_108

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

Carrier, L. M., Rosen, L. D., Cheever, N. A., & Lim, A. F. (2015). Causes, effects, and practicalities of everyday multitasking. *Developmental Review*, *35*, 64–78. Retrieved from http://dx.doi.org/10.1016/j.dr.2014.12.005 doi: 10.1016/j.dr.2014.12.005

Charman, S. C., & Howes, A. (2003). The adaptive user: an investigation into the cognitive and task constraints on the generation of new methods. *Journal of Experimental Psychology: Applied*, *9*(4), 236–48. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14664675 doi: 10.1037/1076-898X.9.4.236

Collins, E. I. M., Cox, A. L., Bird, J., & Cornish-Tresstail, C. (2014). Barriers to engagement with a personal informatics productivity tool. In *Proceedings of the 26th australian computer-human interaction conference on designing futures the future of design - ozchi '14* (pp. 370–379). Retrieved from http://dl.acm.org/citation.cfm?doid=2686612.2686668 doi: 10.1145/2686612.2686668

Constine, J. (2018a). *Facebook prototypes tool to show how many minutes you spend on it.* Retrieved from https://techcrunch.com/2018/06/22/your-time-on-facebook/

Constine, J. (2018b). *Instagram CEO confirms upcoming "time spent" Usage Insights.* Retrieved 2018-05-31, from https://techcrunch.com/2018/05/15/instagram-usage-insights/

Cutrell, E., Czerwinski, M., & Horvitz, E. (2001). Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. In *Proceedings of interact 2001* (pp. 263–269). New York, NY, USA: Springer.

Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A Diary Study of Task Switching and Interruptions. In *Proceedings of the sigchi conference on human factors in computing systems (chi '04)* (pp. 175–182). doi: 10.1145/985692.985715

Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., & Starkweather, G. (2003). Toward Characterizing the Productivity Benefits of Very Large Displays. In *Interact'03* (pp. 9–16). Retrieved from http://research.microsoft.com/en-us/um/redmond/groups/cue/publications/interact2003-productivitylargedisplays.pdf

Dearman, D., & Pierce, J. S. (2008). It's on my other computer!: computing with multiple devices. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 767–776). doi: 10.1145/1357054.1357177

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011, jan). Fortune favors the bold (and the Italicized): effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21040910 doi: 10.1016/j.cognition.2010.09.012

Driscoll, D. L., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging Qualitative and Quantitative Data in Mixed Methods Research: How To and Why Not. *Ecological and Environmental Anthropology*, *3*(1), 19–28. Retrieved from http://digitalcommons.unl.edu/icwdmeea

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. C. (2003). Stuff I've Seen: A System for Personal Information Retrieval and Re-Use. In *Sigir '03.* Toronto, Canada. Retrieved from https://www.microsoft.com/en-us/research/wp-content/uploads/2003/01/siscore-sigir2003-final.pdf

Evans, A. C., & Wobbrock, J. O. (2012). Taming Wild Behavior : The Input Observer for Obtaining Text Entry and Mouse Pointing Measures from Everyday Computer Use. In *Chi* (pp. 1947–1956). Austin, Texas, USA.

Farmer, G. D., Janssen, C. P., Nguyen, A. T., & Brumby, D. P. (2017). Dividing Attention Between Tasks: Testing Whether Explicit Payoff Functions Elicit Optimal Dual-Task Performance. *Cognitive Science*, 1–30. Retrieved from http://onlinelibrary.wiley.com/store/10.1111/cogs.12513/asset/cogs12513.pdf;jsessionid=ECF8C2EAC059D343C30C454B13126DF8.f01t01?v=1{&}t=jcdeo5xz{&}s=5c3623e8e560f2d7bb00babe08289c61f4a84bdc doi: 10.1111/cogs.12513

Furniss, D., & Blandford, A. (2006). Understanding Emergency Medical Dispatch in terms of Distributed Cognition: a case study. *Ergonomics*, *49*(12 & 13), 1174 – 1203.

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *Quarterly journal of experimental psychology (2006)*, *68*(April), 971–92. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4448673{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1080/17470218.2014.972963

Gonzalez, V. M., & Mark, G. (2004). "Constant, Constant, Multi-tasking Craziness": Managing Multiple Working Spheres. In *Proceedings of the sigchi conference on human factors in computing systems (chi '04)* (pp. 113–120). Vienna, Austria. Retrieved from http://delivery.acm.org/10.1145/990000/985707/p113-gonzalez.pdf?key1=985707{&}key2=9709385111{&}coll=GUIDE{&}dl=GUIDE{&}CFID=44938518{&}CFTOKEN=14011566

Gould, S. J., Brumby, D. P., & Cox, A. L. (2013, sep). What does it mean for an interruption to be relevant? An investigation of relevance as a memory effect. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 149–153. Retrieved from http://pro.sagepub.com/lookup/doi/10.1177/1541931213571034 doi: 10.1177/1541931213571034

Gould, S. J., Cox, A. L., & Brumby, D. P. (2016). Diminished Control in Crowdsourcing: An Investigation of Crowdworker Multitasking Behavior. *ACM Transactions on Computer-Human Interaction*, *23*(3), 1–27. Retrieved from message:{%}3C3DF7070D-92E4-400F-B288-C71BA25CC839@ucl.ac.uk{%}3Epapers3://publication/uuid/AC17DF2C-0715-49A0-A788-126336835C8E doi: 10.1145/2928269

Gould, S. J., Cox, A. L., Brumby, D. P., & Wickersham, A. (2016). Now Check Your Input: Brief Task Lockouts Encourage Checking, Longer Lockouts Encourage Task Switching. In *Proceedings of the sigchi conference on human factors in computing systems (chi '16)* (pp. 3311–3323). Retrieved from http://doi.acm.org/10.1145/2858036.2858067 doi: 10.1145/2858036.2858067

Gould, S. J., Cox, A. L., Brumby, D. P., & Wiseman, S. E. M. (2016). Shortlinks and tiny keyboards: a systematic exploration of design trade-offs in link shortening services. *International Journal of Human-Computer Studies*. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1071581916300854 doi: 10.1016/j.ijhcs.2016.07.009

Gray, W. D., & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*(3), 359–382. doi: 10.1016/j.cogsci.2003.12.001

Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006, jul). The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological review*, *113*(3), 461–82. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/16802878` doi: 10.1037/0033-295X.113.3.461

Grudin, J. (2001). Partitioning DigitalWorlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Proceedings of the sigchi conference on human factors in computing systems (chi '01)* (pp. 458–465). Retrieved from `http://dl.acm.org/citation.cfm?id=365312` doi: 10.1145/365024.365312

Grundgeiger, T., Sanderson, P. M., Macdougall, H. G., & Venkatesh, B. (2010). Interruption Management in the Intensive Care Unit: Predicting Resumption Times and Assessing Distributed Support. *Journal of Experimental Psychology: Applied*, *16*(4), 317–334. Retrieved from `https://interruptions.net/literature/Grundgeiger-JEPA10.pdf` doi: 10.1037/a0021912

Hanrahan, B. V., & Pérez-Qu, M. A. (2015). Lost in Email: Pulling Users Down a Path of Interaction. In *Chi'15* (pp. 3981–3984). Retrieved from `http://delivery.acm.org/10.1145/2710000/2702351/p3981-hanrahan.pdf?ip=128.16.5.5{&}id=2702351{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}CFID=1018446498{&}CFTOKEN=44534115{&}{_}{_}acm{_}{_}=1513782721{_}94007b7bb008` doi: 10.1145/2702123.2702351

Healy, A. F., Kole, J. a., Buck-Gengle, C. J., & Bourne, L. E. (2004, sep). Effects of prolonged work on data entry speed and accuracy. *Journal of Experimental Psychology: Applied*, *10*(3), 188–99. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/15462620` doi: 10.1037/1076-898X.10.3.188

Hollan, J. D., Hutchins, E. L., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.*, *7*(2), 174–196. doi: http://doi.acm.org/10.1145/353485.353487

Holtzblatt, K., & Beyer, H. (2014). *Contextual Design: Evolved* (J. M. Carroll, Ed.). Morgan & Claypool Publishers.

Hutchins, E. L. (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science*, *19*, 265–288. Retrieved from `http://doi.wiley.com/10.1207/s15516709cog1903{_}1` doi: 10.1207/s15516709cog1903_1

Iqbal, S. T., & Bailey, B. P. (2005). Investigating the Effectiveness of Mental Workload as a Predictor of Opportune Moments for Interruption. In *Chi '05* (pp. 1489–1492). Portland, OR, USA. Retrieved from https://interruptions.net/literature/Iqbal-CHI05-p1489-iqbal.pdf

Iqbal, S. T., & Bailey, B. P. (2008). Effects of Intelligent Notification Management on Users and Their Tasks. In *Chi'08* (pp. 93–102). Retrieved from http://delivery.acm.org/10.1145/1360000/1357070/p93-iqbal.pdf?ip=128.16.5.5{&}id=1357070{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}{_}{_}acm{_}{_}=1527353089{_}3b962ded143311b06fe699aa288f6ffa

Iqbal, S. T., & Horvitz, E. (2007). Disruption and recovery of computing tasks: field study, analysis, and directions. In *Chi '07* (pp. 677—-686). San Jose, California, USA. Retrieved from http://dl.acm.org/citation.cfm?id=1240730 doi: 10.1145/1240624.1240730

Janssen, C. P., & Gray, W. D. (2012, mar). When, what, and how much to reward in reinforcement learning-based models of cognition. *Cognitive science*, *36*(2), 333–58. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22257174 doi: 10.1111/j.1551-6709.2011.01222.x

Jin, J., & Dabbish, L. A. (2009). Self-Interruption on the Computer: A Typology of Discretionary Task Interleaving. In *Chi'09* (pp. 1799–1808).

Katidioti, I., & Taatgen, N. A. (2013, sep). Choice in Multitasking: How Delays in the Primary Task Turn a Rational Into an Irrational Multitasker. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(4), 728–736. Retrieved from http://hfs.sagepub.com/cgi/doi/10.1177/0018720813504216 doi: 10.1177/0018720813504216

Kim, J., Cho, K. C., & Lee, K. U. (2017). Technology Supported Behavior Restriction for Mitigating Self-Interruptions in Multi-device Environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol*, *1*(21). Retrieved from https://doi.org/10.1145/3130932 doi: 10.1145/3130932

Lazar, A., Koehler, C., Tanenbaum, J., & Nguyen, D. H. (2015). Why We Use and Abandon Smart Devices. In *Ubicomp'15* (pp. 635–646). Osaka, Japan. Retrieved from http://delivery.acm.org/10.1145/2810000/2804288/p635-lazar.pdf?ip=128.16.5.5{&}id=2804288{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5

.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}{_}{_}acm{_}{_}=1526546789{_}9fd8a8e734f6893e47ae49896e708611 doi: 10.1145/2750858.2804288

Lewis, C. (1986). Understanding what's happening in system interactions. In D. A. Norman & S. W. Draper (Eds.), *User centered system design* (pp. 169–185). Hillsdale, NJ: Erlbaum.

Li, S. Y. W., Cox, A. L., Or, C., & Blandford, A. (2016). Effects of monetary reward and punishment on information checking behaviour. *Applied Ergonomics*, *53*, 258–266. doi: 10.1016/j.apergo.2015.10.012

Lyngs, U. (2018). A Cognitive Design Space for Supporting Self-Regulation of ICT Use. In *Chi'18 extended abstracts.* Montreal, Canada. Retrieved from http://delivery.acm.org/10.1145/3190000/3180296/SRC14.pdf?ip=128.16.5.5{&}id=3180296{&}acc=OPEN{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.6D218144511F3437{&}{_}{_}acm{_}{_}=1526464496{_}3d92b7e9ea58478894720c0697523b2a

Lynley, M. (2018). *Google rolls out app time management controls.* Retrieved 2018-05-31, from https://techcrunch.com/2018/05/08/android-rolls-out-a-suite-of-time-management-controls-to-promote-more-healthy-app-usage/

Maior, H. A., Wilson, M. L., & Sharples, S. (2018). Workload Alerts—Using Physiological Measures of Mental Workload to Provide Feedback During Tasks. *ACM Transactions on Computer-Human Interaction ACM Reference ACM Trans. Comput.-Hum. Interact*, *25*(2), 1–30. Retrieved from http://eprints.nottingham.ac.uk/49196/8/a9-maior.pdfhttp://eprints.nottingham.ac.uk/end{_}user{_}agreement.pdfhttps://doi.org/10.1145/3173380

Makri, S. (2008). *A study of lawyers ' information behaviour leading to the* (Unpublished doctoral dissertation). University College London.

Mark, G., Czerwinski, M., & Iqbal, S. T. (2018). Effects of Individual Differences in Blocking Workplace Distractions. In *Chi'18.* Montreal, Canada. Retrieved from http://delivery.acm.org/10.1145/3180000/3173666/paper92.pdf?ip=192.252.136.182{&}id=3173666{&}acc=OPEN{&}key=4D4702B0C3E38B35.4D4702B0C3E38B35.4D4702B0C3E38B35.6D218144511F3437{&}{_}{_}acm{_}{_}=1524844800{_}b6715c698fdec44cb2d7422d2687f12d{#}URLTOKEN{%}23 doi: 10.1145/3173574.3173666

Mark, G., Gonzalez, V. M., & Harris, J. (2005). No task left behind? In *Proceedings of the sigchi conference on human factors in computing systems (chi '05)* (pp. 321–330). Portland, Oregon, USA. Retrieved from http://dl.acm.org/citation.cfm?id=1054972.1055017 doi: 10.1145/1054972.1055017

Mark, G., Iqbal, S. T., Czerwinski, M., & Johns, P. (2014). Capturing the Mood : Facebook and Face-to-Face Encounters in the Workplace. In *17th acm conference on computer supported cooperative work & social computing* (pp. 1082–1094). doi: 10.1145/2531602.2531673

Mark, G., Iqbal, S. T., Czerwinski, M., & Johns, P. (2015). Focused, Aroused, but so Distractible: A Temporal Perspective on Multitasking and Communications. In *Cscw '15 proceedings of the 18th acm conference on computer supported cooperative work & social computing* (pp. 903–916). Retrieved from http://delivery.acm.org/10.1145/2680000/2675221/p903-mark.pdf?ip=128.16.5.5{&}id=2675221{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}{_}{_}acm{_}{_}=1521461269{_}0ffddd2ed42df73f602efe73b93be18b doi: 10.1145/2675133.2675221

Mark, G., Iqbal, S. T., Czerwinski, M., Johns, P., & Sano, A. (2016a). Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Chi'16* (pp. 1717–1728).

Mark, G., Iqbal, S. T., Czerwinski, M., Johns, P., & Sano, A. (2016b). Neurotics Can't Focus: An in situ Study of Online Multitasking in the Workplace. In *Chi '16* (pp. 1739–1744). San Jose, CA, USA. Retrieved from http://delivery.acm.org/10.1145/2860000/2858202/p1739-mark.pdf?ip=128.16.5.5{&}id=2858202{&}acc=CHORUS{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.6D218144511F3437{&}{_}{_}acm{_}{_}=1520681227{_}d66922137944034c993df83641b5b0a7 doi: 10.1145/2858036.2858202

Mark, G., Voida, S., & Cardello, A. V. (2012). "A Pace Not Dictated by Electrons": An Empirical Study of Work Without Email. In *Chi'12* (pp. 555–564). Austin, Texas, USA. Retrieved from http://delivery.acm.org/10.1145/2210000/2207754/p555-mark.pdf?ip=128.16.5.5{&}id=2207754{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}{_}{_}acm{_}{_}=1523544578{_}bf4b7766d0172e85f713049c87328e35

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two. *Psychological Review*, *63*, 81–97.

Monk, C. A., Trafton, J. G., & Boehm-Davis, D. A. (2008). The effect of interruption duration and demand on resuming suspended goals. *Journal of Experimental Psychology: Applied*, *14*(4), 299–313. Retrieved from http://doi.apa.org/getdoi.cfm?doi=10.1037/a0014402 doi: 10.1037/a0014402

Morgan, P. L., & Patrick, J. (2012). Paying the price works: Increasing goal-state access cost improves problem solving and mitigates the effect of interruption. *The Quarterly Journal of Experimental Psychology*(November 2012), 1–19. doi: 10.1080/17470218.2012.702117

Morgan, P. L., Patrick, J., Waldron, S. M., King, S. L., & Patrick, T. (2009, dec). Improving memory after interruption: exploiting soft constraints and manipulating information access cost. *Journal of Experimental Psychology: Applied*, *15*(4), 291–306. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20025416 doi: 10.1037/a0018008

Morgan, P. L., Waldron, S. M., King, S. L., & Patrick, J. (2007). Harder to Access, Better Performance? The Effects of Information Access Cost on Strategy and Performance. In M. Smith & G. Salvendy (Eds.), *Human interface, part i, hcii 2007, lncs 4557* (pp. 115–125). Berlin: Springer-Verlag.

Murphy, H. C., Chen, M.-M., & Cossutta, M. (2016). An investigation of multiple devices and information sources used in the hotel booking process. *Tourism Management*, *52*, 44–51. doi: 10.1016/j.tourman.2015.06.004

O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive psychology*, *35*(1), 34–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9520317 doi: 10.1006/cogp.1997.0676

Oladimeji, P., Thimbleby, H., & Cox, A. L. (2011). Number entry interfaces and their effects on error detection. In *Ifip international federation for information processing* (pp. 178–185). Springer-Verlag. Retrieved from http://discovery.ucl.ac.uk/1308862/

Oladimeji, P., Thimbleby, H., & Cox, A. L. (2013). A Performance Review of Number Entry Interfaces. In *Proceedings of 14th ifip tc13 conference on human-computer interaction (interact 2013), part i* (Vol. 8117, pp. 365–382).

Olsen, K. A. (2008). The $100,000 keying error. *Computer*, *41*. doi: 10.1109/MC.2008.135

Onwuegbuzie, A. J., & Dickinson, W. B. (2008). Mixed Methods Analysis and Information Visualization: Graphical Display for Effective Communication of Research Results. *The Qualitative Report*, *13*(2), 204–225. Retrieved from http://wtgrantmixedmethods.com/sites/default/files/literature/Onwuegbuzie{_}Dickinson{_}2008{_}MMAnalysisandVisualization{_}0.pdf

Patrick, J., Morgan, P. L., Tiley, L., Smy, V., & Seeby, H. (2014). Designing the Interface to Encourage More Cognitive Processing. In *Hci* (Vol. 23, pp. 255–264).

Payne, S. J., & Howes, A. (2013). *Adaptive Interaction: A Utility Maximization Approach to Understanding Human Interaction with Technology* (1st ed.). Morgan&Claypool Publishers.

Potts, C. A., Pastel, S., & Rosenbaum, D. A. (2017, nov). How are cognitive and physical difficulty compared? *Attention, Perception, & Psychophysics*, *80*(2), 500–511. Retrieved from http://link.springer.com/10.3758/s13414-017-1434-2 doi: 10.3758/s13414-017-1434-2

Randall, D., & Rouncefield, M. (2014). Ethnography. In M. Soegaard & R. F. Dam (Eds.), *The encyclopedia of human-computer interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from https://www.interaction-design.org/encyclopedia/ethnography.html

Reason, J. (1990). *Human Error*. Cambridge University Press.

Reddy, M., & Dourish, P. (2002). A finger on the pulse: temporal rhythms and information seeking in medical work. In *Proceedings of the acm conference on computer supported collaborative work (cscw '02)* (pp. 344–353). Retrieved from http://portal.acm.org/citation.cfm?id=587078.587126 doi: 10.1145/587078.587126

Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., & Tan, D. (2005). The large-display user experience. *IEEE Computer Graphics and Applications*, *25*(4), 44–51. doi: 10.1109/MCG.2005.88

Rule, A., & Youngstrom, K. (2013). That Reminds Me: Identifying Elements of Screen Recordings that Cue Contextual Memory. In *Cogs-230'13.* San Diego, CA, USA.

Rzeszotarski, J. M., Chi, E., Paritosh, P., & Dai, P. (2013). *Inserting Micro-Breaks into Crowdsourcing Workflows* (Tech. Rep.).

Salthouse, T. a. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological bulletin*, *99*(3), 303–319. doi: 10.1037/0033-2909.99.3.303

Salvucci, D. D., & Bogunovich, P. (2010). Multitasking and Monotasking : The Effects of Mental Workload on Deferred Task Interruptions. In *Chi 2010* (pp. 85–88).

Sellberg, C., & Susi, T. (2014). Technostress in the office: A distributed cognition perspective on human-technology interaction. *Cognition, Technology and Work*, *16*, 187–201. doi: 10.1007/s10111-013-0256-9

Sellen, A. J., & Harper, R. H. (2003). *The Myth of the Paperless Office*. Cambridge, MA, USA: MIT Press.

Skatova, A., Bedwell, B., Shipp, V., Huang, Y., Young, A., Rodden, T., & Bertenshaw, E. (2016). The Role of ICT in Office Work Breaks. In *Chi'16* (pp. 3049–3060). San Jose, CA, USA. Retrieved from http://delivery.acm.org/10.1145/2860000/2858443/p3049-skatova.pdf?ip=128.16.5.5{&}id=2858443{&}acc=ACTIVESERVICE{&}key=BF07A2EE685417C5.D93309013A15C57B.4D4702B0C3E38B35.4D4702B0C3E38B35{&}{_}{_}acm{_}{_}=1520681051{_}49163cbfb34fb3f374dcaef12aa3a62b doi: 10.1145/2858036.2858443

Smith, G., Baudisch, P., Robertson, G., Czerwinski, M., Meyers, B., Robbins, D., & Andrews, D. (2003). GroupBar: The TaskBar Evolved. In *Ozchi 2003 conference for the computer-human interaction special interest group of the human factors society of australia.* Retrieved from http://www.patrickbaudisch.com/publications/2003-Smith-OZCHI03-GroupBar.pdf

Soboczenski, F., Cairns, P., & Cox, A. L. (2013). Increasing accuracy by decreasing presentation quality in transcription tasks. In *Proceedings of 14th ifip tc13 conference on human-computer interaction (interact 2013), part ii* (Vol. 8118, pp. 380–394).

Sohn, T., Li, K. A., Griswold, W. G., & Hollan, J. D. (2008). A diary study of mobile information needs. In *Chi '08* (pp. 433–442). Retrieved from http://portal.acm.org/citation.cfm?doid=1357054.1357125 doi: 10.1145/1357054.1357125

Su, N. M., Brdiczka, O., & Begole, B. (2013). The Routineness of Routines: Measuring Rhythms of Media Interaction. *Human–Computer Interaction*, *28*(4), 287–334. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/07370024.2012.697026 doi: 10.1080/07370024.2012.697026

Thimbleby, H. (2011). Interactive Numbers : a Grand Challenge. In Blashki, K. (Ed.), *Proceedings of the iadis international conference on interfaces and human computer interaction* (pp. 1–9).

Vertanen, K., Memmi, H., Emge, J., Reyal, S., & Kristensson, P. O. (2015). VelociTap : Investigating Fast Mobile Text Entry using Sentence-Based Decoding of Touchscreen Keyboard Input. In *Chi 2015* (pp. 659–668).

Waldron, S. M., Patrick, J., & Duggan, G. B. (2011). The influence of goal-state access cost on planning during problem solving. *Quarterly Journal of Experimental Psychology*, *64*(3), 485–503.

Waldron, S. M., Patrick, J., Morgan, P. L., & King, S. L. (2007). Influencing cognitive strategy by manipulating information access. *Computer Journal*, *50*, 694–702. doi: 10.1093/comjnl/bxm064

Weir, C., Nebeker, J., Hicken, B. L., Campo, R., Drews, F., & Lebar, B. (2007). A Cognitive Task Analysis of Information Management Strategies in a Computerized Provider Order Entry Environment. *Journal of the American Medical Informatics Association*, *14*(1), 65–75. doi: 10.1197/jamia.M2231.Introduction

Westbrook, J. I., Woods, A., Rob, M. I., Dunsmuir, W. T. M., & Day, R. O. (2010). Association of Interruptions With an Increased Risk and Severity of Medication Administration Errors. *Archives of Internal Medicine*, *170*(8), 683–690. Retrieved from `http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinternmed.2010.65` doi: 10.1001/archinternmed.2010.65

Whittaker, S., Hollis, V., & Guydish, A. (2016). 'Don't Waste My Time': Use of Time Information Improves Focus. In *Chi'16* (pp. 1729–1738). San Jose, CA, USA.

Wiseman, S. E. M. (n.d.). *Designing for Numerical Transcription Typing: Frequent Numbers Matter* (Doctoral thesis). University College London.

Wiseman, S. E. M., Borghouts, J., Grgic, D., Brumby, D. P., & Cox, A. L. (2015). The Effect Of Interface Type On Visual Error Checking Behavior. In *Human factors and ergonomics society.* Los Angeles, CA.

Wiseman, S. E. M., Cairns, P., & Cox, A. L. (2011). A Taxonomy of Number Entry Error. In *The 25th bcs conference on human-computer interaction* (pp. 187–196).

Wiseman, S. E. M., Cox, A. L., Brumby, D. P., Gould, S. J., & Carroll, S. O. (2013). Using Checksums to Detect Number Entry Error. In *Proceedings of the 2013 conference on human factors in computing systems (chi-2013)* (pp. 2403–2406).

APPENDIX A

STUDY 1: DIAGRAMS OF THEMES

The results of Study 1 were analysed using an inductive approach of thematic analysis: there was no pre-existing coding scheme. From this analysis, 51 codes were established, which were grouped into eight themes. Each theme was visualised in a diagram, which shows the theme's main codes and relationships between codes, as well as quotes in dotted squares to exemplify what type of quotes were grouped under this code. The numbers in parentheses indicate the number of quotes, and the number of interviewees who mentioned it. The description of each theme is accompanied with notes and quotes taken from the transcripts to further illustrate when this theme was mentioned. These serve as examples and are not all the instances of a theme. To differentiate notes from verbatim quotes, the quotes are in italics and double quotation marks. Words put in brackets are added by the researcher to make the quote more understandable for the reader, for instance if the interviewee is talking about 'it' or 'them'. The diagrams are ordered according to the number of quotations associated with a theme, with the theme with the most quotations listed first. The only exception is the 'Other' theme which is described last.

A.1 Diagrams of themes

A.1.1 Task

Quotes were grouped under this theme if participants described things that were particular to their task, for instance how they structured their task, whether they switched tasks, and how long they took to complete tasks.

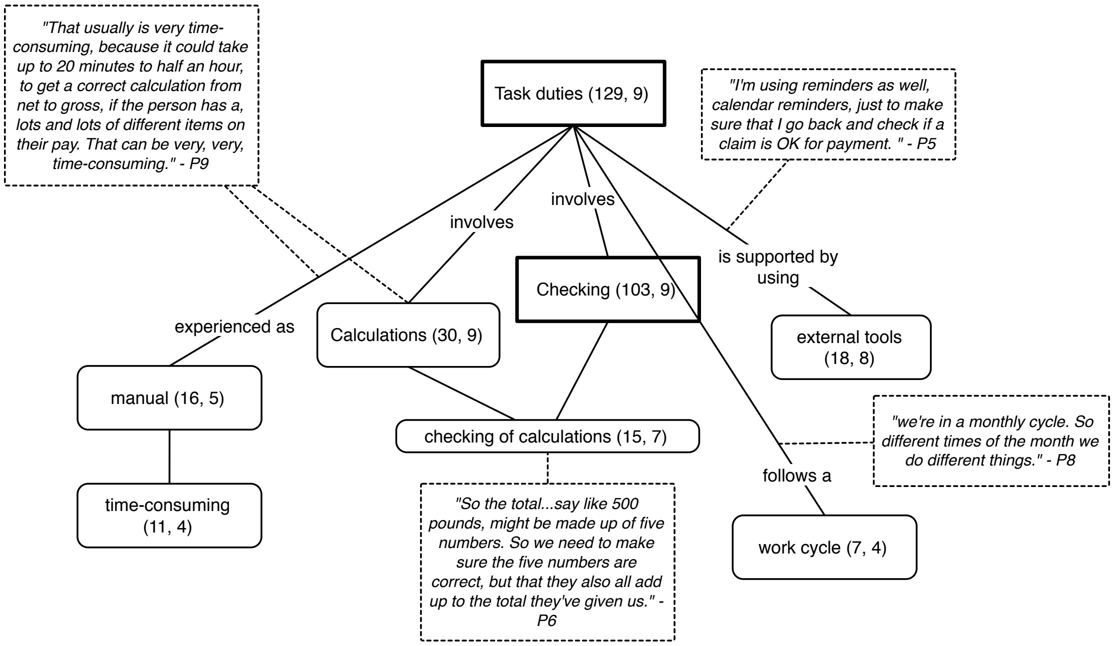


Figure A.1: Diagram showing the theme Task. The numbers in parentheses indicate the number of quotes and the number of participants who mentioned it, respectively.

A.2 Checking

Quotes were grouped under this theme if participants talked about checking data input as part of their job.

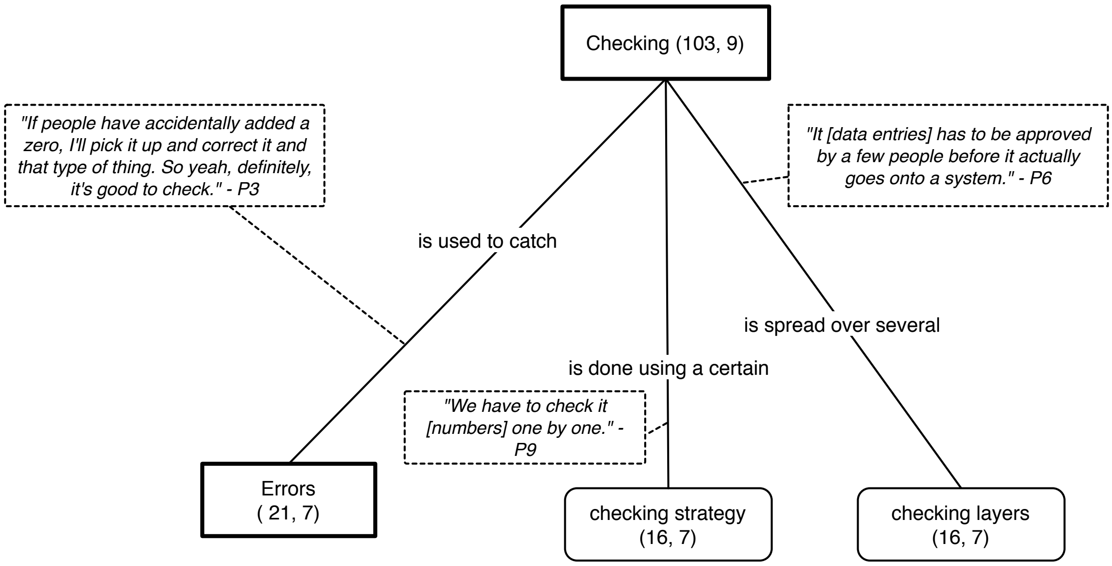


Figure A.2: Diagram showing the theme Checking.

Participant	Quote
P3	*"we try and pick it [errors] up and then obviously there's all the different stages that pick it up as you go along."*
P9	*"the departments actually sometimes treat us as a checking system [laughs], but they shouldn't really, the schools. Because we're here just to make sure that people get paid correctly. But even though we are like a second check, we feel sometimes that we are the first checkpoint."*
P7	*"All this piece of work, when we input in the system, will be actually checked by another person... my manager will print it out, and then check... other colleagues will double-check it for you as well, the calculations."*
P8	*"one of these errors could be things that are missed during the checking."*

Table A.1: Verbatim quotes taken from the interview transcripts that were about checking.

Participant	Quote/note
P1	first puts in all the details, then when done checks everything against the source.
P7	when entering numbers from paper to computer, mostly looked at paper form and the number pad; only looked at screen after finishing entering all the numbers from the form to check.
P5	*"We would go by the receipt, so we would try to make sure that the receipts are in order."*

Table A.2: Checking own input when entering data.

Participant	Quote
P5	*"The numbers on the expense form will be checked individually. So the total will obviously be, say like 500 pounds, might be made up of five numbers. So we need to make sure the five numbers are correct, but that they also all add up to the total they've given us."*
P6	*"The numbers on the expense form will be checked individually."*
P9	*"We have to check it one by one."*

Table A.3: Checking other people's input.

A.3 System

Quotes were grouped under this theme if participants talked about the computer system they were using to input data.

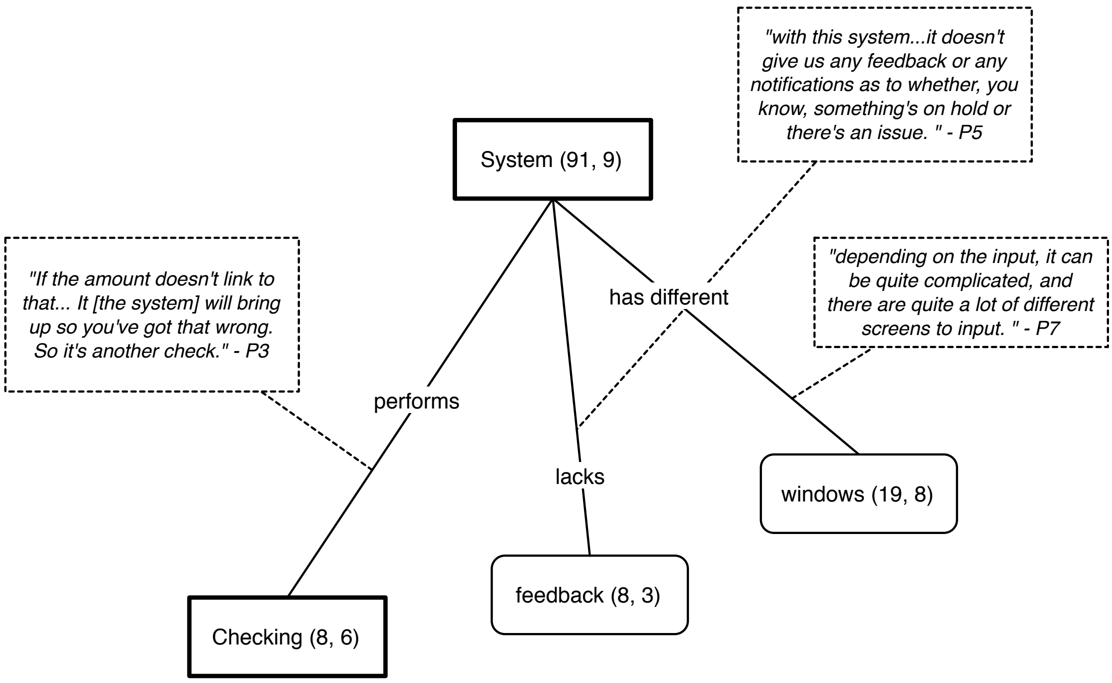


Figure A.3: Diagram showing the theme System.

A.4 Environment

Quotes were grouped under this theme if participants described their environment, for instance if they talked about their physical work setting, and the work culture of their organisation.

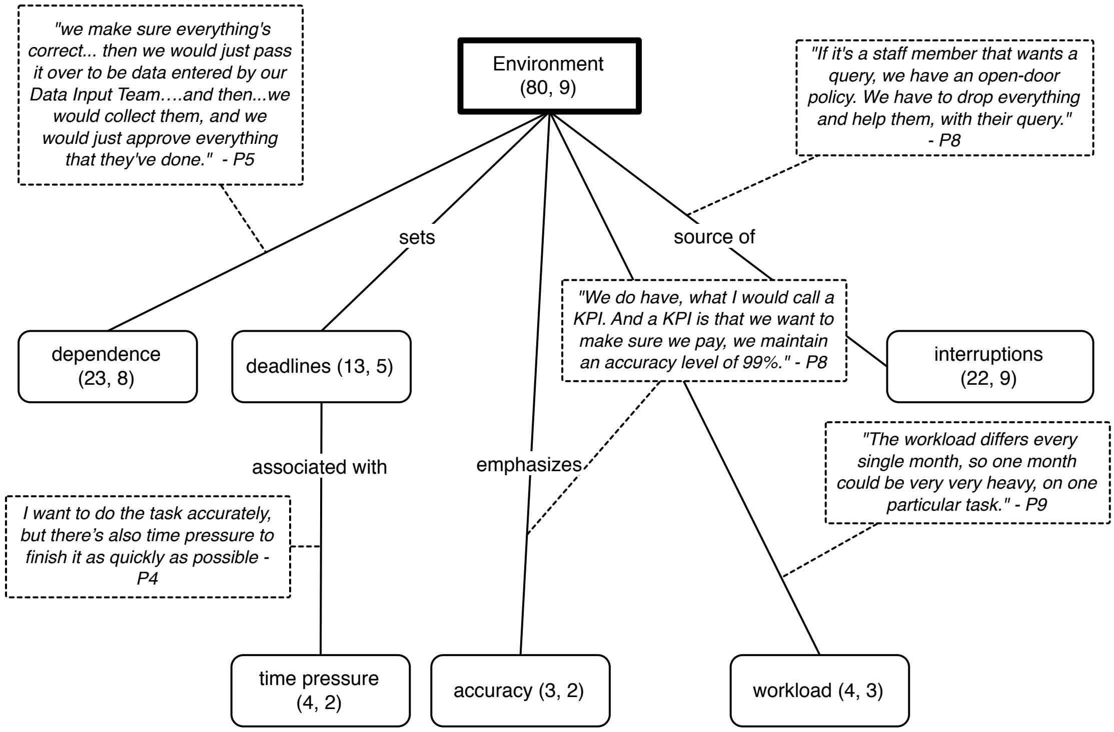


Figure A.4: Diagram showing the theme Environment.

A.5 Data

Quotes were grouped under this theme if participants described the data they were dealing with, for instance the type and length of data items, and from which source they copied data.

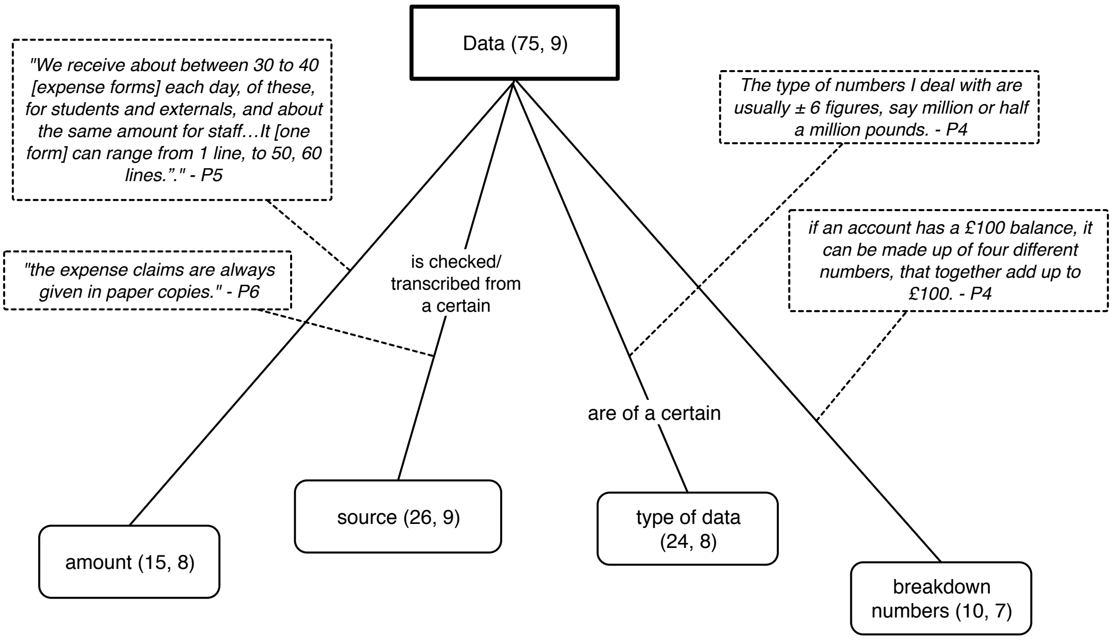


Figure A.5: Diagram showing the theme Data.

A.6 Errors

Quotes were grouped under this theme if participants described situations where errors were made: who made them, why were they made, what were the consequences.

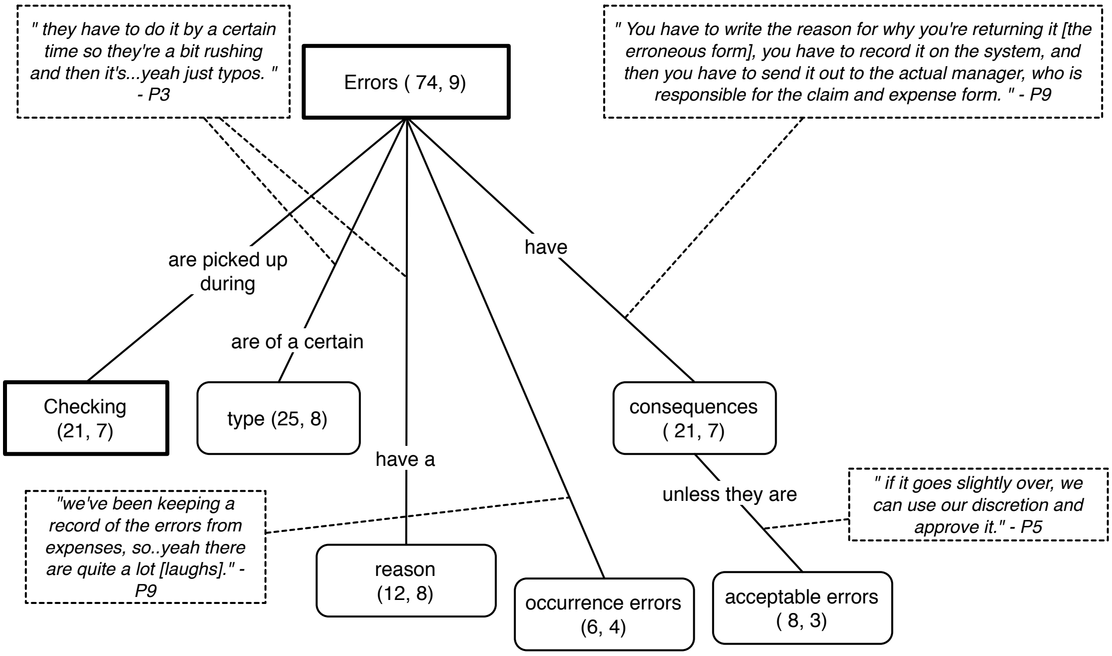


Figure A.6: Diagram showing the theme Errors.

Participant	Quote/note
P6	*"it's quite common that we have to return an expense or payment back to someone. It happens quite often, yeah."*
P4	Yes all the time, lots of typos.

Table A.4: People mentioned errors occur quite frequently.

Participant	Quote
P3	*"sometimes it's because people have done typos, done too many zeroes, or left out a zero."*
P5	*"the expense breakdown doesn't match what (...) whatever they put as the grand total."*

Table A.5: The type of errors.

Participant	Quote
P9	*"Because the departments actually sometimes treat us as a checking system [laughs], but they shouldn't really."*
P7	*"Yeah, human laziness or something [laughs]."*
P8	*"sometimes, you know, through human error, you know, things don't get paid properly."*

Table A.6: The reasons for errors.

Participant	Quote/note
P5	*"generally we tend, we try not to send claims back to departments because they might get lost in the post, and it's an inconvenience as well. So we try to... resolve it ourselves.."*
P4	We allow a certain amount of tolerance; if it turns out the thing you bought has actually decreased value and is now £40, we will allow to return £50
P7	*"we normally e-mail the budget holder to say... what you authorised is actually different. But for this kind of thing, it's only 10 pounds...we normally just process this without contacting them."*

Table A.7: Acceptable errors.

A.7 Strategy

Quotes were grouped under this theme if participants described the strategies they used to carry out their task.

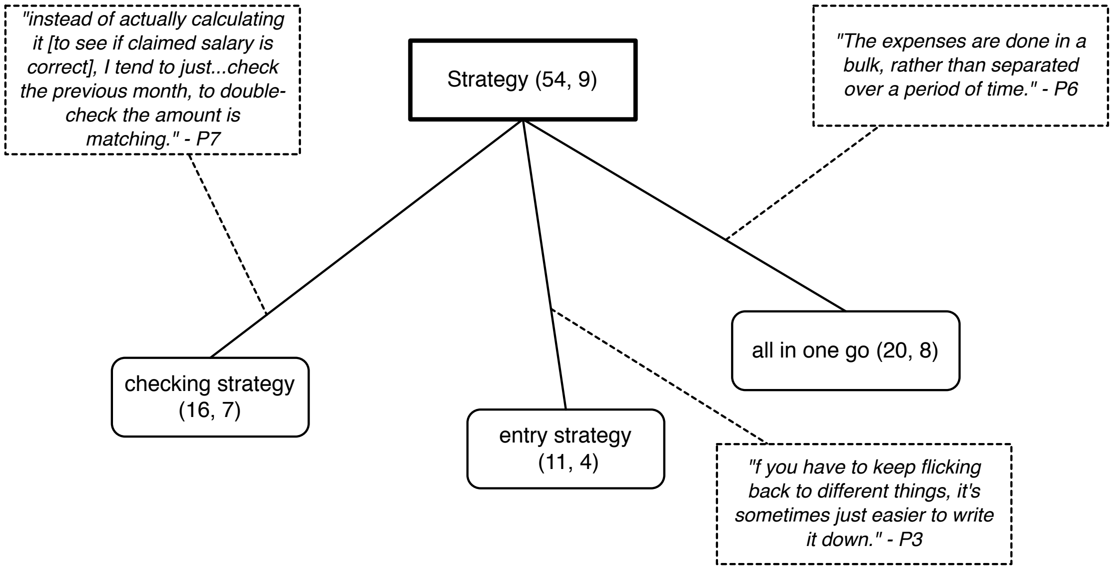


Figure A.7: Diagram showing the theme Strategy.

Participant	Quote/note
P3	*"I just try and do it in the quickest way...It's nice, once you've done it, it's completed, so it's sort your weight lifted [laughs]. So you don't need to think about it again."*
P6	*"the expenses are done in a bulk, rather than separated over a period of time. When I'm doing it lots at a time, I think once you get into sort of the hang of it, it gets done a lot quicker than..you just get used to putting them in, and inputting it all."*
P9	*"I try to concentrate on my task...I try to do one task [i.e. doing all expenses], finish one, and then do another."*
P4	It's difficult to take rests or even switch in-between number entry tasks because of the work pressure, and feels pressure by boss.

Table A.8: Most participants entered all numbers in one go.

Participant	Quote/note
P3	*" I wouldn't necessarily have to [memorise numbers], It's more just if you have to keep flicking back to different things, it's sometimes just easier to write it down, or just try and remember it. But you can obviously take the long version and keep flicking back to the correct screen."*
P2	*"we have different grants and different project codes as a result, but you, because you use them so much, you end up remembering them."*

Table A.9: Examples of strategies people used.

A.8 Importance of accuracy and paper trails

Quotes were grouped under this theme if participants talked about the sensitivity of financial data, which is why not all people were authorised to approve or access financial data, and the importance of a paper trail for data entries.

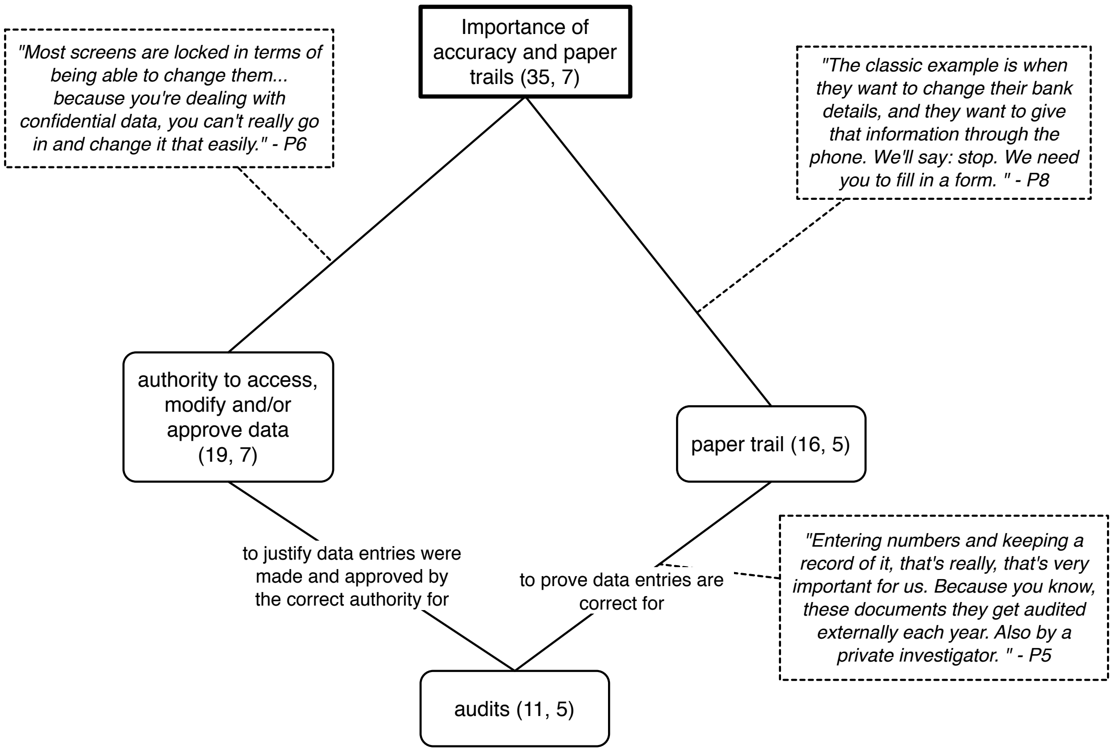


Figure A.8: Diagram showing the theme 'Importance of accuracy and paper trails'.

A.9 Other

Quotes were grouped under this theme if participants talked about things that did not fit into any other category but were still considered relevant, such as issues participants experienced, or queries they often received.

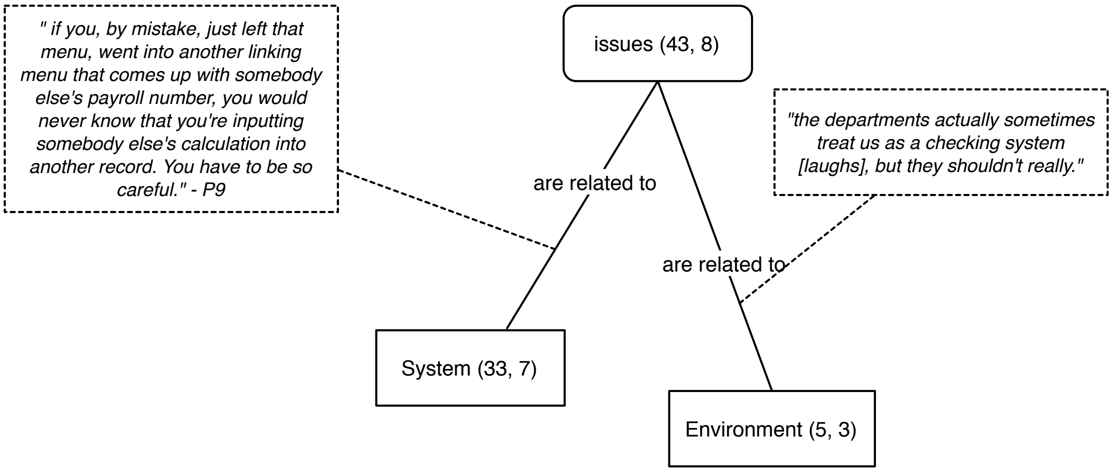


Figure A.9: If people described issues, it usually had to do with the system.

Participant	Quote/note
P5	*"There are other issues. You could say I think hundreds, I mean not just with the work that we do on expenses, but across [university A], across [university A] Finance, the Finance division...We just have to kind of work our way around the system and you know, adapt to it."*
P7	*"It's only the matter of how you get used to the Payroll system. Because companies have different systems, the data inputting can take a while to get used to it."*
P9	*"You know, all systems are a bit funny, I think. But you just gotta get used to it."*

Table A.10: Issues that participants experienced with the system.

Appendix B

Information sheet

The information sheet given to the participants in Study 1 is shown in Figure B.1.

You will be given a copy of this information sheet.

Title of Project: **Data entry in multitask settings**

This study has been approved by UCL Research Ethics Committee
[Project ID Number UCLIC/1415/001/Staff Brumby/Borghouts]

Name, Address and Contact Details of Investigators:

Dr. Duncan Brumby 8.23 Malet Place Engineering Building UCL Gower Street London WC1E 6BT United Kingdom brumby@cs.ucl.ac.uk	**Judith Borghouts** 8.18 Malet Place Engineering Building UCL Gower Street London WC1E 6BT United Kingdom judith.borghouts.14@ucl.ac.uk
+44 (0)20 7679 0689	+44 (0)20 7679 0693

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, please read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or you would like more information.

Details of Study
This study aims to investigate how people perform number transcription tasks. You will be asked about your experience with transcribing numbers, and asked to perform a number transcription task, to show the researcher how you would normally perform this type of task. We will also audio record the interview for further analysis.
The interview will take approximately **30-45 minutes** and you will be reimbursed with a **£10** Amazon voucher as a token of appreciation for your participation.

It is up to you to decide whether or not to take part. If you choose not to participate, you won't incur any penalties or lose any benefits to which you might have been entitled. However, if you do decide to take part, you will be given this information sheet to keep and asked to sign a consent form. Even after agreeing to take part, you can still withdraw at any time and without giving a reason.

All data will be collected and stored in accordance with the Data Protection Act 1998.

Figure B.1: Information sheet

APPENDIX C

CONSENT FORM

The consent form used for Study 1 is shown in Figure C.1.

Title of Project: **Data entry in multitask settings**

This study has been approved by the UCL Research Ethics Committee [Project ID Number UCLIC/1415/001/Staff Brumby/Borghouts]

Participant's Statement

I ……………………………………………......................................

agree that I have

- read the information sheet and/or the project has been explained to me orally;
- had the opportunity to ask questions and discuss the study; and
- received satisfactory answers to all my questions or have been advised of an individual to contact for answers to pertinent questions about the research and my rights as a participant and whom to contact in the event of a research-related injury.

I understand that I am free to withdraw from the study without penalty if I so wish, and I consent to the processing of my personal information for the purposes of this study only and that it will not be used for any other purpose. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Signed: Date:

Investigator's Statement

I ……………………………………………………………………….

confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signed: Date:

Researcher notes, optional clauses (the participant should indicate yes/no to these and initial):

- I understand that my participation will be audio recorded, and I am aware of, and consent to, any use you intend to make of the recordings after the end of the project.
- I agree to be contacted in the future for invitations to participate in follow-up studies.
- I understand that the information I have submitted will be published as a report and I will be sent a copy. Confidentiality and anonymity will be maintained, and it will not be possible to identify me from any publications.
- I understand that I am being paid for my assistance in this research and that some of my personal details will be passed to UCL Finance for administration purposes.

Figure C.1: Consent form

APPENDIX D

STUDY 1: INTERVIEW SCRIPT

The interview script used for Study 1 is given below. This script only served to guide the interview, and does not contain all questions that were asked. Based on what the participant was saying, follow-up questions were asked.

Before the interview

- ensure participant is aware of purpose research
- explain what will happen
- informed consent
- ask for permission to audio record interview

Work

- Tell me something about your work (what do you do)
- How many hours per week (full-time/part-time)
- How long have you been working here (at this company)
- How long have you been doing this type of work

Number entry

- What activities do you do for work that involve transcribing numbers? e.g. filling in expenses, tax returns, setting up invoices
- How often do you do this (per day/week)?
- How many numbers is it roughly that you have to enter?
- How long do you usually take?
- What type of numbers? Usually same numbers, or can it be anything?
- Do you get to enter numbers that are different from your familiar format? e.g. 2,000 or 2.000; 9/15/14 instead of 15/9/14
- Do you deal with foreign currencies?
- Tell me something about how you enter these numbers
- When do you do these tasks? Immediately when you get them, or save them for later? Morning, afternoon?
- Does urgency/time pressure influence how you do the task (if so, how)
- Do you do them in-between other tasks or save a particular part of the day for it?
- Do you do all tasks all at once, or take rests in between? (if rests, what do you do? switch to another task, have a coffee, lunch, break, etc.)
- Do you feel that the way you enter it changes after a while? e.g. you get better at it so it kind of becomes automatic, or less mentally exhausting? Or is it the opposite, becomes more exhausting?
- Do you do other things as well during this task e.g. listening to music, attending to another task
- Do you sometimes have to briefly store numbers in memory, or calculate them from numbers you already have? If so, do you use external tools to offload memory?
- Where do you copy them from? Paper, digital files, combination?
- Do numbers get checked, to see if they're correct? Do you or anyone else check these numbers?

- Do you ever get entered numbers from someone else, that you then have to check if they are correct?
- What is your general experience with transcribing numbers? e.g. easy, boring, part of the job

Environment

- Do you always work in the same environment, or sometimes work in different places, such as at home, or when you're on the train, or working at a cafe? What about number entry tasks?
- Do you do your work on a desktop, laptop, tablet, anything else? Are some devices harder or easier?
- How is your desk organized?
- Do you organise it differently when doing number entry tasks?
- Do you have notifications on (e.g. e-mail, work-related instant messaging); if you do get new notification, do you attend to it straight away or finish task first?
- Do you get interrupted in other ways, for example when the phone is ringing, or when a colleague or your boss asks you something? How do you deal with these interruptions? What is your experience with these interruptions?
- Critical incident: Has there ever been an incident where a mistake in entering a number went undetected, and was discovered later on?

Demonstration

- Could you show me the software you use to transcribe numbers? What is your experience with this system, works well? (If negative, how do you deal with that? do you use any strategies to make it more optimal for yourself?)
- Do you feel confident entering the numbers?
- How do you place your windows?
- Could you show me how you perform a typical number transcription task (do it how you would normally); if you feel uncomfortable about sharing work data, you can enter any

type of numbers, as long as it somewhat resembles data you would normally enter for work

After the interview

- Thank participant
- explain what will happen to their data
- do they have any more questions
- clarify when they will be compensated
- Ask if participant knows any further people who might be suitable and willing to participate

APPENDIX E

STUDY 2: DISTRIBUTED COGNITION MODELS

To aid the analysis of the contextual inquiry data collected in Study 2, I developed three Distributed Cognition models using Furniss & Blandford (2006)'s guidelines:

- The physical model: this model describes the physical layout of the task environment
- The information flow model: this model describes how information flows through all users involved in the task
- The artefact model: this model describes all artefacts involved in the task

These models were used to gain insight how information sources were distributed in the task environment, and were used to understand differences in inquiries. Each model consists of a diagrammatic representation that visualises the data, and a narrative representation which verbally describes the data.

The three models are described below. I used the DC principles as described by (Furniss & Blandford, 2006) as guidelines to decide what to include in the models. The principles of each model are described at the start of each section, and marked in italics (e.g., *horizon of observation*) in the narrative descriptions.

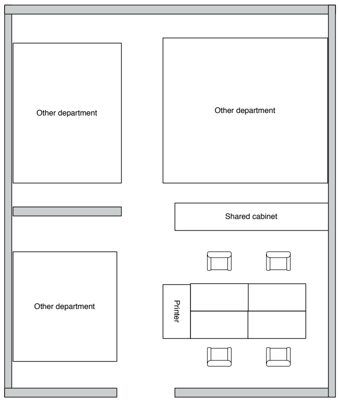


Figure E.1: Physical model diagram showing a typical physical layout of people's work environment at room level.

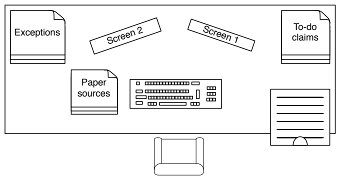


Figure E.2: Physical model diagram showing the physical layout of people's work environment at desk level.

E.1 Physical model

The physical model describes what the individual can physically hear, see, and access, and how information sources are laid out in the physical environment. In developing the model the following was considered: what is the proximity of, and access to, devices and people: what can be seen and heard from the individual's point of view?

Physical model principles

- Space and Cognition: how do people use the physical space to support their work
- Subtle bodily supports: do people use their body to support their work
- Situation awareness: are people informed of what is going on
- Horizon of observation: what are people able to see and hear
- Arrangement of equipment: how do people arrange their equipment

The physical layouts of the four offices were not identical, but shared a number of characteristics. All participants had their own desk and worked in an open office with two or more colleagues. P3 was the only person in the office responsible for data entry tasks. The other eight participants had colleagues in the same room who dealt with similar tasks. P1, P2, P8 and P9 also had colleagues in the same room that worked on different tasks. Colleagues that were working on different tasks were situated further away from them than other colleagues. As an example, the physical environment of the office of P1 and P2 is depicted in Figure E.2 at room (a) and desk (b) level. The number of people in one room ranged from two to sixteen.

The open layout made it easy for colleagues to interact with each other and share information between themselves. They could see when a colleague was present and available to consult. Participants regularly consulted colleagues in their room to retrieve information they could not find on any other sources. This information is given verbally, or a colleague directs the participant to the correct information source.

The offices had an open-door policy, meaning that workers could at any time be interrupted by people walking in. During the observations, participants were regularly interrupted by colleagues. They responded to it but used *subtle bodily supports* to not lose track of where they were in the expenses task. For example, P3 responded verbally to a colleague but kept his kept his visual attention on the computer and tried to continue with the expenses task. When P1 was interrupted by a colleague, he placed a finger on the computer screen to remember where he was in the task.

All participants worked with both paper and digital information sources. Several participants used the physical *space* to organise their paper sources. P2 maintained separate physical locations on his desk for expense claims to be processed, expense claims that had been completed, and exceptional claims that required further attention. P7 also had paper sources fixed on his wall, which were visible from his desk.

Other physical sources were located in an individual's drawer, in a shared bookcase or drawer in the same physical room. These were not visible from their desk, and participants had to physically move from their desk to retrieve and view this information. Participants could see whether a colleague is present or not, and whether they can consult them at that moment to request information. Recent physical files are kept in a closet in the same room, whereas older files and employee files are physically kept in another location. These older files were used less frequently. If P6 received expense claims, she placed these in her drawer to return to later. If there were a lot of claims to be processed or if the participant started processing claims, these were placed on their desk.

Participants become aware of new claim requests by batches of claim forms or receipts on their desk with a handwritten note, the claimant asking them in person, or when they check their e-mail and see a new e-mail by the claimant. For physical claims, they browse through them and place them either at a dedicated area on their desk or in their drawer, to return to later. P2, P3 and P6 re-ordered claims based on urgency. Claims that were more urgent were processed first. These could be claims by their boss, or claims with an explicit instruction that it was urgent. In addition, P2 sometimes grouped claims according to category. The reason for grouping them is that different categories of expenses can require other data fields to be filled in, and P2 felt it to be easier to do these particular type of claims together. Examples are travel expenses for which participants have to fill in the departure point and destination, or external lunches for which the name of the restaurant has to be filled in. Participants place exceptional claims in a separate physical location. For P2, this was in a separate tray in the top-left hand corner of his desk. By putting it in his *horizon of observation*, he can see if there are still claims in this tray to be processed and return to at a later point in time. P2 starts each day by seeing if there are claims in this tray, and processes these first. P5 places exceptional claims in a separate drawer. They are not visible and she does not use any reminders, but returns to these when she sees fit.

Participants also received requests digitally through email. They could not place these in their horizon of observation. However, whether the claim requests were in physical or digital form, they always needed the physical receipts of expenses, before they could process a claim. Similar to physical claim forms, they had these receipts either in their drawer or on their desk, as a reminder that they still needed to process these claim requests.

Participants also created their own artefacts to aid them in their work. This includes physical artefacts, such as a spreadsheet with frequently used codes, as well as digital artefacts, such as a digital form to look up codes by person name. Participants could type in a name of a claimant, and the form populated codes related to one person. This made it easier for people to look up which account codes to charge the expenses to.

Digital information sources included the data entry system, intranet, and external websites via their computer. Furthermore, they had digital files, such as PDF files, Excel spreadsheets, and Word documents stored on their desktop computer. The *arrangement of equipment* was that seven out of nine participants had two screens. All seven participants mainly used one screen, both to enter and retrieve information. Some participants used a second screen to display their email inbox. If they received a notification of a new email, they would briefly glance to determine the urgency and importance of the email, but would try to continue with the expenses task.

Sometimes a claim was checked by multiple people in the same room. Participants had *situation awareness* of the progress of the claim as long as it was still in the same physical office. They could see the progress of expense claims by the pile on colleagues' desks, and could ask colleagues in person. Furthermore, they can overhear conversations on the phone and become aware of claims that require further attention. Participants can see their own screens and desks. They have to walk over or do additional physical actions in order to view what is on colleagues' desk.

As soon as an expense claim was submitted to another office and physical location, the user would have little *situation awareness* of the progress of the claim. The system did not have a visible status update, meaning participants can not see the progress of their claim once it has been submitted to Central Finance. As they do not receive updates, administrators and claimants often forget to keep paying attention to it and do not query it until they realise payment has not been processed yet. Often an error will have occurred and at this point the project may have finished and payment is no longer possible. If participants needed more information on the status of a claim, they needed to contact colleagues from another office via email and telephone, or they had to visit the other physical location. Participants called colleagues from other departments with queries such as errors, and outstanding claims that had not been processed yet. They emailed claimants with queries such as if they do not agree with the expenses claimed, if they need further information, or if they have spotted an error.

E.2 Artefact model

The artefact model describes the artefacts that are used.

Artefact model principles

- Mediating artefacts: do people use any artefacts to support their work
- Creating scaffolding: do people use the environment to simplify tasks, e.g. set reminders
- Representation:goal parity: do people use artefacts to display the explicit relationship

between the current and goal state of their work

- Coordination of resources: how do people coordinate their information

Participants work with both physical and digital artefacts. Table **??** shows an overview of the information sources that were involved in an expenses task. For each instance of the task, several artefacts can be used, but there are two main artefacts that are used in each instance and are central to the expenses task: the paper receipts and the expenses entry system. In addition to information sources, participants use several *mediating artefacts* to support their work. Calculators are used to aid in calculating sums. Multiple tools were consulted to convert currencies: an external website, a tool on the intranet, and a tool included in the data-entry system itself. They use a physical tray on their desk to hold exceptional claims. They use a pen to annotate receipts and highlight which items on receipts to claim back.

Some of the participants *created external scaffolding* to simplify their task. In particular, participants had difficulties remembering codes they needed to enter. P5 and P6 had made a personal spreadsheet with codes they used most frequently. P4 remembered old codes but had difficulties remembering new codes since they had changed 18 months ago. To look up codes, she used a spreadsheet created by the departmental manager where she could fill in old codes, that would populate the correct new code.

The parity between the current and goal state was displayed on the expenses system. Once a claim had been submitted, there would be a status update in the expenses system. This status reveals whether a claim is Pending, has been Paid, or whether Original receipts are required. As long as a claim was Pending however, there was however no insight into what is happening on the other side. Often a claim would be received but would be held because there was information missing. Participants would not know about this unless they explicitly contacted the office, and depending on the situation, they would often still not receive information on what was happening with a particular claim because the information was not centralised in Central Office and the person on the phone would not be able to look up what was happening with a particular claim.

In order to *coordinate resources*, the intranet was intended as a central point for claimants, administrators and Central Finance officers to access the same information resources. Participants find the resources difficult to use, so people often end up making their own local copies and working with these instead. This supports them better in their work, but they can end up working with old and incorrect information if information gets updated. Examples of information sources used locally are claim forms: claimants keep working with local copies on their computer and do not download the new forms. They keep using old information such as old salaries and old project codes. Another example are spreadsheets with budget codes: administrators created and used their own spreadsheets with codes. If codes were updated, they were not aware and ended up working with old codes. There was an instruction manual to do expenses, to ensure everyone carried this routine task in the same way. Newcomers usually learnt how to do it from other people rather than this written manual, as the experience was that it was often easier and faster to learn it this way. This way of learning the activity again had the risk that some people were doing it in the old, incorrect way, and passed on this incorrect way of doing work to others. In order to provide proof of expenses, people still had to provide hard-copy receipts. These need to be sent to Central Finance, but often get damaged and lost and the claimant and administrator will not know, but Central Finance will not know either so nothing will happen unless the administrator or claimant takes action and chases it.

In order to prevent people from interrupting an expenses task, people were logged out of the data entry system after a period of inactive use and they had to restart the task from the beginning. This added cost to resume the task kept participants focused on the data entry task, and they were less likely to interrupt and switch to unrelated tasks. However, people often did not know beforehand what the cost to access information was going to be. Furthermore, it was also not clear after how long the timeout would occur.

E.3 Information flow model

This model describes the flow of information through several actors for an expenses task.

Information flow principles

- Information movement: how does information move throughout the system
- Information transformation: does the representation of information undergo changes
- Information hubs: what are the main points where different information channels meet

- Buffering: are there any buffers to uphold information
- Communication bandwidth: how does communication take place
- Informal communication: does informal communication take place in addition to formal communication
- Behavioural trigger factors: are there any local factors that individuals respond to

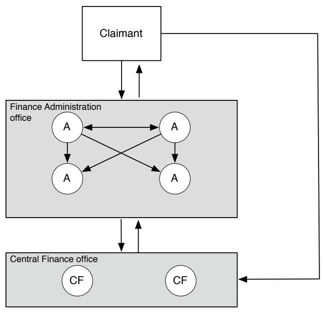


Figure E.3: Model of the information flow.

An expenses claim *moves* through several actors, and moves between actors via email, phone, physical post, and face-to-face communication. The actors who contribute to the processing of an expense claim have limited visibility on the overall status and progress of the claim. For example, once claimants submit a claim request to the administrator, they do not know what the status is of that claim until they receive an email notification that it has been completed. Similarly, once administrators submit a claim to the Central Finance office, they do not know what the status is of the claim. They do not know when or whether it has been processed and if not, what the reasons are for holding it. The workers at the Central Finance office know the reasons for withholding a claim, but often receive incomplete information of a claim, and for example do not know the justification behind a claim, whether the expenses are made correctly and if there is an error in the project code entry, they do not know the correct project to charge it to.

Information transformation takes place when calculations have to be carried out. At the beginning the individual numbers are saved, as well as the calculations on those numbers. Once a claim is submitted only the end result will be saved on the system. For example, if one claim request involves multiple expenses, each individual amount has to be checked by the administrator. Administrators are then free to choose whether to type each amount or only the sum total on the system. Once they have processed and submitted the claim, only the sum amount will be available for auditors.

There are two main *information hubs*: the first hub is the office of the administrator, who deals with incoming claims and hard-copy receipts from claimants. These claims are processed at the administrator office and then sent off to the Central Finance office, the second information hub. Workers at this office deal with incoming claims and hard-copy receipts from administrators, match and process these and submit them for payment.

The administrator is the main *information buffer* between the claimant and the Central Finance officers. Claimants submit a request to administrators. This claim is upheld until the administrator decides to process it and send it to Central Finance. If there was an issue with a claim, claimants contacted the administrator, who then contacted Central Finance. Though claimants could also contact Central Finance directly, administrators said it was often easier if they contacted Central Finance on their behalf, as they knew who to contact [quote].

Communication between the claimant and administrator takes place face-to-face, over the phone, via email and via handwritten notes. Communication between colleagues takes place face-to-face. Communication between the administrator and Central Finance solely takes part via email or over the phone, though it is possible to take place face-to-face.

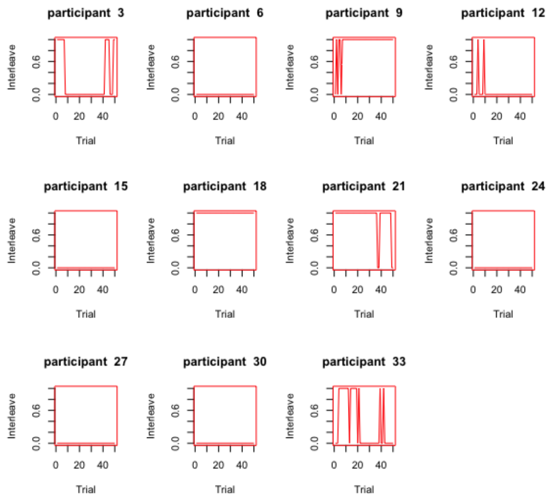
Instructions are mostly *communicated informally* through word of mouth. Knowledge of how to use the system sits with the employees, and it is often faster to explain newcomers how to do it rather than go through the written instructions. A consequence is that when information gets updated, not everyone is aware of it and keeps using the old and incorrect way, or learn the incorrect way from someone else who is still using the old way.

Receiving a claim request from another actor are the main *factors triggering behaviour.* Participants collected claim requests and saved them to return to later. Some participants kept claims to be completed on their desk. The size of this pile acted as a trigger to decide whether to start processing them. Furthermore, the payroll deadline was another trigger. Participants tried to complete claims before the deadline so claimants were reimbursed in time.

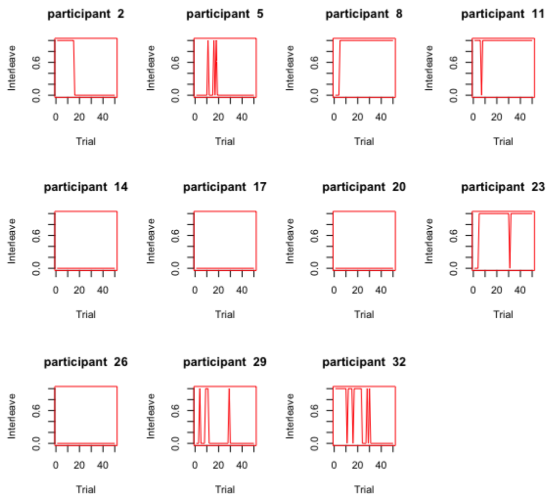
Appendix F

Study 4: interleaving rates per participant

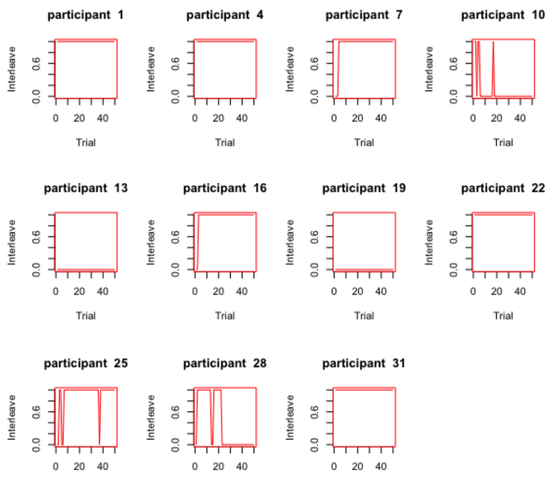
Study 4 investigated the effect of time costs on the timing of inquiries. For this purpose, on a trial-by-trial basis it was considered whether participants entered expenses in sequential order, or whether they entered items with a low cost of each expense first. Across conditions, participants were mostly consistent in strategy choice, and interleaved on either all or no trials. The consistency in strategy choice per participant is further illustrated in Figure F.1, which displays a plot for each participant across trials. The x axis plots the trial number, and the y axis displays whether they interleaved on that trial or not: a value of 0 means they did not interleave, and a value of 1 means they did interleave. These plots further illustrate that most participants were consistent in interleaving on no or all trials, as the majority of plots have a flat line (see for example Participant 6, who interleaved on all trials). A subset of participants switched between strategies at the first couple of trials before sticking with one strategy, such as Participants 9 and 12: at the start of the x axis, their lines go up and down between 0 (no interleaving) and 1 (interleaving) before becoming a straight line. Lastly, participants 29, 32 and 33 seemed to switch between the strategies throughout the experiment and did not stick with a particular strategy: their lines continue to go up and down between 0 and 1 along the entire x axis.



(a) Participants in the Control condition.



(b) Participants in the High-Amount condition.



APPENDIX G

STUDY 7 MATERIALS

LONDON'S GLOBAL UNIVERSITY

UCL

CONSENT FORM FOR RESEARCH PARTICIPANTS IN RESEARCH STUDIES

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study:	Tasks, document and window management in the workplace
Department:	UCL Interaction Centre
Name and Contact Details of the Researcher(s):	Judith Borghouts Judith.borghouts.14@ucl.ac.uk
Name and Contact Details of the Principal Researcher:	Dr. Duncan Brumby brumby@cs.ucl.ac.uk
Name and Contact Details of the UCL Data Protection Officer:	Nico Preston n.preston@ucl.ac.uk
This study has been approved by the UCL Research Ethics Committee:	**Project ID number:** UCLIC/1415/001/Staff Brumby/Borghouts

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I confirm that I understand that by ticking/initialling each box below I am consenting to this element of the study. I understand that it will be assumed that unticked/initialled boxes means that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element that I may be deemed ineligible for the study.

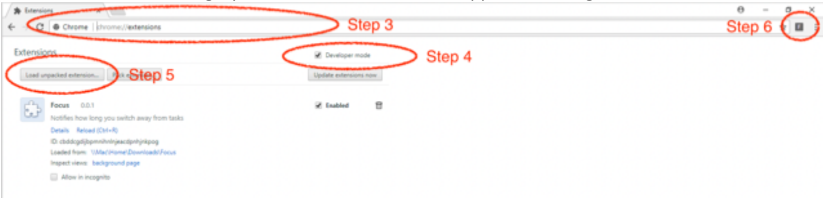
		Tick Box
1.	*I confirm that I have read and understood the Information Sheet for the above study. I have had an opportunity to consider the information and what will be expected of me. I have also had the opportunity to ask questions which have been answered to my satisfaction and would like to take part in an individual interview.	
2.	*I understand that I will be able to withdraw my data up to 4 weeks after the interview.	
3.	*I consent to the processing of my personal information for the purposes explained to me. I understand that such information will be handled in accordance with all applicable data protection legislation.	
4.	*I understand that all personal information will remain confidential and that all efforts will be made to ensure I cannot be identified. I understand that my data gathered in this study will be stored anonymously and securely. It will not be possible to identify me in any publications.	
5.	*I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason. I understand that if I decide to withdraw, any personal data I have provided up to that point will be deleted unless I agree otherwise.	
6.	I understand the potential risks of participating and the support that will be available to me should I become distressed during the course of the research.	
7.	I understand the direct/indirect benefits of participating.	
8.	I understand that the data will not be made available to any commercial organisations	

University College London, Gower Street, London WC1E 6BT
Tel: +44 (0)20 7679 2000
email@ucl.ac.uk
www.ucl.ac.uk

Figure G.1: Study 7 consent form given to participants.

Instructions to install Focus browser extension

How to install the browser extension
1 Download the Focus.zip here: https://www.dropbox.com/s/q3iyno8f7fyfbfc/Focus.zip?dl=0 Unzip it and save the Focus folder somewhere on your computer.
2 Open Google Chrome. If you do not have Google Chrome, download it here: https://www.google.co.uk/chrome/
3 In Google Chrome, enter **chrome://extensions/** in the URL bar, as in the picture below (or click on the 3 grey dots in the right-hand corner and go to More Tools > Extensions).
4 Check the box for **Developer mode** at the top of the page.
5 Click the **Load unpacked extension** button and select the unzipped Focus folder to install it.
6 Next to the URL bar, a grey icon with an F should now appear in the right-hand corner.



How to use
• The browser extension gives you additional time information on a particular task you are working on at the moment.
• To select a task that you want to focus on, go to the relevant webpage and click on the extension F icon. A pop-up should appear saying 'This is now the main task page'.
• Every time you switch away from this page, a notification will appear showing you how long an average you go away for, before you return to this window.
• You are free to choose if, when and how often to use it for any task. If you do choose to use it, ideally you use it at least once for a task involving entering expenses.
• If you want to stop the extension, simply close or refresh the page.

To be able to use the extension on a webpage:
• The URL of the webpage has to start with https://
• You have to allow notifications on the webpage. When you click on the F icon of the browser extension, a popup should appear asking you to enable this.
• The extension can be used for any other task which is conducted in the Google Chrome browser, such as composing an email in Outlook, etc. Even though we have tried to test it on as many websites as possible, it may not work (yet) on some webpages (e.g. Google Docs and Office Online).

Data collection
• The browser extension does not store any data.
• The ManicTime application stores how much time you spend on each computer window. At the end of the study, you will be asked to share your ManicTime database.
• If the data contains any sensitive data, I will assist you to make necessary amendments, such as removing any (or all) names of applications and websites visited. The main focus of this study is how many windows are involved in office work, rather than what these particular windows are.
• All your data will be kept confidential and used for research purposes only, and will not be shared with anyone outside the research project.

Figure G.2: Study 7 instructions to install the browser extension.