
Understanding how people look up and enter information for a data entry task in financial office settings

Author:
Judith BORGHOUTS

Supervisors:
Dr. Duncan BRUMBY
Dr. Anna COX

May 11, 2017

UCL Interaction Centre,
Department of Psychology,
University College London

Contents

Abstract	1
List of Tables	4
List of Figures	5
1 Introduction	6
2 Background	8
2.1 Data entry	8
2.1.1 The perception stage	9
2.1.2 The encoding stage	11
2.1.3 The execution stage	13
2.1.4 The checking stage	15
2.2 Office work	18
2.2.1 Reducing information access costs: multiple and larger screens	18
2.2.2 Managing information needs	20
2.3 Conclusion	21
3 Entering expenses in a financial setting	22
3.1 Study 1: Understanding data entry work in a financial office	22
3.1.1 Introduction	22
3.1.2 Method	23
3.1.3 Results	26
3.1.4 Discussion	41
3.2 Study 2: Information sources for entering expenses	43
3.2.1 Introduction	43
3.2.2 Method	44
3.2.3 Findings	47
3.2.4 Discussion	54
4 Understanding the effect of IAC on looking up information	57
4.1 Study 3: Copying numbers from one source	57
4.1.1 Introduction	57
4.1.2 Method	59
4.1.3 Results	62
4.1.4 Discussion	68
4.2 Study 4: Copying data from multiple sources	70
4.2.1 Introduction	70
4.2.2 Method	71
4.2.3 Results	74
4.2.4 Discussion	78
4.3 Study 5: Interleaving between data entry tasks	79

4.3.1	Introduction	79
4.3.2	Method	80
4.3.3	Results	81
4.3.4	Discussion	83
4.3.5	Conclusion	83
5	Evaluating design changes to improve performance in data entry with varying IACs	78
5.1	Introduction	78
5.1.1	Contributions	79
	References	80
	Appendix A Information sheet	86
	Appendix B Consent form	88
	Appendix C Interview script	90

List of Tables

3.1	Study 1 participant information	24
3.2	Study 1 interview topics	25
3.3	Study 1 Themes	27
3.4	Study 1 checking quotes	30
3.5	Study 1 checking own input	31
3.6	Study 1 checking other people's input	31
3.7	Study 1 errors quotes	35
3.8	Study 1 type of errors quotes	36
3.9	Study 1 reasons for errors quotes	36
3.10	Study 1 acceptable errors quotes	36
3.11	Study 1 batching quotes	38
3.12	Study 1 strategy quotes	38
3.13	Study 1 issues quotes	40
4.1	Study 3 dependent variables	59
4.2	Study 3 descriptive measures	63
4.3	The rates (and frequencies) of ordering strategies per condition.	74
4.4	The rates (and frequencies) of grouping strategies per condition.	75
4.5	The rates (and frequencies) of grouping strategies for trials where people entered items in column order.	76
4.6	The rates (and frequencies) of grouping strategies for trials where people entered items in row order.	76
4.7	The most common grouping strategy was to chunk the items into three groups. The strategies are shown graphically in Figure 4.8.	76
4.8	The means (and standard deviations) of error rate and trial completion time the three conditions. The rates are percentages calculated as the ratio of occurrences to the number of opportunities.	77
4.9	Results of the dependent variables for the three conditions. The rates are percentages calculated as the ratio of occurrences to the number of opportunities.	81
4.10	The number of trials in which participants interleaved for the three conditions.	82

List of Figures

2.1	Different stages of a data entry task	8
2.2	An incremental and keypad number entry interface.	15
3.1	Study 1 Task diagram	29
3.2	Study 1 Checking diagram	30
3.3	Study 1 System diagram	32
3.4	Study 1 Environment diagram	33
3.5	Study 1 Data diagram	34
3.6	Study 1 Errors diagram	35
3.7	Study 1 Strategy diagram	37
3.8	Study 1 Importance of accuracy and paper trails diagram	39
3.9	Study 1 Other diagram	40
3.10	Study 2 Room model	47
3.11	Study 2 Desk model	47
3.12	Study 2 Information flow model	51
4.1	Study 3 task lay-out	60
4.2	Study 3 number of visits	64
4.3	Study 3 duration of first visit	65
4.4	Study 3 number of blocks correctly placed	66
4.5	Study 3 placeholders	67
4.6	Participants had to enter four data items in a data entry form (a). Each item had to be retrieved from a separate window (b) and entered into the form (c).	72
4.7	The order strategies. (a) If people entered Amount 1 and then Account1 (or Amount2 and then Account2), this was categorised as a 'Row' strategy. (b) If people entered Amount 1 and then Amount2 (or Account1 and then Account2), this was categorised as a 'Column' strategy.	75
4.8	The sequence of the most common grouping strategies. V = visit to the data source, E = entry of the data item. For example, in Strategy (a) a participant first viewed Amount1 and Account1, and then entered Amount1 and Account1. He/she then viewed Amount2 and entered it, and then viewed Account2 and entered it.	77
4.9	Participants had to enter two data entry forms per trial, each containing two items. (a). The amounts for both forms had to be retrieved from one window (b) and entered into the form (c), while the account codes had to be retrieved from a second window (d) and entered (e).	81
A.1	Information sheet	87
B.1	Consent form	89

Chapter 4

Understanding the effect of IAC on looking up information

Chapter outline

The previous chapter demonstrated that for an expenses task, people have to copy data from multiple sources that differ in terms of their IAC. This chapter describes two studies that explore the extent to which IACs influence strategies, speed and accuracy.

Study 3 tests how IAC affects switching behaviour when copying from one source. It tested whether the effect of IAC on copying colours, as shown in previous studies, extends to copying numbers. Study 4 tests how IAC affects switching behaviour when information for one entry task has to be collected from multiple sources. Together these studies intend to show that IAC affects how people switch between entering and looking up information.

4.1 Study 3: Copying numbers from one source

4.1.1 Introduction

Study 1 and 2 illustrated that people had to retrieve and copy data from multiple sources with varying information access costs (IAC). People were not always able to look at the data source while entering data, and information sometimes had to be briefly kept in memory before people could enter it. In this situation, people have various options: they can decide to keep going back to the target source, reduce their memory load by writing it down, or keep as much of it in memory. Participants in Study 1 explained they tried to do it in the quickest way possible, and preferred to keep data in memory as this was often a faster strategy than writing it down, printing it out, or going back and forth between the source and the input window.

This is in line with previous lab studies, which showed that as the cost to access information increases, people increasingly rely on the information in their head rather than the information in the environment (e.g. Gray, Sims, Fu, & Schoelles, 2006; Morgan, Patrick, Waldron, King, & Patrick, 2009). People look at the target source longer before copying anything, but after one look, they do enter more items in one sequence before looking back.

Gray & Fu (2004) conducted a study where people had to copy over VCR programming information. Participants either had permanent access to the information, the information was covered by a grey box which could be uncovered by hovering over it with the cursor, or they were explicitly instructed and trained before the trial to memorise the information. The people in the latter condition were more accurate than the other conditions in entering the information. This shows that a deeper encoding of the information in memory makes people more accurate in entering the information.

In later studies that further explored the effect of access costs on use of memory, the Blocks World Task (BWT) was used as a task paradigm, which requires people to copy a 3x3 grid of coloured blocks. In these studies, participants relied more on memory as the cost to access the target window increased. In one study, this memory-based strategy made participants better able to resume after interruptions by copying more blocks before having to revisit the target window (Morgan et al., 2009). Looking at overall task performance overall however, they did make more errors overall 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) which is more familiar and easy to memorise. This suggests that the type of information to copy influences the effect that IAC might have on task performance, though the studies differed in task paradigm making it hard to compare their findings: in Gray & Fu (2004) people were explicitly instructed to memorise the information and conducted 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 given this training or instruction.

Rather than training participants to memorise the information, Soboczenski, Cairns, & Cox (2013) used an alternative approach to motivate people to encode the information more deeply. They conducted two studies 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. Participants made fewer data transcription errors if data was shown in the harder-to-read font colour, both for transcribing text and numbers. In line with Gray & Fu (2004), these studies showed that when people do make the effort to more deeply encode information in a copying task, it can improve their accuracy.

In order to design data entry interfaces that support how people enter expenses from information sources with varying IACs, it is important to understand how these differences in IAC affect strategy and task performance. Relying on information in the head over information in the world can make it more likely for people to make errors in a copying task (Morgan et al., 2009), though other studies have shown that a deeper encoding of information can also improve accuracy if memorised well (Gray & Fu, 2004; Soboczenski et al., 2013).

The study reported in this chapter replicates the BWT study, by using both coloured and numbered blocks. The purpose of this study is to see whether the effect on IAC on people's strategies and performance, as reported in previous BWT studies, holds when people have to copy numbers instead of colours.

It is expected that the effect of IAC on strategy will be similar and that an increase in IAC will encourage people to adopt a more memory-based strategy. However, the expectation is that numbers are easier to memorise and people will be able to memorise more items. Furthermore, based on previous findings that a deeper encoding of numbers in memory can reduce errors, it is expected that an increased IAC will improve accuracy for copying numbers.

4.1.2 Method

Participants

Forty-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.

Design

A mixed design was used with two independent variables: IAC and block type. The between-participants variable was the level of IAC which had three levels. If the IAC was Low, the target pattern was permanently visible. In the Medium and High IAC 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 IAC 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, 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. 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 dependent variables were used to measure 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.

Strategy measures

Number of visits to target window

Visit duration of first visit (s)

Average duration of visits (s)

Number of blocks copied after first visit

Number of blocks copied correctly after first visit

Global task performance measures

Number of incorrectly placed blocks (per trial)

Number of incorrectly submitted trials (per experiment block)

Trial completion time incl. and excl. lockout (s)

Table 4.1: Dependent variables used in the study.

Trial: 1/15
Please copy the pattern on the left into the window on the right, using the blocks at the bottom of the screen.

3	4	7
5	9	6
1	8	2

Submit

Resource Window

(a) The number condition.

Please copy the pattern on the left into the window on the right, using the blocks at the bottom of the screen.

Submit

(b) The colour condition.

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

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.

For the Low IAC condition, eye fixations were used to measure the number and duration of visits to the target window. Eyetracking data was also obtained for the Medium and High IAC conditions. However, this data was not used due to the fact that people were able to also view the target window area whilst the target pattern was covered. Therefore, in accordance with previous IAC studies (e.g. Gray & Fu, 2004), for the Medium and High IAC conditions the number and duration of uncovering the mask was taken as a measurement for visits to the target window. These uncoverings were measured by Javascript. The usefulness and limitations of using these measures are discussed in the Discussion.

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.

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. 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. They were then asked to read and sign a consent form and 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 study took around 20-30 minutes to complete.

Ethical considerations

The study was undertaken with ethical approval from the UCL Research Ethics Committee [Project ID Number UCLIC/1415/001/Staff Brumby/Borghouts]. At the start of each study, participants were first briefed verbally about the study. They were asked to read and sign a consent form, and were given an information sheet to keep. This information sheet contained

a summary of the study information and the researchers' contact details. It was explained that an eyetracker would record their eye fixations and movements, but that these recordings were anonymous and that they would not be directly identifiable. After participants had completed the first part of the experiment, a prompt appeared on the screen advising them to take a short break. Participants could take a break as long as they wanted and could decide themselves when to continue with the second part of the experiment.

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.

4.1.3 Results

The means and standard deviations of all dependent variables are shown in Table 4.2.

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.

Task strategies

Number of visits to the target window

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 (see Figure 4.2), 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, $F(1,30) = 3.05$, $p = .09$, $\eta^2 = 0.09$. 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, $F(2,30) = 5.70$, $p < .01$, $\eta^2 = 0.28$ (see Figure 4.3). 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 ($M = 1.49$, $SD = 1.01$), $t(11) = 2.76$, $p = 0.02$.

1.3						
	Colours			Numbers		
	Low	Medium	High	Low	Medium	High
Strategy measures						
Number of visits to target window	6.36	4.24	2.98	5.10	2.03	2.05
	2.28	1.62	0.90	2.48	0.63	0.67
Visit time of first visit (s)	0.39	0.04	2.18	0.51	0.04	1.49
	0.23	0.02	1.59	0.45	0.05	1.01
Average time of visits (s)	0.29	0.04	1.54	0.35	0.04	1.07
	0.13	0.02	0.95	0.15	0.03	0.77
Number of blocks copied	1.90	3.55	4.52	2.44	6.18	6.33
after first visit	1.84	1.93	1.43	1.89	1.61	1.67
Number of blocks copied correctly	1.86	3.22	4.07	2.36	5.96	5.98
after first visit	1.75	1.83	1.20	1.74	1.52	1.55
Global task performance measures						
Number of incorrectly placed blocks	0.15	0.67	0.79	0.17	0.31	0.46
(per trial)	0.18	0.40	0.44	0.19	0.18	0.16
Number of incorrectly submitted	0.27	1.9	2	0.36	0.5	0.83
trials (per experiment block)	0.65	2.51	2.13	1.21	1.08	1.03
Trial completion time incl. lockout	19.60	25.40	31.80	19.47	20.83	25.95
(s)	2.98	5.16	6.08	3.03	3.08	4.21
Trial completion time excl. lockout	19.60	25.40	28.84	19.47	20.83	23.89
(s)	2.98	5.16	6.34	3.03	3.08	4.06

Table 4.2: The effect of IAC on copying colours and numbers. The means are shown in bold, the standard deviations are below the means.

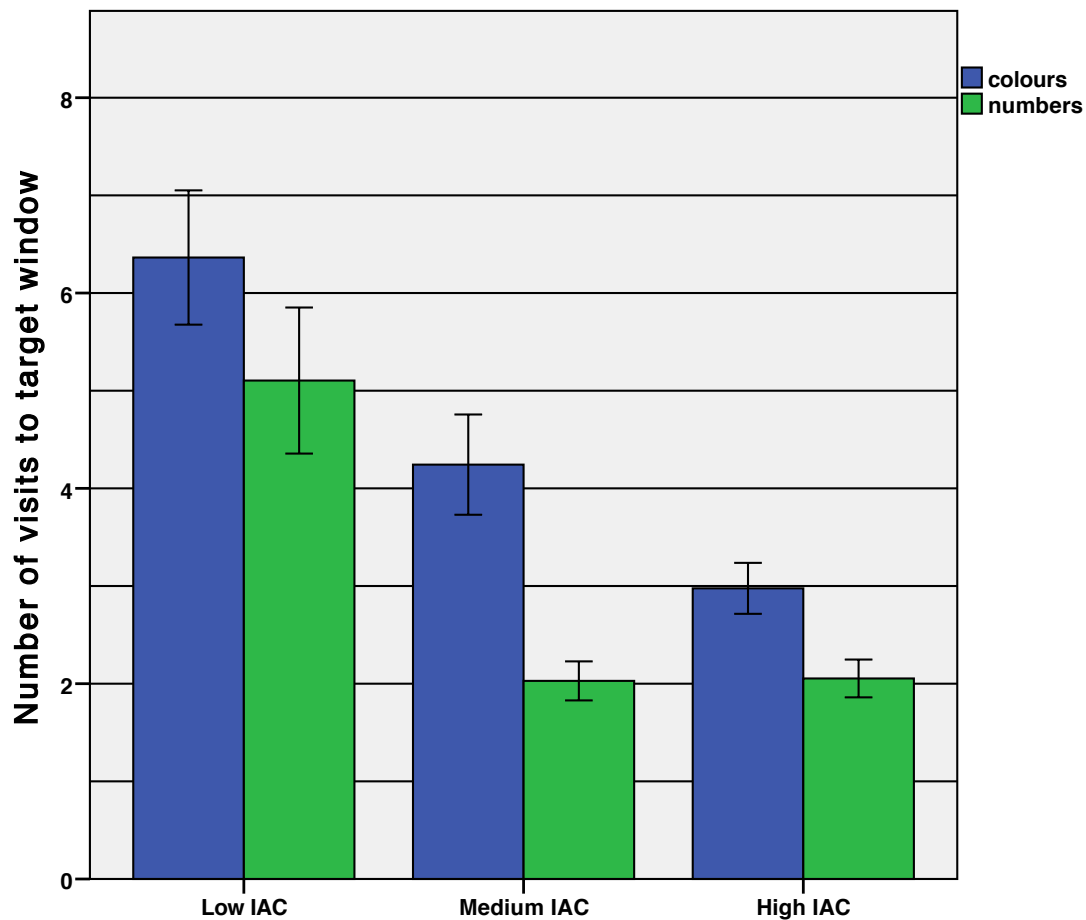


Figure 4.2: The interaction between block type and IAC for number of visits to the target window. The error bars represent ± 1 standard error.

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, $F(2,30) = 12.54$, $p < .001$, $\eta^2 = 0.46$. 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, $F(2,30) = 8.96$, $p < .01$, $\eta^2 = 0.37$ (see Figure 4.4). 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.

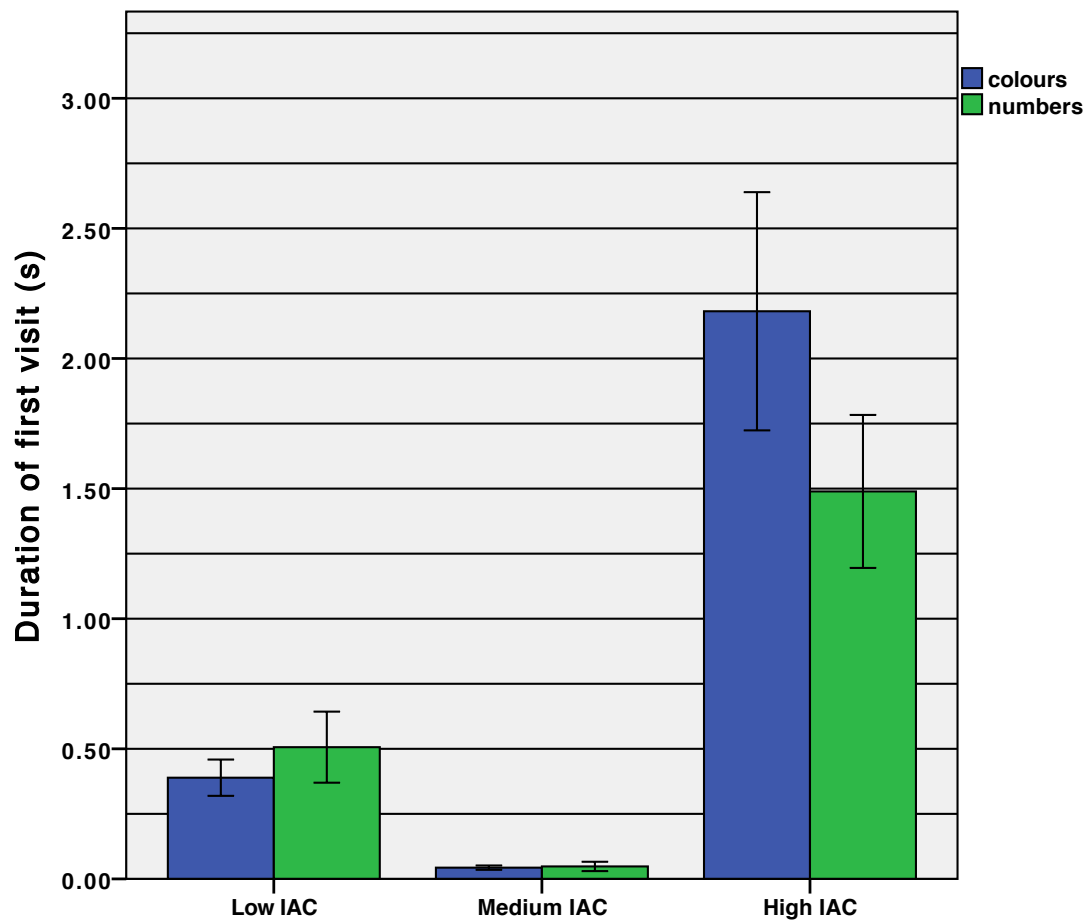


Figure 4.3: The effect of IAC on the duration of the first visit to the target window. The error bars represent ± 1 standard error.

Number of incorrectly placed blocks

Participants placed more blocks incorrectly for colours ($M = 0.54$, $SD = 0.45$) than numbers ($M = 0.32$, $SD = 0.21$), $F(1,30) = 10.71$, $p = .003$, $\eta^2 = 0.26$. 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, $F(2,30) = 1.65$, $p = .2$, $\eta^2 = 0.10$.

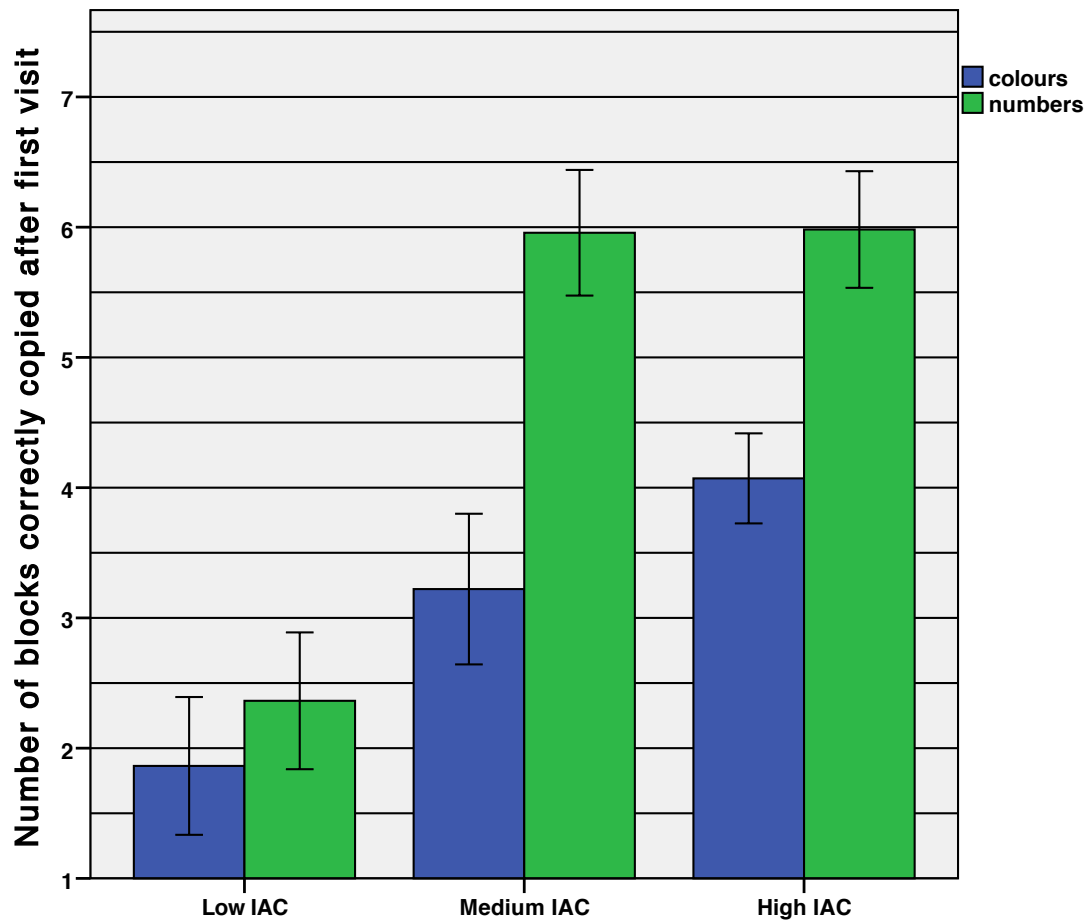


Figure 4.4: The interaction between block type and IAC for number of blocks correctly placed after the first visit to the target window. The error bars represent ± 1 standard error.

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 ($M = 22.24$, $SD = 4.47$), $F(1,30) = 44.08$, $p < .001$, $\eta^2 = 0.60$. As IAC increased from Low to Medium to High, participants took longer to complete a trial, IAC, $F(2,30) = 15.91$, $p < .001$, $\eta^2 = 0.52$. 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.

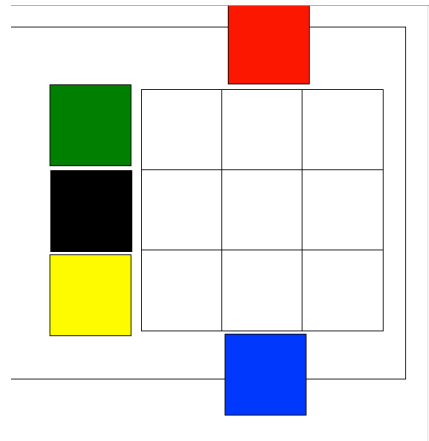


Figure 4.5: Participants placed blocks outside of the output window as ‘placeholders’.

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 4.5: 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.

4.1.4 Discussion

This study replicated the Blocks World Task with the aim to see if the effect on IAC on a copying task, as previously found when using colours, can be extended to copying numbers, and if a deeper encoding of numbers into memory makes people more accurate.

The main findings are:

- Numbers are easier to copy than colours, but only if IAC is increased
- Increases in IAC make people adopt memory-intensive strategies
- Increases in IAC increases errors

The effect of block type on task strategies and performance

Changing the block type from colours to numbers made it easier to memorise the blocks: people made shorter and fewer visits than when copying colours, but were still able to place more blocks. This difference in performance is only found in the Medium and High IAC conditions, which further suggests the difference between numbers and colours is due to the memorability of the information, and the interactions on most of the dependent variables further show that the difference between block types was dependent on the level of IAC.

The difference between numbers and colours fit well with the distinction in Baddeley's (1986) seminal model of working memory between processing visuo-spatial and verbal information. As numbers can be rehearsed and therefore refreshed in working memory, they are likely easier to memorise. After the study had ended, several participants in this study explained they used some sort of rehearsal and tried to say the numbers out loud in their head. For colours, some indicated they tried to either capture a mental picture of the colours in their mind, while others rehearsed the word for each colour (e.g. 'red' and 'blue') but they felt they were able to remember fewer items using this strategy. This is consistent with Chincotta, Underwood, Ghani, Papadopoulou, & Wresinski (1999)'s finding that people have a greater memory span for Arabic numerals (e.g. 1, 2, 3) than corresponding words (e.g. one, two, three).

Though an increase in IAC made people rely on memory and people seemed to memorise numbers more easily than colours, an increase in IAC did not make people more accurate in copying numbers as in Gray & Fu (2004), where people had to memorise VCR programming information. The error rate for numbers was however already quite low in all conditions, and moreover people were not explicitly instructed or trained to memorise the information before each trial as in Gray & Fu (2004).

The effect of IAC on task strategies and performance

The effect of IAC on people's cognitive strategies, as found in previous studies, is replicated in this study: as the cost of accessing the target window increased, people increasingly relied on memory (Gray et al., 2006; Morgan et al., 2009; Waldron, 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 strategy worsened their global performance as they took longer to complete the task and placed more incorrect blocks throughout the trials.

An increase in IAC also made people rely more on memory when copying numbers. When a mask was added over the target window, people visited the target window less often and they placed more blocks after the first visit. In both the Medium IAC and High IAC conditions, people on average placed around six blocks correctly after the first visit, after which they needed

one more visit to look at the last three blocks. The average number of items people can memorise is 7 ± 2 items (Miller, 1956), and potentially six blocks was the maximum that people could reasonably memorise.

In previous studies, adopting a memory-based strategy when copying text and numbers improved accuracy (Gray & Fu, 2004; Soboczenski et al., 2013). In the current study, the hypothesis that memorising numbers due to an increased IAC would make people more accurate is not supported. 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 Gray & Fu (2004) and Soboczenski et al. (2013), people typed in data using a computer keyboard.

With the BWT, it was difficult to measure visits to the target window in the same manner for all conditions. For the Low IAC conditions, eye fixations were used, whereas for the Medium and High IAC conditions, uncoverings of the mask were used. This introduced several problems. 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, 2007). For the Medium and High IAC conditions, an interaction was required and a conscious decision had to be made to reveal the target in these conditions. It would therefore seem likely that uncoverings more reliably measure visits to the target window. However, the uncoverings for the Medium IAC conditions were suspiciously short. Playing back the screen recordings suggests participants often accidentally uncovered the window, and it is therefore less clear if these are a reliable measure of actual visits.

Summary

The purpose of this study was to study the effect of IAC on strategy, speed and accuracy when copying from one source, and investigate if the effect of IAC, as found in previous experimental studies, would extend to copying numbers. In order to be able to compare results of this study with previous IAC studies the same task paradigm of the Blocks World Task was used.

The hypothesis that people would need fewer switches and would perform better for numbers than colours is supported, indicating that numbers are easier to memorise. The hypothesis that memorising numbers due to an increased IAC would make people perform better is not supported. 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. Furthermore, in Study 1 people saved up data to enter a lot in one sequence, and errors increase as people have to copy more (Healy, Kole, Buck-Gengle, & Bourne, 2004). Potentially the setup of this experiment was not suitable to reliably measure an increase in errors.

The information was copied from one source, and for each participant the IAC was the same throughout the experiment: it was either low, medium, or high. Therefore, while using the BWT task paradigm was appropriate in comparing its results with previous IAC studies, it did not resemble the data entry tasks observed in the financial office setting of Study 1.

Considering these limitations, it was decided after this study to not continue with the same task paradigm, but certain learning points can be taken away from this study.

First, the entry method matters. An entry method that is sufficiently slow may make people accurate, regardless of the level of IAC. Second, the effect of IAC generalises, but it is not clear how the results of this experimental paradigm can be translated in suggestions in how people dealing with differing levels of IAC can best be supported in their work. In contrast to this study, where information came from one source, and always had the same level of IAC for each participant, Study 1 showed that for entering expenses, people have to enter different types of information, these do not all come from the same source, and each source can have a different access cost.

Third, the measured short visits in the Medium IAC conditions may have confounded the results. In future studies, the setup of the experiment should be designed so that participants do not accidentally access the source when they do not intend to.

Lastly, for each dependent variable, the same consistent measure should be used to compare across conditions. This could be data provided by an eyetracker, or interaction measures such as mouse clicks, key presses or interkey intervals.

The next study will try to more closely simulate the expenses task observed in the financial office setting of Study 1 and 2. People will have to enter and collect information from multiple sources with different IACs. The aim is to see how these differences in IAC influence people's switching strategies between entering and looking up data when copying from multiple sources.

4.2 Study 4: Copying data from multiple sources

4.2.1 Introduction

Study 3 showed that as it takes longer to access an information source needed for a copying task, people spend a longer time looking at that source. They try to group and memorise as much information, in order to minimise the number of revisits to this source. Though people copied over more items after one visit using this strategy, they also made more errors overall. In the experiment, all information was to be found on a single source. People may have tried to memorise too much items in one visit, and upon entry relied on incorrectly memorised information in the head.

Data entry in office workplaces often does not involve a single source, but information can be spread over various sources. These sources are often not all equally easy or hard to access. How do people prioritise how they enter and look up information from these different sources? Do they enter the easy items first? And will they still try to group and memorise as many items? Or will they look up and enter items from one source at a time? In order to support people in looking up information for data entry work, it is important to first understand how people currently manage these information tasks.

In the contextual inquiry of Study 2, participants started data entry by first collecting all physical sources first and placing these on their desk. They then entered this information, which was nearby, first. They postponed accessing other sources until they needed to enter it during the task. If information took too long to access, they would set aside the task and continue with other tasks. One of the factors that influenced the different strategies appeared to be the time cost to retrieve the information.

This study tests the assumption that observed strategies from Study 2 are influenced by the time costs to access information sources. Whilst prior studies have demonstrated that various tasks can involve the use of multiple information sources (Cangiano & Hollan, 2009; Murphy, Chen, & Cossutta, 2016; Su, Brdiczka, & Begole, 2013), there are limited studies that measure how people use these sources, and to what extent the time cost to access a source influences these decisions.

The following questions will be addressed:

- RQ1. *How does IAC affect the number of visits to look up information from multiple sources with different IACs?*
- RQ2. *How does IAC affect the order of visits to look up information from multiple sources with different IACs?*
- RQ3. *Do the number and order of visits affect data entry accuracy and speed?*

The soft constraints hypothesis predicts that people choose and adapt their task strategies in order to minimise time (Gray et al., 2006). Study 3 found that the longer it takes to access information, the more items people try to memorise in one visit. Based on these findings, the following hypothesis is made:

H1. As IAC increases, people will try to group and memorise multiple items to minimise visits.

In O'Hara & Payne (1998)'s study on the effect of IAC on problem solving tasks, participants in Low-IAC and High-IAC conditions initially performed the same type of strategies. However, over time participants in a High-IAC condition learnt more efficient strategies, whilst participants in a Low-IAC condition continued to use the same strategy. Prior work has also shown that people who are exposed to High-IAC situations will adopt and continue to use a certain strategy, even in situations where the cost to access information is no longer high (Patrick, Morgan, Tiley, Smy, & Seeby, 2014). It is therefore expected that once participants learn it is more efficient to group high-IAC items, they may adopt this strategy for low-IAC items as well. In Study 2 of this thesis, people tried to enter items that were nearby in the environment first, and postponed looking up other information until later. Based on these findings, the following hypothesis is made:

H2. As the experiment progresses and people become aware how costly it is to access certain sources, they will learn and choose to enter all the Low-IAC items first, in a batch, and then the High-IAC items second, also in a batch, rather than looking up each item as they need it.

Lastly, if people group items and spend a longer time looking up information, this means they will be interrupted from their data entry task for a longer time. The longer people are interrupted from a primary task, the slower they are to resume that task after the interruption (Monk, Trafton, & Boehm-Davis, 2008), and the more likely they are to make resumption errors (e.g. Brumby, Cox, Back, & Gould, 2013). Based on this, the following hypothesis is made:

H3. As IAC increases, people will be slower and make more errors.

4.2.2 Method

Participants

Thirty-three participants (12 male) with a mean age of 26 years ($SD=8$) took part in the experiment. They were recruited from a university subject pool and received £4 for their participation.

Task

The experimental data entry task was framed around the expenses task from Studies 1 and 2. For this task, the user has to complete a number of data entries regarding incurred expenses in order to claim back the money. They enter this into a claim form, which looks similar to a spreadsheet. Users can choose to either fill in multiple expenses in one sheet, in which each row corresponds to one expense, or have separate spreadsheets for each expense. The current experiment will use the scenario where users enter multiple expenses in one sheet.

The experimental task can be viewed in Figure 4.6. Participants had to switch between five computer windows on the computer screen. One window contained the data entry form, in which the data for two expenses had to be entered (see Figure 4.6a). For each expense, they had to enter a financial amount to specify the expense that was made, and

Claim form Amount Account code Amount Account code

Claim form

Claim number	Amount	Account code
1	Am1	Acc1
2	Am2	Acc2

Submit

(a)

Claim form Amount Account code Amount Account code

Amount

Claim number	Amount
1	11.49

(b)

Claim form Amount Account code Amount Account code

Claim form

Claim number	Amount	Account code
1	11.49	
2		

Submit

(c)

Figure 4.6: Participants had to enter four data items in a data entry form (a). Each item had to be retrieved from a separate window (b) and entered into the form (c).

an account code to determine from which account to claim back the expense. The four data items had to be retrieved from four separate windows (see Figure 4.6b), and entered in the corresponding fields on the data entry form (see Figure 4.6c). A window was revealed 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 one window covered the full screen.

Materials

The data items were designed to look similar to the numbers seen in Studies 1 and 2. People usually worked with the same 20 or 30 account codes. They were aware they worked with the same set, but still had to look them up each time they needed them. The codes had a length of six digits, and the string was random with no pattern.

An experimental session consisted of 50 trials, divided into 5 blocks of 10 trials. Each trial consisted of four data entries, so in total participants made 200 entries. 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. There were two practice trials before the experiment began. The numbers used in the practice trials were not used for the experimental trials. The data items always had the same length: the amounts consisted of two digits on the integer part and two digits on the fraction part (e.g. 11.94), and the account codes were six digits (e.g. 654273).

The experiment was conducted in a web browser, on a desktop computer with a 24-inch monitor with a resolution of 2048 x 1152 pixels. People used a computer mouse and number pad to view and enter the numbers, and the option to copy and paste information was disabled. The cursor stayed in the same cell of the data entry form if the participant switched to other windows and back. The task interface was developed in HTML, CSS, Javascript and PHP. The browser window was maximised to avoid distractions. All mouse clicks, keypresses and their timestamps were recorded using Javascript.

Design

The experiment was a between-subjects design with one independent variable, the IAC of the information source windows. In the control condition, there were no delays in opening the windows. In the other two conditions, there was a two-second delay before two of the four windows were revealed. In the High-Account condition, there was a delay for viewing the two account codes, whereas in the High-Amount condition there was a delay for viewing the two amounts. All window switches and key presses were logged to determine the order and grouping of visits and entries. Other dependent variables were trial completion time, number of visits to information sources, and data entry errors.

Procedure

The experiment took place in a closed quiet room. Participants were welcomed and briefed verbally about the study. They were instructed to complete the task as fast and accurately as possible, but other than this instruction, it was emphasised they were free to use whichever strategy they wished. Examples of possible strategies were demonstrated: they could enter it per row, per column, and could go back to the information sources as often as they wished. After reading an information sheet and signing a consent form, participants completed two practice trials. After this they completed 50 experimental trials. After every 10 trials, there was the opportunity to take a break. A prompt appeared on the computer screen and time stopped. Participants were free to choose the length of each break, and continued the experiment by pressing a 'Next' button on the screen. The experiment took approximately 30 minutes.

Pilot study

Two pilot studies were conducted with colleagues of the researcher to test the experimental design. In particular, the pilot studies aimed to see if the length of the experiment was long enough for participants to learn and develop strategies, but not too long to tire the participant.

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 tabs had an increased IAC. As a result, he did not adapt his strategies according to anticipated IACs and entered the numbers row by row. The second participant mentioned that the increased IACs 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 real experiments, the breaks were reduced to happen after every 10 trials. In addition, the names of information sources with an increased IAC were underlined in the horizontal menu. This visual feature was added to help users see which windows had an increased time cost to open.

Analysis

In order to analyse if there was a significant difference in ordering and grouping strategies, a bottom-up approach was taken to manually categorise people's strategies into different strategy categories.

All window switches and key presses were recorded and saved in a spreadsheet. This spreadsheet showed the order of actions for each participant and trial. There were eight different possible actions: viewing the first amount (V-Am1), viewing the second amount (V-Am2), viewing the first account code (V-Acc1), viewing the second account code (V-Acc2), 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).

There were 33 participants who each completed 50 trials, so in total there were 1650 trials to be categorised. Each instance was first categorised according to the order in which items were entered: per row, column, or another order. This was to determine whether people grouped and entered data per type of data item (Amount and Account code), if they entered data in the sequence in which it was presented on the form (Claim1 and Claim2), or if they grouped according to the IAC (Low-IAC and High-IAC). The instances were then categorised according to whether the participants grouped multiple data items in one visit, or if data items were viewed and then entered one by one.

The occurrences of each strategy type was calculated for each condition and summarised in a frequency table. One-way ANOVAs were used to determine if there were significant differences between conditions in frequency of strategies, number of visits to the information sources, trial completion time, and errors.

4.2.3 Results

Ordering strategies

Condition	Row	Column	Other	Total
High-Account	33.27% (183)	66.55% (366)	0.18% (1)	100% (550)
High-Amount	64.73% (356)	34.18% (188)	1.09% (6)	100% (550)
Control	64.18% (353)	34.00% (187)	1.82% (10)	100% (550)
Total	54.06% (892)	44.91% (741)	1.03% (17)	100% (1650)

Table 4.3: The rates (and frequencies) of ordering strategies per condition.

Table 4.3 shows the number of trials where people grouped per row, per column and where they had another order. The Row strategy meant that the participant first entered Amount1 and Account1, and then Amount2 and Account2, or the other way around (Figure 4.7a). The Column strategy meant that the participant first entered Amount1 and Amount2, and then Account1 and Account2, or the other way around (Figure 4.7b).

(a)

(b)

Figure 4.7: The order strategies. (a) If people entered Amount 1 and then Account1 (or Amount2 and then Account2), this was categorised as a 'Row' strategy. (b) If people entered Amount 1 and then Amount2 (or Account1 and then Account2), this was categorised as a 'Column' strategy.

There was a significant difference in choice of strategy, $F(2,30) = 3.32$, $p = 0.04$. People most often entered data in the order in which they were presented, per row. In the High-Amount and Control condition, this strategy was chosen on 64.73% and 64.18% of the trials, respectively. In the High-Account condition, the strategy was chosen 33.27% of the time, and people more often entered items per column. In the 741 trials where people adopted a Column strategy, they predominantly entered the amounts first, even in the High-Amounts condition. There were only fifteen trials where account codes were entered first, and these were by the same person in the High-Amounts condition.

Grouping strategies

Condition	Grouped	One-by-one	Total
High-Account	37.52% (206)	62.48% (343)	100 (549)
High-Amount	55.70%(303)	44.30% (241)	100 (544)
Control	67.78% (366)	32.22% (174)	100 (540)
Total	53.58% (875)	46.42% (758)	100 (1633)

Table 4.4: The rates (and frequencies) of grouping strategies per condition.

There was no significant difference between conditions on total number of visits to information sources, $F(2,30) = 2.71$, $p = 0.08$. On average people needed five visits to the sources per trial. To see if people tried to reduce number of visits by viewing multiple data items before entering them, strategies were categorised into so-called 'Grouped' and 'One-by-one' strategies. If participants grouped visits and made multiple visits before entering them, they were placed in the 'Grouped' category. If they visited and then entered each item one by one, they were placed in the 'One-by-one' category.

Overall, there was an even distribution of trials where participants tried to group items (53.58%) and trials where they entered them one by one (46.42%) and this did not differ

Condition	Grouped	One-by-one	Total
High-Account	24.32% (89)	75.68% (277)	100 (366)
High-Amount	65.43% (123)	34.57% (65)	100 (188)
Control	50.80% (95)	49.20% (92)	100 (187)
Total	41.43% (307)	58.57% (434)	100 (741)

Table 4.5: The rates (and frequencies) of grouping strategies for trials where people entered items in column order.

Condition	Grouped	One-by-one	Total
High-Account	26.78% (49)	73.22% (134)	100 (183)
High-Amount	24.44% (87)	75.56% (269)	100 (356)
Control	13.88% (49)	86.12% (304)	100 (353)
Total	20.74% (185)	79.26% (707)	100 (892)

Table 4.6: The rates (and frequencies) of grouping strategies for trials where people entered items in row order.

significantly between conditions, $F(2,30) = 0.39$, $p = 0.68$. (see Table 4.4). However, looking at the trials where people did group per column, in Table 4.5 it can be seen that in the High-Amount condition, people more often tried to group items. A chi-square test showed that looking at trials where people ordered per column, grouping strategy was significantly associated with condition, $\chi^2(2) = 95.55$, $p < 0.05$.

Finally, for the trials where people decided to enter items row by row, they most often entered them one by one and did not group them, as can be seen in Table 4.6:

Condition	Row		Column			Total
	First row (a)	First & Second row (b)	Amounts (c)	Amounts & Accounts (d)	Other	
High-Account	34% (48)	4% (6)	57% (80)	2% (3)	3% (4)	100 (141)
High-Amount	20.99% (44)	16.57% (35)	49.72% (104)	9.39% (20)	3.31% (7)	100 (210)
Control	11.2% (16)	21.6% (32)	54.4 (81)	12% (18)	0.8% (1)	100 (148)
Total	21.18% (190)	15.02% (73)	52.96% (265)	8.37% (41)	2.46% (12)	100 (499)

Table 4.7: The most common grouping strategy was to chunk the items into three groups. The strategies are shown graphically in Figure 4.8.

Figure 4.8 illustrates the most common Grouping strategies. People most often chunked items into three groups (Strategies (a) and (c) in Figure 4.8). For example, they first visited Amount1 and Amount2, and then entered Amount1 and Amount2. They then viewed Account1 and entered it, before they viewed Account2 and entered it. Table 4.7 shows the frequency with which these strategies were chosen per condition.

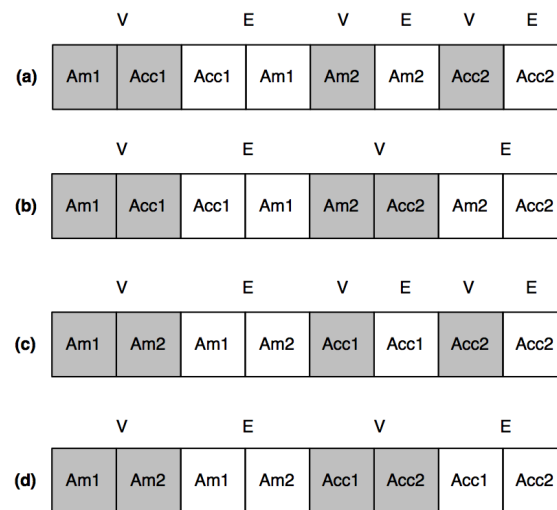


Figure 4.8: The sequence of the most common grouping strategies. V = visit to the data source, E = entry of the data item. For example, in Strategy (a) a participant first viewed Amount1 and Account1, and then entered Amount1 and Account1. He/she then viewed Amount2 and entered it, and then viewed Account2 and entered it.

Condition	Error rate	Trial completion time (s)
	M (SD)	M (SD)
High-Account	4.82 (3.56)	29.23 (7.44)
High-Amount	3.68 (2.73)	28.88 (6.92)
Control	7.75 (9.81)	29.99 (10.94)

Table 4.8: The means (and standard deviations) of error rate and trial completion time the three conditions. The rates are percentages calculated as the ratio of occurrences to the number of opportunities.

Performance measures: errors and speed

There was no significant difference in the time it took to complete a trial, $F(2, 30) = 0.05$, $p = 0.9$). On average, participants took about 29 seconds per trial across conditions.

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. In total, 101 errors were made in the High-Account condition, 82 in the High-Amount condition, and 160 errors in the Control condition. The mean error rates were 4.82%, 3.68 %, and 7.75%, respectively. Despite the differences in means, due to high sample variance no significant difference was found in errors, $F(2,30) = 1.25$, $p = 0.3$.

An alternative error rate was calculated based on the number of erroneous trials. If a participant made a mistake on all four entries of a trial, this rate would consider it as one error, rather than four errors. Looking at the number of trials on which an error was made, there was no significant difference either, $F(2, 30) = 0.96$, $p = 0.4$. In total, there were 87 erroneous trials in the High-Account condition, 41 trials in the High-Amount condition, and 65 trials in the Control condition. In other words, errors were made on 14.91%, 9.76% and 22.91% of the trials, respectively.

Qualitative findings

Participants mentioned adapting 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, several 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 at a time. As a result, most participants ended up viewing and entering each account code one by one.

Several 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 had entered before, they would use the 'Grouped' strategy: they viewed another window and item before returning to the entry form. Furthermore, some participants started a trial by not viewing any other windows, but re-entering the data items of the previous trial in the data entry form. They would then visit the information source windows to check if the items were the same, or if they needed to change some of the items.

4.2.4 Discussion

The aim of this study was to investigate the effect of IAC on how people manage subtasks to look up information from multiple information sources with varying IACs. IAC influenced the order in which people entered data and whether and how they chunked data items.

Strategies

IAC had an effect on people's ordering strategies, but only in the condition where the IAC was increased for accessing account codes. People first looked up and entered the low IAC items, the amounts, before they entered the high IAC items, the account codes. However, if the cost to access the amounts was increased, people entered items row by row, as participants did in the control condition.

IAC also had an effect on how often people switched between visiting and entering information. In the conditions where the amounts were low, people often visited multiple items before returning to the entry interface. Though there was no difference in errors in the current study, it can be imagined this strategy is more risky as people are away for longer so resuming a task is more difficult, and they are storing more information in memory, so people may misremember information.

In the condition where the IAC for amounts was high, this strategy was not used as frequently. IAC thus does not have to be bad if the information is easy enough to briefly hold in memory (condition amounts high).

IAC made people change strategies even if only two of the four information sources had an increased IAC. This is in accordance with (?), who showed that a more memory-based strategy can be trained for future situations. When people were exposed to an interface with an increased access cost, they adopted a memory-based strategy and retained this strategy, even when they then interacted with an interface with lower access costs.

In previous studies, people would postpone or not address information needs if there was a high time cost associated with it (Sohn, Li, Griswold, & Hollan, 2008). This was also found in Study 2. People initially tried to do tasks as they needed it, but if it took too long, they would put it aside.

In the current study, this was also found but the cost to memorise an item overrode the cost to access it: people would first enter the amounts, even if the cost to access them was higher. If the cost was high, they would try and memorise them both before entering them. This shows people try to minimise interleaving between viewing and entering information.

People made more visits to information sources in the control condition. However, as opposed to Study 3, this did not improve accuracy: on the contrary, people made significantly more errors in the control condition than when one of the sources had an increased IAC. Differences in IAC can thus help people be more efficient and effective in scheduling subtasks of looking up information, and reduce the number of visits they need to make.

Contributions

Demonstration that:

- IAC also affects people's strategies of switching between entering and looking at the target source when multiple sources are involved.
- IAC affects how people manage subtasks of looking up information in a data entry task.
- Depending on users' experience and awareness of how costly it is to access information, they optimise costs by processing all the low-IAC tasks first, in a batch, and then the high-IAC tasks second, also in a batch, rather than switching between the two different types of task.

Conclusion

This study investigated the effect of IAC on looking up information from multiple sources, and showed that, if all information has the same IAC it is better if this is low, differences in IAC between sources can make people schedule their subtasks more efficiently and effectively. People needed less visits and made fewer errors. However, for this study people only had one data entry task at a time to focus on, and the sources only contained information for this single task. In Study 1 and 2, people often batched and saved up multiple data entry tasks, and sources included more than just the information for one task: for example, if people had to look up an account code, this was often retrieved from a spreadsheet with all account codes. They thus not only had to manage subtasks of one data entry task, but had to coordinate multiple data entry tasks. In order to save visits, people may therefore look up information for several tasks and interleave. One of the participants from Study 1 shared that this did occur and said she and colleagues had to be very careful to not input information in the wrong form. If people group information subtasks per IAC for a single task, would they also group these subtasks per IAC when dealing with multiple tasks? In order to answer this question, a follow-up study was run to see the effect of IAC on interleaving behaviour between two claim forms.

4.3 Study 5: Interleaving between data entry tasks

4.3.1 Introduction

Study 4 showed that people group subtasks of looking up information according to the information's IAC: if all information sources were equally easy to access, participants looked up and entered information for a task in sequential order. However, if some sources were

more costly to retrieve, people first retrieved and entered all low IAC items, before they retrieved high IAC items. If people deal with multiple data entry tasks, would this also make people more likely to look up and enter low IAC items across tasks, rather than completing one task at a time? In Study 2, participants saved up data entry tasks to do in one session and occasionally had multiple data entry tasks open at the same time. Multi-tasking between data entry tasks is a common occurrence, but this strategy can be prone to errors. The aim of the current follow-up study is to see if differences in IAC between information sources makes people more likely to interleave between data entry tasks.

The following question will be addressed:

What is the influence of IAC on interleaving between data entry tasks?

4.3.2 Method

Participants

Thirty-two participants (seven male) with a mean age of 25 years ($SD=8$) 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 two aspects. Instead of filling in one data entry form per trial, people filled in two forms per trial. Furthermore, instead of four sources, participants in the current study had to retrieve information from two sources: one source to retrieve amounts, and a second source to retrieve account codes. The aim of this follow-up study was to investigate if differences in IAC of the two sources makes people more likely to interleave between data entry tasks.

Design

The experiment was a between-subjects design with the IAC of the information source windows as the independent variable. In the control condition, there were no delays in opening the windows. In the second condition, there was a delay for opening the window that contained the account codes. In the third condition there was a delay for viewing the amounts. All window switches and key presses were logged to determine the order in which participants accessed the windows and entered the data items. In addition, the number of errors and trial completion time were measured.

Procedure

The experimental setup was similar to Study 4. For each experimental trial, participants had to complete two forms with two entries each, an account code and an amount. They were instructed to complete it as fast and accurately as possible. It was explained that they could use any strategy they wanted, but that they had to complete both forms, before continuing to the next trial. The experiment took approximately 30 minutes.

(a)

(b)

(c)

(d)

(e)

Figure 4.9: Participants had to enter two data entry forms per trial, each containing two items. (a). The amounts for both forms had to be retrieved from one window (b) and entered into the form (c), while the account codes had to be retrieved from a second window (d) and entered (e).

Condition	Interleaving rate	Error rate	Trial completion time (s)
	M (SD)	M (SD)	M (SD)
High-Accounts	66.6 (3.39)	6.5 (1.04)	33.96 (5.08)
High-Amounts	85.4 (2.17)	2.9 (0.25)	33.67 (8.68)
Control	40.88 (4.03)	4.9 (0.45)	27.66 (3.40)

Table 4.9: Results of the dependent variables for the three conditions. The rates are percentages calculated as the ratio of occurrences to the number of opportunities.

4.3.3 Results

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 removed from the data. Data of the remaining 29 participants was taken into the data analysis.

Table ?? 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.

Condition	Interleaving	No interleaving	Total
High-Account	66.6% (333)	33.4% (167)	100% (500)
High-Amount	85.4% 427	14.6% (73)	100% (500)
Control	40.88% (184)	59.12% (266)	100% (450)
Total	65.1% (944)	34.9% (506)	100% (1450)

Table 4.10: The number of trials in which participants interleaved for the three conditions.

Interleaving strategies

Participants' data entry strategies were categorised in one of two categories: they either interleaved between the two data entry forms (e.g. entering amount on Form1, followed by entering the amount in Form 2), or they filled the forms in sequential order (entering amount and account code on Form1, followed by entering the amount and account code on Form2).

Table 4.10 shows the frequency of strategies per condition. In total, the strategy was chosen on 333 trials in the High-Account condition, 427 trials in the High-Amount condition, and 184 trials in the Control condition. In other words, the strategy was chosen on 66.6%, 85.4% and 40.88% of the trials, respectively. There was a significant difference in interleaving behaviour, $F(2,30) = 4.41$, $p = .02$. A post-hoc comparison showed people interleaved significantly more often in the High-IAC conditions than the Control condition ($U = 138$, $p = 0.02$). There was no significant difference between the High-Account and High-Amount condition ($U = 32$, $p = 0.18$).

Speed and errors

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.

In total, 130 errors were made in the High-Account condition, 58 in the High-Amount condition, and 88 errors in the Control condition. The mean error rate was 6.5%, 2.9% and 4.9%, respectively. Looking at the total number of data entries, there was no significant difference in errors, $\chi^2(2, N = 29) = 0.95$, $p = 0.62$.

An alternative error rate was calculated based on the number of erroneous trials. If a participant made a mistake on all four entries of a trial, this rate would consider it as one error, rather than four errors. Looking at the number of trials on which an error was made, there was no significant difference either, $\chi^2(2, N = 29) = 0.80$, $p = 0.67$. In total, there were 87 erroneous trials in the High-Account condition, 41 trials in the High-Amount condition, and 65 trials in the Control condition. In other words, errors were made on 17.4%, 8.2% and 14.4% of the trials, respectively.

There was a significant difference in the time it took participants to complete a trial, Kruskal-Wallis chi-squared = 7.087, $df = 2$, p -value = 0.02891. Mann-Whitney U tests show participants took longer in the High-IAC conditions than the control condition ($U = 145$, $p = 0.008$). There is no significant difference between the High-Account and High-Amount conditions ($U = 57$, $p = 0.63$).

4.3.4 Discussion

The aim of this study was to see if an increase in IAC makes people interleave more between data entry tasks. In contrast with Back, Cox, & Brumby (2012), who found that an increase in IAC made people less likely to interleave between two data entry tasks, participants in the current experiment interleaved more as IAC increased.

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. However, in many situations, such as the office setting of Chapter 3, information is not in one location, but different information sources have to be consulted for different types of information. For an expenses task, amounts and account codes are not on one sheet, but people have to consult one spreadsheet for account codes, and another source to retrieve the amounts. The current study looked at people's interleaving behaviour when retrieving items from multiple sources. If there were no delays in accessing these sources, participants completed one task at a time on 59% of the trials. If there was a delay in accessing either one of these sources, people tried to enter all information from this source after one visit, so they did not have to open it again. They chunked the data items associated with one source, rather than task. They first entered either Amount1 for Form1, Amount2 for Form2, and then the accounts, or first entered Account1 for Form1, Account2 for Form2, and then the amounts.

Whereas in Study 4 people became more accurate by chunking data items according to IAC, there no longer was a difference in errors in the current study. Chunking by data items in this set-up meant people interleaved between tasks and started a second task before completing the first task, a strategy which can be prone to errors (). People may forget steps, or enter correct information in the wrong fields.

It can be argued that the design of the materials encouraged participants to always group per source, regardless of the condition. However, in the Control condition there was an almost even distribution of strategies, and participants interleaved 40% of the trials. The majority of the time participants still chose to complete one task at a time.

4.3.5 Conclusion

People have to regularly switch between looking up information for a data entry task and entering it. The three studies described in this chapter showed how strategies to look up and enter information are influenced by the time cost to access information sources. It also showed that certain strategies are more accurate or efficient than others. The main effect of an increased IAC is that people try to minimise (re)visits. If the time to access a source increases, people will try to copy over more information after one visit. If they do not memorise it well, errors increase (Study 3). If information is spread across different sources and the IAC differs between these sources, people group visits and first look up and enter low IAC items, before entering high IAC items. This not only made them more efficient, but it also reduced errors (Study 4). However, if they have to manage multiple data entry tasks, this strategy means that they will interleave between tasks (Study 5).

These results are partly in line with observations from the first two studies. Whereas people would look at the physical receipts while typing it in, they would hold other items in memory and barely used tools to offload memory. They would first enter all items on the physical receipts. For digital information however, they would look it up as they

needed it, even if IAC differed between these sources, and it could sometimes take a while before they had retrieved the information.

It seems that it is better to be able to reduce IAC and have task information ready at hand, so people do not need to switch back and forth to a source that takes time to access. There are solutions, such as increased screen space, multiple screens, or having a physical copy and placing it nearby. It was interesting to see in the first two studies that people did decrease IAC for physical items, but not for digital ones. People had a second screen, but used their primary screen to look up information because they perceived it as quicker. In this case, the cost to decrease IAC by placing information on a second screen, outweighed the cost to look it up, hold it in memory, and go back to the primary screen. However, they often did not know the associated time cost to access it, so could be away from the screen for a longer time than anticipated.

There is a need to better support people in decreasing IAC without tasking them with the added responsibility of re-arranging different tasks, information sources, devices and screens. They should be able to do their job and have task information at hand more seamlessly.

References

- Back, J., Brumby, D. P., & Cox, A. L. (2010). Locked-out : Investigating the Effectiveness of System Lockouts to Reduce Errors in Routine Tasks. In *CHI 2010: 28th ACM Conference on Human Factors in Computing Systems*.
- Back, J., & Cox, A. L. (2013). Artifacts for programmable devices: The good, the bad and the ugly. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1731–1736). New York, NY, USA: ACM. doi: 10.1145/2468356.2468666
- 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. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Barchard, K. A., & Pace, L. A. (2011). Preventing human error: The impact of data entry methods on data accuracy and statistical results. *Computers in Human Behavior*, 27(5), 1834–1839.
- Barchard, K. A., & Verenikina, Y. (2013). Improving data accuracy: Selecting the best data checking technique. *Computers in Human Behavior*, 29(5), 1917–1922. doi: 10.1016/j.chb.2013.02.021
- Bardram, J. E., Bunde-Pedersen, J., & Soegaard, M. (2006). Support for activity-based computing in a personal computing operating system. In *Chi '06* (pp. 211–220). Retrieved from <http://doi.acm.org/10.1145/1124772.1124805> doi: <http://doi.acm.org/10.1145/1124772.1124805>
- Beyer, H., & Holtzblatt, K. (1998). *Contextual design: Defining customer-centered systems*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Bi, X., & Balakrishnan, R. (2009). Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. In *Chi '09* (pp. 1005–1014). Retrieved from <http://dl.acm.org/citation.cfm?doid=1518701.1518855> \delimeter"026E30F\$npapers2://publication/doi/10.1145/1518701.1518855 doi: 10.1145/1518701.1518855
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Brumby, D., Cox, A., Back, J., & Gould, S. (2013). Recovering from an interruption: Investigating speed-accuracy tradeoffs in task resumption strategy. *Journal of Experimental Psychology: Applied*, 19(2), 95–107.
- Byrne, M. D., & Bovair, S. (1995). A Working Memory Model of a Common Procedural Error. *Cognitive Science*, 21(1), 31–61.
- 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: First international conference, hcd 2009, held as part of hci international 2009, san diego, ca, usa*,

- july 19-24, 2009 proceedings (pp. 945–954). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-642-02806-9_108 doi: 10.1007/978-3-642-02806-9_108
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum Associates,.
- Chincotta, D., Underwood, G., Ghani, K. A., Papadopoulou, E., & Wresinski, M. (1999). Memory Span for Arabic Numerals and Digit Words: Evidence for a Limited-capacity, Visuo-spatial Storage System. *The Quarterly Journal of Experimental Psychology*, 52A(2), 325–351.
- Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A Diary Study of Task Switching and Interruptions. In *Chi '04* (Vol. 6, 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>
- De Luca, A., Langheinrich, M., & Hussmann, H. (2010). Towards understanding ATM security: a field study of real world ATM use. *Proc. SOUPS 2010*, 1–10. doi: 10.1145/1837110.1837131
- 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). Fortune favors the bold (and the Italicized): effects of disfluency on educational outcomes. *Cognition*, 118(1), 111–5.
- 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'12* (pp. 1947–1956). Austin, Texas, USA.
- Furniss, D., & Blandford, A. (2006). Understanding Emergency Medical Dispatch in terms of Distributed Cognition: a case study. *Ergonomics*, 49(12 & 13), 1174 – 1203.
- Gong, Y., & Zhang, J. (2009). Effects of Information Displays for Hyperlipidemia. In G. Salvendy & M. Smith (Eds.), *Human interface and the management of information. information and interaction* (Vol. 5618, pp. 503–512). Springer Berlin Heidelberg.
- Gould, S. J. J., Cox, A. L., & Brumby, D. P. (2015). Task lockouts induce crowdworkers to switch to other activities. In *CHI 2015 Extended Abstracts, Republic of Korea, April 18 - 23, 2015* (pp. 1785–1790).
- 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, July). 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
- Green, R. a., Hripcsak, G., Salmasian, H., Lazar, E. J., Bostwick, S. B., Bakken, S. R., & Vawdrey, D. K. (2014). Intercepting Wrong-Patient Orders in a Computerized Provider Order Entry System. *Annals of emergency medicine*.

- Grudin, J. (2001). Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Chi 2001* (pp. 458–465). Retrieved from <http://dl.acm.org/citation.cfm?id=365312> doi: 10.1145/365024.365312
- Healy, A. F., Kole, J. A., Buck-Gengle, C. J., & Bourne, L. E. (2004). Effects of prolonged work on data entry speed and accuracy. *Journal of experimental psychology. Applied*, 10(3), 188–99.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.*, 7(2), 174–196.
- Hutchins, E. (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science*, 19, 265–288.
- 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, March). 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
- Katidioti, I., & Taatgen, N. A. (2013). 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. doi: 10.1177/0018720813504216
- Kawado, M., Hinotsu, S., Matsuyama, Y., Yamaguchi, T., Hashimoto, S., & Ohashi, Y. (2003). A comparison of error detection rates between the reading aloud method and the double data entry method. *Controlled Clinical Trials*, 24, 560–569.
- Kohn, L., Corrigan, J., & Donaldson, M. (2000). *To err is human: building a safer health system*. National Academy Press.
- Koppel, R., Wetterneck, T., Telles, J. L., & Karsh, B.-T. (2008). Workarounds to Barcode Medication Administration Systems: Their Occurrences, Causes, and Threats to Patient Safety. *Journal of the American Medical Informatics Association*, 15, 408–423.
- Li, S., Cox, A., Blandford, A., Cairns, P., Young, R., & Abeles, A. (2006). Further investigations into post-completion error: the effects of interruption position and duration. In *Proceeding Cognitive Science Conference 2006* (pp. 471–476).
- Li, S. Y. W., Blandford, A., Cairns, P., & Young, R. M. (2008). The effect of interruptions on postcompletion and other procedural errors: an account based on the activation-based goal memory model. *Journal of Experimental Psychology: Applied*, 14(4), 314–328.
- Li, S. Y. W., Cox, A., Or, C., & Blandford, A. (2016). Effects of monetary reward and punishment on information checking behaviour. *Applied Ergonomics*, 53, 258–266.
- Lin, C.-J., & Wu, C. (2011). Factors affecting numerical typing performance of young adults in a hear-and-type task. *Ergonomics*, 54(12), 1159–1174.
- Lin, C.-J., & Wu, C. (2013). Reactions, accuracy and response complexity of numerical typing on touch screens. *Ergonomics*, 56(5), 818–31.
- Lin, C.-J., & Wu, C. (2014). Predicting numerical data entry errors by classifying EEG signals with linear discriminant analysis. *Behaviour & Information Technology*(November 2014), 1–12.
- MacKenzie, I. S., & Soukoreff, R. W. (2002). Text Entry for Mobile Computing: Models and Methods, Theory and Practice. *Human-Computer Interaction*, 17(May 2015), 147–198.

- Mangen, A., & Velay, J.-L. (2010). Digitizing literacy: reflections on the haptics of writing. In M. H. Zadeh (Ed.), *Advances in haptics* (pp. 385–402). InTech.
- Mark, G., Gonzalez, V. M., & Harris, J. (2005). No task left behind? In *Chi '05* (pp. 321–330). Portland, Oregon, USA. Retrieved from <http://dl.acm.org/citation.cfm?id=1054972.1055017> doi: 10.1145/1054972.1055017
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The 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, December). 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, hci 2007, Incs 4557* (pp. 115–125). Berlin: Springer-Verlag.
- Mueller, P. A., & Oppenheimer, D. M. (2014). The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological science*, 25(6), 1159–1168.
- 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
- Nakata, T. (2014). Improving Human Reliability On Checking. In *Proceedings of International Symposium on Interaction Design and Human Factors*.
- Norman, D. A. (1993). Cognition in the Head and in the World : An Introduction to the Special Issue on Situated Action. *Cognitive Science*, 17, 1–6. doi: 10.1207/s15516709cog1701
- Norman, D. A. (2002). *The design of everyday things*. New York, NY, USA: Basic Books, Inc.
- 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. (2011). Number entry interfaces and their effects on error detection. *IFIP International Federation For Information Processing*, 178–185.
- Oladimeji, P., Thimbleby, H., & Cox, A. (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). Customer Errors in Internet Banking. In *Proceedings of Norsk Informatikkonferanse*.
- 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.

- 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., & Howes, A. (2013). *Adaptive interaction: A utility maximisation approach to understanding human interaction with technology*. Morgan Claypool.
- 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.
- 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., Tabard, A., Boyd, K., & Hollan, J. (2015). Restoring the Context of Interrupted Work with Desktop Thumbnails. In *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 2045–2050). Pasadena, USA.
- 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.
- Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological bulletin*, 99(3), 303–319.
- Sandnes, F. E. (2013). An Error Tolerant Memory Aid for Reduced Cognitive Load in Number Copying Tasks. In *Universal Access in Human-Computer Interaction. User and Context Diversity* (pp. 614–623). Springer-Verlag.
- 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
- Smith, M., Lewis, R., Howes, A., Chu, A., & Green, C. (2008). More than 8,192 ways to skin a cat: Modeling behavior in multidimensional strategy spaces. *Proceedings of the 30th Annual Conference of Cognitive Science Society*, 1441–1446.
- 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).
- Tu, H., Oladimeji, P., Wiseman, S., Thimbleby, H., Niezen, G., & Cairns, P. (2014). Employing Number-based Graphical Representations to Enhance the Effects of Visual Check on Entry Error Detection. In *International Symposium on Interaction Design and Human Factors (IDHF 2014)*. Kochi, Japan.

- 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. (2007). *The Availability and Cost of Accessing Information from the Interface: Consequences for Memory, Planning and Learning* (Unpublished doctoral dissertation). Cardiff University.
- 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.
- Wiseman, S. (2014). *Designing for Numerical Transcription Typing: Frequent Numbers Matter* (Unpublished doctoral dissertation). University College London.
- Wiseman, S., Brumby, D. P., Cox, A., & Hennessy, O. (2013). Tailoring Number Entry Interfaces To The Task of Programming Medical Infusion Pumps. In *Proceedings of the Human Factors and Ergonomics Society* (pp. 683–687).
- Wiseman, S., Cox, A. L., & Brumby, D. P. (2013). Designing Devices With the Task in Mind: Which Numbers Are Really Used in Hospitals? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 55(1), 61–74.
- Zhang, J., & Wang, H. (2009). An exploration of the relations between external representations and working memory. *PloS ONE*, 4(8), e6513.