SMARTPHONES: "EASY", "PORTABLE" AND "FAST"? THE IMPACT OF DEVICE STEREOTYPES ON COGNITION

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF SCIENCE AND ENGINEERING

2022

Madeleine S. Steeds

School of Computer Science

Contents

Word Count: 50,905

List of Tables

List of Figures

Glossary

Cognition: The mental processes which contribute to our awareness of the world around us; an umbrella terms for higher order behaviours such as memory, attention, and decision- making.

Devices: Computing hardware which humans interact with to facilitate various tasks (e.g. a smartphone).

Stereotypes: A stereotype is made of two key elements: a *belief* and the *group* about which that belief is held [72]; while often related to negative beliefs, they can also be positive or neutral.

Associations: Constructs an individual relates to an object; a stereotype is an association but an association is not necessarily a stereotype.

Memory: The ability to store and retrieve information from one's own mind; an umbrella term for storing and retrieving different types of information.

Encoding: The memorisation of information. In memory experiments the learning phase may be referred to as the 'encoding phase', and participants may be said to 'encode' information.

Retrieval: The act of remembering information. In memory experiments, the 'retrieval phase' is when participants are asked to remember the target information, and participants may be said to 'retrieve' information.

Abstract

Stereotyping in Human-Computer Interaction (HCI) has traditionally been applied to the end-user, to exemplify the types of people who may use a given technology or device. However, recent research has indicated that devices themselves may be stereotyped by users, but this prior research has been limited to just one stereotype model. Research from psychology suggests that stereotypes can be used as cognitive shortcuts to preserve cognitive resources, and allow other information to be processed faster. A study has further suggested that perceptions of the technological device being used can cause cognitive processes to be biased towards behaviour in keeping with that stereotypical view. Research in this domain is limited though, and there is little investigation into the extent to which devices are stereotyped, as well as the extent to which these stereotypes can influence human cognition. This thesis therefore aims to investigate the stereotypes and associations held about four everyday computing devices, and how these associations influence human cognition. To do this, the thesis first investigates existing literature, comparing the effect of different devices on one cognitive process (memory). A survey is then carried out investigating how different devices are used, and how those devices are stereotyped in terms of hedonic and utilitarian qualities, the stereotype content model, and free association. The results of this survey are then used to investigate the impact of two devices on two cognitive process (information processing and memory). The findings of this thesis suggest that there are common trends in the ways devices are stereotyped and perceived. Furthermore, performance in tasks assessing cognitive abilities is affected when the stimuli used in the task is associated with the device being used to complete the task. The findings of this thesis therefore expand current understanding of the stereotypical properties of everyday computing devices. They further provide evidence of these stereotypical views influencing human cognition in cognitive domains not previously explored in this fashion.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

I would like to thank Sarah Clinch and Caroline Jay for being such supportive supervisors over the course of this project. Thank you Sarah for taking me on as a PhD student despite my lack of computer science background, and for taking the time to teach me to code and use version control software. Thank you Caroline for your support through the writing process, and for giving me pep talks when I was stressing myself out. A further thank you to other staff in the department, who have been very supportive and helpful throughout my PhD.

Thank you to my mum and dad for supporting me throughout my studies and my PhD. I promise I won't go back to being a student again! Thank you to my sister Vivienne who has been very supportive, and who let me borrow her pet rabbits Toffee and Mouse as emotional support bunnies. Thank you Toffee and Mouse for all the cuddles!

Thank you to all my lab mates who have helped shape this project and given me tech support. In particular thank you Lukas, Jonathan, Miguel, Alex, and David. You have all been such good friends and I would not have made it through this project without your support. Thank you to my other friends in and out of the department, Ryan, Mary, Martina, Ruba, Julia, Jake, and Michael, to name just a few.

Chapter 1

Introduction

As computing devices become more pervasive within society, there are concerns that they may have negative effects on a person's cognitive abilities. However, limited research has been conducted into these effects, the exact nature of these effects, and particularly into how these effects may vary between devices. Research that does investigate this has primarily been focussed on the tangible features of a given device such as its screen size or interaction method. This project instead aimed to investigate *intangible* features of the device, namely the stereotypes and associations held about the device. The research considered whether the associations and stereotypes may serve as a cue (i.e. a prompt within the environment), influencing cognition. As cognition is a broad research domain, the present project focussed on two cognitive processes: information processing, and memory.

1.1 Motivation

A large number of different devices are available for purchase by consumers globally, with wealthier countries having greater access to household computing devices than less economically developed countries. For example, in 2016 the percentage of households in the US which had a computer was 89.3% [3], whereas across the continent of Africa, the number of households with a computer was 7.4% [4]. A more recent study by Ofcom in 2018 [103], investigated the uptake of various technological devices by adults aged

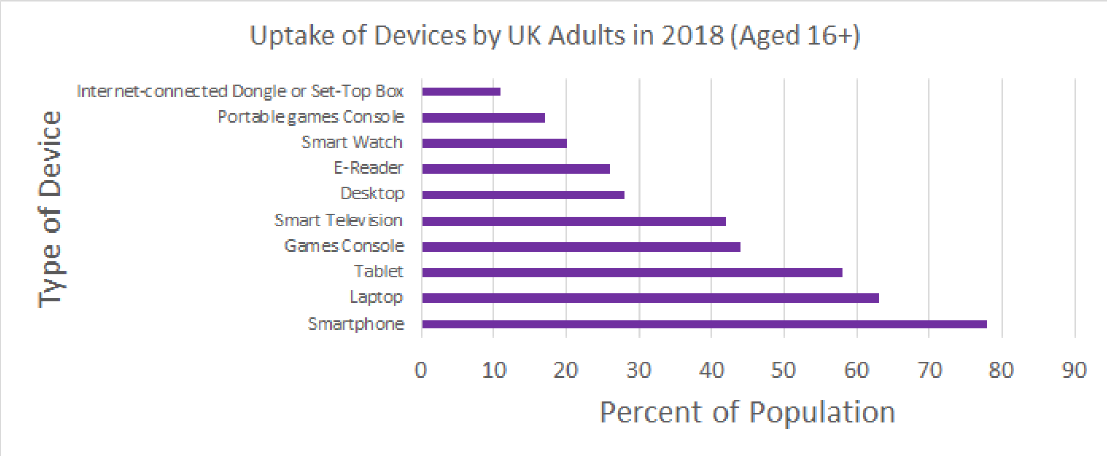


Figure 1.1: The uptake of devices by UK adults in 2018 as a percentage of the population– adapted from Ofcom (2019) [103].

16 and over in the UK. The results showed that 78% of adults aged 16 and over had a smartphone. Laptops and tablets were also pervasive with 63% and 58% of participants uptaking these devices respectively. The full results of this study can be seen in Figure 1.1. This suggests that the majority of adults have access to at least one form of internet enabled device. Furthermore, the statistics from Ofcom [103] suggest that in 2018 UK adults aged 18 and over, spent an average of 3 hours a day online, which indicates that the use of internet-enabled devices is commonplace. Given this, any effect these technologies have on individuals (e.g. effects on cognition) are likely to be widespread. If these effects are detrimental to the individual, then this could lead to large scale, negative consequences.

All the devices identified in Figure 1.1 can be used to access the internet and many have access to similar software and applications via app stores such as the 'Google Play Store'. However, each device has slightly different qualities which make it desirable to consumers. Rieger and Majchrzak (2017) [111] suggested a taxonomy (a classification system) to categorise devices. This was done based on three properties of the device: the media richness of the input, the media richness of the output, and how portable the device was. By mapping devices onto a matrix based on these properties, different devices occupied different (but sometimes overlapping) areas. This shows that despite similarities in applications and software between devices, they do have inherent differences.

Furthermore, this categorisation of devices focusses on tangible differences between devices, and does not consider how intangible features– such as the stereotypes held by the end user about the device– may influence how the device is classed. However, research in human-computer interaction has yet to fully understand the effects these intangible device differences may have on the user, particularly in relation to cognition. Therefore, the importance of these intangible factors are not yet known, and research is required to establish the potential effect of these factors.

The lack of research into these intangible effects is in spite of evidence that whilst many of these devices have similar capabilities, an individual will use a given device for different activities. For example, Kawsar and Brush [73] found that people prefer to complete short tasks (such as scrolling through a social media application) on a smartphone. This suggests a feature of the device is influencing this choice, as other devices are capable of allowing such activities to be undertaken. This may be due to tangible features such as the interface or screen size, and research into the effects of these features has been undertaken (for example in regards to the impact of screen size on memory [74]). However, there is limited research into whether intangible influences, such as the associations an individual makes between their device and a given activity, could also impact this behaviour. As such this thesis addresses a gap within the current work by investigating how stereotypes held about a device influence cognitive performance.

Understanding the differences between devices is also important as not all people have access to every type of device. This is exemplified by the variations in the percentage of the population with access to each device identified in the Ofcom statistics [103]. Consequently, even within countries with a large overall level of access to household computing such as the UK, there is still an issue of the 'digital divide'. This is a term meaning '*a division between people who have access and use of digital media and those who do not*' (van Dijk [143], p.1). The digital divide has been suggested to be comprised of two levels: firstly, those who have access to the internet and those who don't, and secondly, differences in usage of the internet (the usage gap) [139]. Of these two levels, the usage gap is particularly pertinent in developed countries, because although access to internet-enabled devices is common, some individuals are reliant on a single device as an access point rather than having an array of options to choose from. Given this, while there may be a preference to undertake a given task on a particular device, not all individuals will be able to do so. In terms of the present thesis, if some devices are more advantageous for cognitive performance than others, then those negatively impacted by the digital divide may find themselves at a disadvantage when carrying out certain tasks, compared to those on the other side of the digital divide. It is therefore important to consider all potential features that may impact an individual utilising a device. This means going beyond the existing research into interfaces and hardware features, and considering how the human end-user may influence their interaction, through mechanisms such as stereotypes held about an individual's device. Doing so will consequently allow a better understanding of the impact of the usage gap.

Despite the pervasiveness of devices within society, the cognitive effects of using and interacting with these devices in daily life is relatively unknown in terms of experimental research. This is despite concern over these effects, which have been reported in both the media [79] and academic publications [48]. Furthermore, limited differentiation between devices means there is an assumption that all devices will influence cognition in the same way, which may not be true. The main study which has considered these device differences was by Liu and Wang [88], who investigated the effects of using a desktop or a tablet to make a decision. They hypothesised that participants would make a decision that was congruent with the hedonic and utilitarian stereotypes of the device being used. In order to investigate this, Liu and Wang presented 80 participants with two hotels, and asked them to indicate in which they would choose to stay on a fictional trip to Singapore. One hotel represented the hedonic choice, whereby, it was new and well-decorated but further away from amenities. The other hotel represented the utilitarian choice, whereby, it was less aesthetically pleasing but was in a central location. Once participants had indicated the hotel they would choose to stay in, they were asked how they made their choice. Participants were then asked to rate the hotels on a scale of one to seven, where one was 'utilitarian' and seven was 'hedonic'. The participants were also asked to use this scale to indicate the extent to which they considered desktops and tablets to be hedonic or utilitarian. The results suggested that when a decision was made on a hedonic device (in this case a tablet), the hedonic hotel was chosen. They also found that the more salient the stereotype was to the participant, the more the likely the participant was to make the choice in-keeping with the device's stereotype. This supports the notion that intangible device features may influence human behaviour, but more research is required to fully explore this phenomenon within decision making, and in other cognitive domains.

As noted at the beginning of this chapter, cognition is a wide research domain. Therefore this thesis aimed to investigate two other cognitive abilities: information processing and memory. Information processing gives insight into how quickly people respond to a task. This also gives insight into how the information they are responding to is impacting the allocation of attentional resources. This then has an influence on memory, whereby greater attention being paid to a stimulus should lead to a better memory of that stimulus. Memory also underpins many daily tasks and interactions, and thus is an important cognitive ability to understand. Given the lack of research into the effects of different devices on these domains, and the importance of these cognitive abilities, the present work opted to focus on these processes in particular. This work also aimed to consider stereotypes held about devices beyond the hedonic and utilitarian qualities investigated by Liu and Wang. This allowed greater insight into the variety of stereotypes held about devices, rather than focussing on just two qualities.

Given the numerous types of device that exist in society, it is not feasible within the scope of this thesis to investigate them all. As such, the present work investigates 'everyday devices'. This was defined as devices commonly used in the UK, as suggested by the statistics presented in Figure 1.1, with similar computing capabilities.

It is also important to understand the psychological processes underlying this area, before moving onto applied research where these processes may have wider consequences. As this project is exploring the extent to which the stereotypes held about a device can serve as a cue impacting cognitive processes, other potential cues would confound the results. Therefore, in this project, the phenomena are explored within controlled settings, to assess the extent to which these effects occur. This approach allows for larger control [1] over the variables studied in the experiments, to allow clearer observation of the effects of different devices and the stereotypes held about these devices, on cognition.

1.2 Research Questions

Given the motivations outlined, the overall aim of this project is to investigate what stereotypes and associations are held about everyday devices, and whether these devices and the stereotypes held about them can serve as a cue influencing cognitive functions.

[1] Implications of the COVID-19 pandemic not withstanding.

In order to investigate this, four high-level research questions (RQs) were established. Here the high level research questions are reported, and are broken down to lower-level research questions in the subsequent chapters.

RQ1: What is currently known about how memory is affected by the device being used? This research question aims to establish the current state of the art in terms of the impact of different devices on one cognitive process, by reviewing the existing literature. This is achieved by conducting a systematic review of the work comparing the effect of different devices on memory processes, and synthesising current understandings of the effects of different devices on memory processes. This is reported in Chapter 2.

RQ2: How are everyday devices used, and what associations or stereotypes does the user hold about these devices? Answering this research question will give an understanding of how people are using their devices, what they associate with the devices, and how the devices are stereotyped. The aim of this is to establish the context within which people use their devices, and if the usage trends relate to the stereotypes held about each device. The associations and stereotypes attached to the devices found by answering this research question may then be used as stimuli to answer the subsequent research questions. This is explored in Chapter 4 by conducting an online survey which asked participants about how they used laptops, desktops, smartphones and tablets, and what stereotypes they held about these devices.

RQ3: What effect do devices have on information processing, and is this influenced by whether the stimuli are related to the device being used? This question aims to look at the effect of different devices on information processing speeds. Specifically, this question will investigate whether information processing speeds are affected when the stimuli being presented is associated with the device being used. This was investigated using an online, mixed measures experimental study and is reported in Chapter 5.

RQ4: How is memory for word lists affected by devices, and is this influenced by whether the words are associated with the device being used? Similarly to RQ3, this question aims to investigate whether different devices impact memory performance, and whether this performance is affected by the stimuli being related to the device being used.

This was also conducted using an online, experimental study, and had a between-groups design. The study is reported in Chapter 6.

1.3 Publications

Some of the work from this thesis has been published in academic venues over the course of the project. For clarity these will be listed here, with the chapters of the thesis they pertain to.

Steeds, "The Impact of Device Associations on Human Memory Performance." BCSHCI Doctoral Symposium, 2020 [130] – Overview of the project and the related work, and preliminary findings described in Chapter 4.

Steeds, Clinch, Jay, "Device Uses and Device Associations" Computers in Human Behavior Reports, 2021 [132] – Related work reported in Chapters 3.1 and 3.2, introduction, methods, results, findings, and conclusions reported in Chapter 4.

1.4 Thesis Structure

The rest of the thesis is structured as follows. Chapter 2 presents a systematic review of the research, in order to investigate RQ1. Chapter 3 then reviews the related work relevant to the thesis that was not identified in the systematic review. Chapter 4 presents an online study conducted to investigate RQ2. Chapters 5 and 6 investigate the impact of device associations on cognitive processes to address RQ3 and RQ4 respectively. Chapter 7 concludes the thesis, highlighting the contributions and implications of the work. Limitations of the present work are discussed, along with directions future research may take to expand upon the knowledge presented here, with reference to specific application domains.

Chapter 2

Systematic Literature Review

Given the motivations outlined in the previous chapter, this chapter aims to gain an in-depth view of the existing literature regarding the impact of devices on a specific cognitive process: human memory. This was done by conducting a systematic review of the literature which compared two or more devices, with the dependent variable measuring human memory performance. The papers included in the review were then categorised by the type of memory they investigated, and summarised. The trends in the findings were then analysed thematically. The themes identified six potential mechanisms which may have underpinned the effects on memory observed in prior research.

The results suggested that the effect of technology on memory was not uniform, and different devices had varying effects on different forms of memory. These findings suggest that technology cannot be grouped into one category, as devices and their effects on memory vary greatly. Furthermore, memory itself cannot be considered a single concept. Rather it is a construct made up of multiple facets, and research should be careful to appropriately identify the aspect of memory being studied. The six mechanisms identified in the review suggest that the ways in which technology impacts memory, and indeed other cognitive abilities, are diverse. The findings of the review also highlighted areas that future research may wish to explore.

2.1 Introduction

A range of technologies now play an everyday role in supporting human memory (e.g. to do list and reminder applications[1]), with emerging devices and trends set to provide further memory support [55]. However, there is also longstanding concern about the potential negative impacts of technology on memory, both in the academic [48] and popular press [79], as noted in Chapter 1.1.

Understanding the effects of technology on memory is complex, not least because of the wide variety of devices in common use. For example, the *Computer and Internet Use* survey [114] indicated that the majority (over 70%) of US households in 2016 owned a laptop or desktop computer and at least one additional device (e.g. tablet, smartphone). Other studies (summarised in Chapter 1.1) have reported similar findings. Furthermore, there's significant evidence that the way in which people utilise these technologies varies by device [73] (for more discussion of this see Chapter 3.1). Given the diversity of devices and their usage, it is highly likely that the impact of technology use on memory is non-uniform, with some devices affecting memory (both positively and negatively) in a different manner to other devices.

The present work aims to consolidate our understanding of *how different technological devices impact human memory* by conducting a systematic search and review of peer-reviewed literature across four databases. Specifically, a core challenge in this area is addressed by deliberately *sourcing and synthesising research from two distinct disciplines (i.e. psychology, and computer science)*. Studies of technology and memory are prevalent in both of these disciplines, with only limited evidence of knowledge transfer between them (see Table 2.1). In particular, it appears that research in this area from computer science is overlooked, despite four papers included in this review being published in computer science venues. This cross-discipline view is integral to developing a robust understanding of the effect of technology on human memory, that will facilitate future research and innovation.

[1] Searches on the Google Play store for terms such as “reminder", “memories" and “reminiscence" yield hundreds of results as of April 2021. These applications provide a range of memory-related functionality including both generalised and specific memory prompts.

	Papers		Citations					
	Count	Venues	Comp. Sci	Psychology	Interdisc.	Education	Oth. Disc.	Oth. Format
Computer Science	4	4	34	25	12	12	1	18
Psychology	8	5	14	148	86	47	5	90
Interdisciplinary	4	4	9	66	41	4	6	72
Other Format	1	1	1	68	15	2	2	24
Total	17	14	58	307	154	65	14	204

Table 2.1: The disciplines of the citations included in the papers studied in this review. Papers were categorised by the discipline they were published in. The citations in each paper were then categorised by their discipline. ‘Other formats’ was categorised as citations of books, theses, and web resources.

From an initial pool of 3474 papers identified across four databases, 17 were found to meet the inclusion criteria of studying the effects of a technological device on a measure of human memory (through a comparison with another device or an analogue intervention). Multiple papers from the final sample focussed on five areas of memory: narrative memory, working memory, language acquisition, semantic memory and spatial memory. One paper also concerned itself with verbal learning/memory. Irrespective of the area/type of memory under study, the effects of technological devices varied in terms of whether they were beneficial, neutral or detrimental to human memory. To better understand the reported effects of devices on memory, a thematic analysis was conducted which identified six underlying mechanisms which may have led to these differences in memory performance: screen size, attention, effort, embodied cognition, environmental cues, and use case.

This comprehensive, cross-discipline review is a valuable tool in helping to understand common and differential effects of different technologies on a wide variety of memory types and processes. By identifying specific mechanisms by which these effects may occur, a set of key considerations are provided for novel technologies as they continue to emerge. As such, this review synthesises understanding of psychological processes (beneficial to psychology) and also provides information relevant to the design and development of new technologies likely to influence human memory (beneficial to computer science), and the application of those technologies to specific domains (beneficial to interdisciplinary research).

The remainder of this chapter is structured as follows. Chapter 2.2 gives an overview of what is meant by memory as a construct, as well as defining the types of memory included in the review. The background and core concepts section also gives consideration to the features of technological devices which may influence memory performance. The methodology of the review is then described in Chapter 2.3, in accordance with the PRISMA guidelines [85]. In Chapter 2.4, the analysis is presented, grouping included papers by the type of memory that they investigate, and the trends seen in the results are discussed. The thematic analysis is reported in Chapter 2.5, which explores possible causes of the effects on memory, reported in the papers described in the previous section. Chapter 2.6 discusses the limitations of both the present study and the papers included in the review, as well as the overall implications of the findings of the review. Finally, Chapter 2.7 reports the conclusions of the chapter.

2.2 Background and Core Concepts

This section reviews previous work establishing what is meant by memory as a construct, as well as detailing the specific types of memory addressed in the review. Work pertaining to features of technological devices, and how they may influence memory is also discussed.

2.2.1 Memory

Memory refers to the ability to store and recall information. Memory can be split into types based on its temporal nature (i.e. if it is long term or short term) and content (i.e. if it is factual information or lived experiences). Each of these memory types is itself the independent subject of significant study, theorising, and intervention. For example, Baddeley [12] proposes a model of *working memory* which identifies different systems involved in short term memory storage and manipulation of information.

Given their differing roles in real life scenarios and processes, it is important to treat specific forms of memory independently, rather than as one overarching construct. Thus, when asking which elements of memory are being impacted by technology, this work identifies a number of distinct types of memory that commonly emerged in the studies included in the present review. These are as follows:

Narrative Memory: refers to the retention of information contained within a story.
Working Memory: relates to short term information storage and manipulation.
Language Acquisition: is the learning of language structures and vocabulary, either as a first language or as a foreign language.
Semantic Memory: involves the recollection of facts.
Spatial Memory: refers to our knowledge of routes within the environment around us.

With regards to semantic memory, it should be noted that the studies included in this review may also be considered episodic memory, as the information learnt may not be independent of the context it was learnt in. Determining whether a memory is semantic or episodic can be done using a method known as the remember-know paradigm [140]. In this paradigm, when a participant recognises a stimulus, you ask them if they remember seeing it before, or if they just *know* that it feels familiar. This allows semantic and episodic memories to be differentiated, as feelings of familiarity indicate that the memory is not tied to the context, and is thus semantic. As the included studies did not utilise this paradigm, it is unclear the extent to which the participants' memories of the information are tied to the event in which the information was learnt. Therefore, for the purposes of this review, studies concerned with the acquisition of facts are categorised as semantic memory, although they could also represent an episodic memory.

Memory Interventions

Ideas of how technology could serve as a tool to outsource mental functions such as memory have been considered for millennia [104]. Whilst some of these uses have emerged naturally through technology appropriation, many have been the result of an intentionally designed memory interventions (e.g. apps to remind people to take medication [129]). Such applications are highly prevalent on everyday devices.

In addition to these everyday technologies, specialist devices have been developed to aid and augment memory, a prominent example of which is *lifelogging*. Lifelogging devices use cameras and other sensors to produce a '*digital record ... of an individual's experiences ... stored permanently as a multimedia archive*' [38]. In particular, lifelogging cameras such as the Microsoft SenseCam [62] can capture images continuously throughout the day, providing a rich record of experiences that are especially valuable in supporting autobiographical memory (i.e. memory for the events that have happened to us).

Visual lifelogging technology has been utilised in memory interventions in populations such as those suffering from depression [106] and autobiographical memory impairments [2]. Research has also been conducted investigating how lifelogging can mediate memory performance [71]. Given the extent of research in this area, the specialist nature of the devices used, and the presence of an existing review of these technologies [55], it was decided lifelogging studies would not be included in the present review. Instead, the literature search focused on *everyday technologies with widespread adoption*: laptop and desktop computers, phones, tablets, and games consoles.

2.2.2 Device Features

Technology comes in a variety of forms, and this may cause factors other than the device itself to influence memory. In this section two potential confounding variables are discussed.

Immersion

Immersive virtual, augmented, and mixed reality devices have become more pervasive in recent years, and are readily obtained for home-gaming, as well as industrial purposes. As these displays are head-mounted and near-eye, it limits the participant's ability to look away from the device's display. This aids in giving these devices the property of a sense of presence, which is defined as the sense of being within the environment [123]. These devices can thus cause a greater sense of presence than raster displays such as desktops. A comparison between a desktop environment and a virtual reality (VR) environment by [121] found both quantitative and qualitative differences in how participants experienced these environments. Specifically, the participants found VR more immersive, and commented that the sense of realism caused them to form stronger memories of the experience. Due to this significant difference in immersion, it is possible that comparing these VR devices– which are known to cause strong sense of presence– to devices with raster displays will impact memory due to feelings of presence rather than the device itself. For this reason, studies utilising virtual, augmented, or mixed reality were excluded from the study.

Screen Size

A further difference between devices is screen size. Over the course of technological developments and trends in design, the standard sizes of devices such as smartphones have decreased and increased. Consequently, mobile and desktop versions of software have been made to optimise user experiences for the size of the device. In this section, an overview of the prior work relating to the effect of screen size on memory is given, but a more detailed account can be seen in Chapter 3.5.

There is conflicting evidence regarding how screen size can impact memory performance. A dissertation by Heo [57] suggested that in general, larger screen sizes led to better memory performance for television content. However, Kelley [74] conducted a study into the effect of screen size on memory for television news, and found no significant differences. As the latter research was conducted in 2007, it may be the case that watching content on smaller screens became more common, as camera phones became more prevalent than they were during the former study, conducted in 2003. This prevalence may have then reduced the impact of smaller screens on memory performance, as people were accustomed to watching content on small screens. Given these conflicting findings, the present review did not aim to restrict comparisons to similarly sized devices, although it is recognised that there is the potential for this to influence memory performance.

2.3 Methodology

This review assesses the present state of research in the disciplines of psychology and computer science. The review was conducted in accordance with the PRISMA guidelines [85]. The inclusion criteria, search strategy and method of data analysis are described in this section.

2.3.1 Search Strategy

The search was deployed on four databases – **APA PsycInfo**, **Elsevier's Scopus**, **IEEE Xplore**, and the **ACM Digital Library** – to investigate the literature published across psychology and computer science venues. These databases were searched using a string search that was adapted to fit the search style of each database. Key terms were truncated with an asterisk to gather articles with variations of the word and alternate spellings (e.g. 'memor*' was used to capture variations such as 'memory', 'memorise', 'memorize'). Four key concepts were searched for, linked together with 'AND'.

Within these concepts, 'OR' was used to search for synonyms and related words. Concept 1 related to memory, the main dependent variable for articles eligible for this review. Concept 2 pertained to cognition. As the search strategy includes computer science papers, cognition was added as a key concept to reduce the number of papers relating to computer memory in the results. Concept 3 was that of technological devices, the main independent variable of papers in this review. To focus on comparisons between devices, in cases where the term ‘device’ was not used, smartphones were used as the main device to search for in conjunction with a secondary device, due to their prevalence in society. The final concept related to comparisons to eliminate studies that only looked at singular devices.

The full search string for each database was as follows:

```
TS = ((memor* OR recogni* OR recal*) AND (cogniti* OR
metacogniti*) AND (device* OR (mobile* OR phone OR cell*
OR smartphone* AND (laptop* OR computer* OR tablet* OR
console*)) AND (compar* OR effect* OR review*)))
```

The number of papers generated from this string for each database can be seen in Figure 2.1.

2.3.2 Inclusion Criteria

After searching the databases, the abstracts were screened to assess which papers met the inclusion criteria. Articles were eligible if they met the following criteria:

1. Participants in the study were human;
2. A technological device was compared with at least one other device, or with an analogue medium such as pen and paper;
3. The study’s dependent variable was a measure of memory, not a measure of cognition as a whole;
4. The full paper was freely available[2] in English;
5. The paper was not a systematic review or meta-analysis.
6. The paper’s comparison did not include immersive devices, or specialist devices.

As smartphones were identified as the primary device within the search terms, only papers from 2008-2020 (inclusive) were considered, that is, papers published after the first release of the iPhone (in 2007) which launched the modern smartphone era. The

[2] This included papers available to the general population (i.e. gold open access), and those available as a result of publisher/publication subscriptions of the authors’ host institution.

data was collected on the 5th of June 2020, so papers published after this date were not included.

2.3.3 Data Extraction and Management

Two reviewers screened the abstracts and categorised the papers as either *include*, *exclude*, or *maybe*. One reviewer screened all the abstracts, while the second screened a random sample of 155 papers. The inter-rater reliability of this sample was 96.1% (Cohen's $\kappa = .480$, moderate agreement). Papers where categorisation disagreements occurred were discussed, and an agreement in classifying the papers was reached. Papers categorised as *maybe* were included for full-text analysis so as to avoid false exclusions. The full papers were then screened, and those that met the inclusion criteria were included in the final sample.

2.3.4 Data Analysis Procedure

Included papers were categorised by the type of memory studied (e.g. working memory), resulting in six groupings (i.e., the five memory types noted in Chapter 2.2.1 plus a grouping for all 'other' memory types). One paper studied two types of memory and is therefore included in both groups. Analysis on a per-group basis is reported in Chapters 2.4.1-2.4.6 (inclusive). The findings of included papers were then analysed qualitatively using thematic analysis to investigate trends in the findings regarding the effect of different devices on human memory (reported in Chapter 2.5).

2.3.5 Search Results

Across the four databases, 3474 records were found. These were screened to remove duplicates (exact matches of the authors, titles and abstracts), leaving 3472 unique papers. A further 2035 papers were excluded due to failure to meet the original search terms within the title, abstract or author keywords (e.g. the IEEE database additionally returns matches against its own automatically-generated keywords). A random sample of 20 of the excluded papers were additionally screened by one of the reviewers who was not told that these papers had already been excluded, to ensure they were not excluded

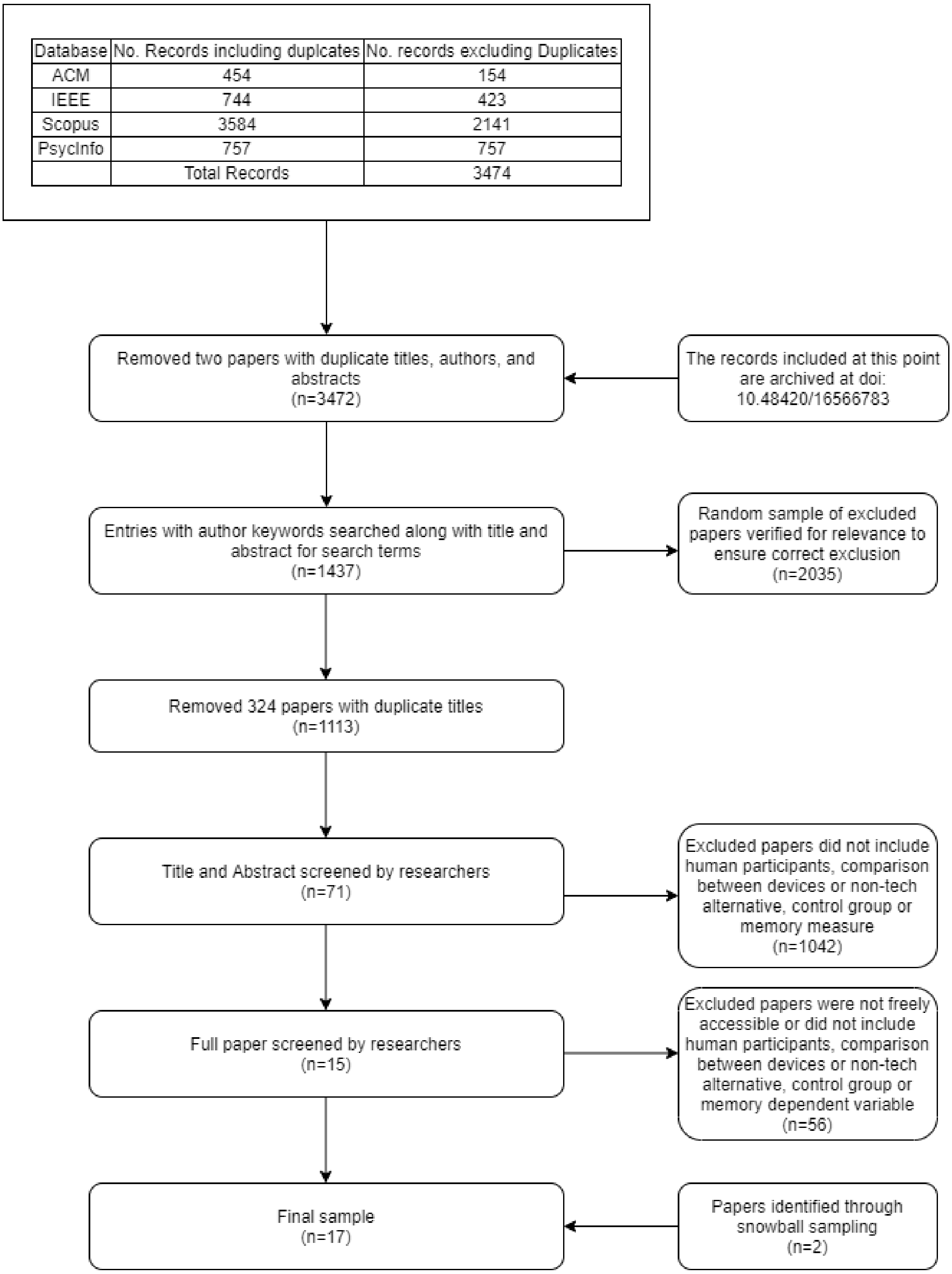


Figure 2.1: A PRISMA flowchart for study inclusion.

incorrectly. The reviewer excluded all papers in this sample suggesting that they were excluded correctly.

This led to 1437 papers remaining, which were then screened to remove papers with identical titles, leaving 1113 papers. This was necessary because some of the databases would append the abstract field with the database's copyright information, leading to the same abstract not being an exact match, and thus not considered a duplicate in the previous duplicate removal process. These papers then went through the title and abstract screening process, after which 71 papers remained. The full texts were then screened leading to 15 papers being retained for the review. Two papers were further identified via snowball sampling, leading to a total of 17 papers being included. A summary of the included papers can be seen in Table 2.2. A summary of the overall process can be seen in Figure 2.1.

2.4 Analysis and Findings

Seventeen papers met the criteria to be included in the analysis. This is a small sample, but it is a reflection of the fact that this is not a well explored area of research. This section presents these papers, grouped by the type of memory they investigate.

2.4.1 Narrative Memory

Five papers and one dissertation were identified as evaluating the participants' memory for narratives. Of these, two papers compared tablets and paper books, one compared e-readers and paper books, and one compared all three mediums. The fifth paper compared participants writing down a story as it was read to them using either pen and paper, a laptop, a tablet using a keyboard input, or a smartphone. The dissertation compared children's memory for story content using a desktop computer, tablet, or television.

The two papers comparing tablets and paper books focussed on story recall in children. Richter and Courage [110] asked 79 children aged 3-5 years old to listen to an adult read one story from a tablet, and a second story from a paper book (or vice versa). After each reading, the children were asked nine questions about the content of the story. Similarly, Yuill and Martin [156] asked 24 children aged 7-9 to read books with their mother, with either the child or parent reading, and with the book presented as a paper copy or on a

Reference	Devices Evaluated	Memory Type	Sample Size
[6]	Paper, Computer*, Tablet, None	Semantic Memory	50
[7]	Satellite Navigation System, Natural Language Interface, Human Navigator	Spatial Memory	61
[10]	Tablet, Paper	Working Memory	$23, 22$
[32]	Desktop, Smartphone	Semantic Memory	117
[42]	Laptop, Tablet, Smartphone, Paper	Narrative Memory	$63, 43$
[65]	Console/Computers, Mobile devices	Working Memory, Verbal Learning	88
[67]	GPS, Map, Experience	Spatial Memory	66
[68]	Smart Device*, Paper	Language Acquisition	$20, 20$
[80]	Tablet, E-reader, Paper	Narrative Memory	57
[81]	Tablet, Paper	Semantic Memory	86
[90]	Tablet, Laptop, Paper	Semantic Memory	36
[91]	E-reader, Paper	Narrative Memory	50
[95]	Tablet (Keyboard, Stylus), Paper	Language Acquisition	147
[98]	Television, Tablet, Desktop	Narrative Memory	132
[110]	Tablet, Paper	Narrative Memory	79
[147]	Paper, Computer*, None	Semantic Memory	127
[156]	Tablet, Paper	Narrative Memory	24

Table 2.2: Summary of papers included in the review. The sample size for each experiment presented in the paper is reported.
* denotes that the specific device is not stated. More information can be seen in the paper summary in Section 2.4.

tablet. After two stories on one medium, the children were asked to recall what happened in the text, before moving onto the next medium, leading to two recall sessions for each child. Neither study found statistically significant differences in recall between stories read from tablets and stories read from books.

Kretzschmar *et al.* [80] investigated 57 adults' recognition of events from short texts presented on either a tablet, e-reader, or paper book. Participants read three texts on each device, and were asked questions about the texts before moving to the next device. Participants also had their eye-movements tracked and EEG readings were taken. The results suggested that medium did not have a significant effect on participants' memory performance.

The fourth paper investigating memory for story content compared e-readers and paper books. Mangen *et al.* [91] asked 50 participants to read a 28-page story from either an e-reader or a paper book. Participants were then asked to indicate which words and sentences they recognised from the text. This was followed by a recall task, asking about the story content in terms of the characters, locations, objects, and timeline. Participants were then asked to indicate in which part of the book an event occurred (the first, second, or third section), and to reconstruct the order of events in the plot.

Results for Mangen *et al.*'s recognition tasks indicated there was no significant difference between the e-reader and paper groups. The recall results also suggested similar performances between the two groups for all question categories except timeline, where the paper group performed significantly better ($p < .05$, $\eta^2 = .22$). Regarding where in the book an event occurred, the second and third sections of the book saw similar recall between the groups. However, the paper book group recalled significantly more ($p < .05$, $\eta^2 = .06$) about the first part of the story than the e-reader group. Finally, when asked to reconstruct the plot, participants in the paper book group were significantly closer to identifying the correct order than the e-reader group ($t(48) = 2.03$, $p < .05$, $\eta^2 = .08$). The significant results had medium to large effect sizes. These results suggest that temporal-based narrative information is better recollected when encoded from paper books rather than technological devices.

The paper by Frangou *et al.* [42] asked participants to transcribe a story as it was read to them using different writing modalities. In the first experiment, the 63 participants were children aged between 10 and 11 years old. They were tasked with writing down three stories, one on paper, one on a laptop and one on a tablet. The story was then repeated so the children could check they had transcribed it correctly. A week later, the participants were asked to say what they remembered about each story in a free recall task. This was marked by assessing how many details the children recalled out of 20. In the case that nothing could be recalled, a one word cue was given. Experiment two used the same procedure except the 43 participants were aged 16 and they used a smartphone instead of a tablet. The recall test in this experiment was out of 25.

The results of experiment one suggested that for 10 year olds, there were no significant effects of device on memory performance. For 11 year olds however, the stories written on paper were recalled significantly better than the stories recorded on a laptop ($t(31) = 3.32, p < .01$) or a tablet ($t(31) = 2.15, p < .05$). The results of experiment two also found significantly better memory performance when the story was recorded on paper than on a laptop ($p = .011$). However, this significant difference was not found between the smartphone group and other modalities. It should be noted that there are two issues with the statistical analysis in this study. First, the difference between the device for the 11 year olds was calculated using t-tests rather than a post-hoc test with appropriate corrections. As such, this result may be incorrectly rejecting the null hypothesis (type 1 error). Secondly, the analysis of experiment two appears to have been reported twice, giving two different ANOVA results, one ANOVA followed up by post-hoc tests, and the other t-tests. In the present review, the more conservative results (i.e. the higher p-value) are reported, but it should be noted that there appears to be conflicting analysis. Despite this, taken together, Frangou *et al.*'s results suggest that memory for narrative declines when recorded using technology – something that the authors attribute to the amount of effort expended hand writing information relative to typing on devices. The authors further suggest that the lower typing proficiency in 10 year olds led to similar levels of effort for all modalities, leading to non-significant differences in memory performance.

Finally, the dissertation by Menkes [98] investigated children's memory for narrative content displayed on a television, desktop, or tablet. In this study, 132 children aged 4 or 6 years old were assigned to either the television group, the tablet group, or the desktop group. In the television group, the participants were asked to watch a children's cartoon. In the other conditions, the children were asked to play a game to familiarise themselves with the devices controls (a touchscreen, or a keyboard and mouse), before watching an interactive version of the cartoon viewed in the television condition. Following this, they were asked questions about the plot of the cartoon in a free-recall task. After the free-recall task, the children took part in a cued-recall task, followed by a task where they were asked to put images from the cartoon in the order in which they occurred. These memory scores were then combined to give an overall *media comprehension* score. Measures of executive functioning ability were taken to control for extraneous variables such as working memory capacity which may have also affected the media comprehension score.

The results showed that when controlling for other factors, the television and tablet groups showed significantly better media comprehension than the desktop group ($t(115)=-2.81$, $p = .006$, and $t(115)=-1.99$, $p = .05$ respectively). This suggests that increasing interaction between an individual and a narrative will not necessarily improve memory for the narrative. Menkes suggests that the desktop may have been harder for children to interact with, causing them to attend to the narrative less. However, similarly to Frangou *et al.*, t-tests were used rather than post-hoc tests. Consequently this may also have lead to the finding of significant differences, where a more conservative test would not have.

Overall, these papers suggest that generally there is little difference in memory for narratives based on the device said narrative is presented on. However, there is a difference with regard to memory for the order the events took place, and of the time frame for the events in the text. Mangen *et al.* suggest that this may be due to the physical indicators present with a paper book, whereby (in Western societies) the right side of the book is thicker at the beginning and becomes thinner the further through the book you read. This may be increasing the awareness participants have of their progress through the story, causing them to encode the temporal information to a greater extent. It may then be the case that displaying an indicator in an e-book of the readers progress, would reduce this difference in memory performance. Similarly, when tasked with recording narrative information, the reduced effort involved in typing may lead to reduced awareness or attention being paid to the content. As such, when recording information with technology, increasing interaction and attention to the task may improve memory performance. However, the study by Menkes suggests that for younger children, interacting using a keyboard and mouse may be more effortful, leading to detriments in memory performance. This suggests that a balance needs to be struck between something being effortful and effortless to maximise attention and information retention.

2.4.2 Working Memory

Two papers examined the effect of different devices on working memory (WM). Huang *et al.* [65] investigated the effect of device on WM in the context of gaming, which has been established to aid WM ability [18, 28]. In their study, participants were divided into three groups: non-video game players ($n = 29$, those who play less than five hours a week), mobile video game players ($n = 25$, play games for five or more hours a week, predominantly on a mobile device), and console/computer gamers ($n = 34$, those who play five or more hours a week on a console or computer). This study had multiple dependent variables, one of which was WM. WM was measured using the N-Back Task [76] which was performed as a 1-Back and 2-Back, where a lower N represents a smaller WM load. The N-back was scored in terms of speed (response time) and accuracy.

Significant effects of platform were seen for WM in terms of both response time and accuracy. Mobile video game players showed significantly better accuracy on both the 1-Back and 2-Back than the non-video game players. Console/computer players did not significantly differ from either of the other two groups in terms of accuracy. In terms of reaction time on the N-Back, mobile players were significantly faster than non-video game players on both N-Back tasks. Console/computer players were also significantly faster than non-players on the 1-Back but not the 2-Back. While mobile players performed better than console/computer gamers on both speed and accuracy, the differences did not appear to be significant. However, it should be noted that the p-values indicating the difference between the three devices in the post-hoc tests were not reported. Furthermore, the type of post-hoc test utilised was not reported. As such, the analysis undertaken to obtain these results is not clear.

Arreola *et al.* [10] investigated WM performance when a tablet or paper was available to offload information. They conducted two experiments, following the same procedure. Participants were asked to learn word-pairs of an academic subject and a testing format (e.g. PSY-QUIZ), and in the practice and choice trials were able to offload the information to either a piece of paper or a tablet device. The offloading device was assigned. This was followed by a no-choice trial where participants were not allowed to offload information. The only differences between experiment 1 and 2 were the number of word pairs shown before recall. For experiment 1 it was 2, 4 or 6, and in experiment 2 it was 4, 6, or 8. The experiments had 23 and 22 participants respectively. The results were analysed using a mixed ANOVA.

For all practice trials and choice trials there was a significant effect (all $p <= .006$) of the number of word pairs on the amount of information offloaded to a device, where more trials led to more information being offloaded. In experiment 1 there was also an effect of the number of word pairs on WM accuracy, where accuracy reduced as the number of pairs increased ($p <= .001$). There was no main effect of offloading device, although there was an interaction in the no-choice trial which suggested the paper group had a naturally better WM ability for the two word pairs trial ($p = .01$), but this did not persist when more word pairs were added. In experiment 2, there was a main effect of device ($p = .009$), where the paper group offloaded more than the tablet group in the practice trial, however this did not lead to a difference in WM performance. In the choice trial, the amount of information offloaded between paper and tablet users was not significant, but there was a significant effect of device on accuracy whereby participants offloading onto paper were more accurate ($p = .019$). The results suggested that overall no device was more beneficial for WM, however utilising paper methods sometimes led to more offloading than using a tablet device.

Taken together these studies suggest that the context under which a device is used impacts the effect it has on WM performance. In terms of gaming, mobile technologies are not significantly better for WM capacity than other gaming devices, but do have benefits for WM capacity compared to non-gamers. The increased benefit of mobile platforms for working memory may be due to the smaller screen sizes of mobile devices being able to display less information at once, leading to more information being held in the WM. Conversely, in scenarios where technology is used for cognitive offloading to aid WM, technology did not benefit WM more or less than traditional methods. This suggests then that it is not technology itself that impacts working memory performance but rather how it is used. These results further suggest that some technology may hold qualities that lend themselves to training working memory, such as reduced screen size. However, it should be noted that these studies investigated different aspects of WM, and therefore the extent to which they can be compared to each other is limited.

2.4.3 Language Acquisition

Two papers investigated the role of different technologies on language acquisition. One study compared paper-based learning materials with 'smart devices', defined as a smartphone, tablet or e-reader [68]. The second study compared pen and paper with tablets with a stylus input and tablets with a keyboard input [95].

Jo and Lim [68], asked university students in Korea to complete a Lexical Decision Task (LDT), where they were shown a word in English and asked to indicate whether it was a real word or not. Learning is measured by the number of correct answers and the speed of response (where a faster response is indicative of knowledge). Ten students used a smart device to learn and ten learnt using paper-based methods. The specific smart device is implied to be a smartphone, however this is not explicitly stated. The material was studied across two learning phases. The LDT was conducted three times: at baseline; after learning phase 1; and after learning phase 2.

Jo and Lim's results indicate that both groups showed improvements in both the number of correct answers and in response speed over the course of the study. The smart device group appeared to make slightly larger gains, however this is not shown to be statistically significant. A further comparison was conducted between 20 elementary school students and university students when using paper and smart devices. The elementary school students showed faster response times for the smart device group but not the paper group. However, this difference was also not shown to be statistically significant, and the response speeds following learning phase 2 became similar between the groups, suggesting the gains may be due to differing ability levels at baseline.

Mayer *et al.* [95] asked 145 German children aged 4 to 6 years old to take part in a literacy training intervention. The children undertook this training on either pen and paper, a tablet with a stylus, or a tablet with a keyboard. The participants were assessed on their ability to recognise eleven letters (letter recognition), write sixteen letters (letter writing), read five words (word reading), and write five words (word writing). The latter was marked as a percentage of the number of correct letters out of the total number of letters in the word. Letter recognition involved the presentation of the letter correctly as well as three variations in terms of mirroring the letter or changing the lines. Letter recognition and writing was measured at baseline (T1), after seven weeks of training (T2), and four to five weeks post-training (T3). Word reading and word writing were measured at T2 and T3. It is unclear if and how literacy skills were trained between T2 and T3, therefore changes in skill between T2 and T3 may have been influenced by further literacy skill training.

In terms of letter recognition, the results suggested there was a significantly larger change in ability at T1 and T3 for the pencil and paper group compared with the tablet with keyboard group ($p = .043$), indicating the paper group recognised more letters than the other group. The tablet and stylus group did not differ significantly from either of the other groups. For letter writing, there were no significant differences between the groups. There was a statistically significant difference in word reading ability between the tablet with keyboard group and tablet with stylus group at T2 ($p = .033$), with the tablet and keyboard group performing better. This difference did not persist to a significant degree at T3. Finally, in terms of word writing, children in the tablet with keyboard group performed better than the tablet with stylus group at T2 ($p = .036$). Again, this difference did not maintain its significance at T3. Overall these results suggest that there are minor short-term advantages in using a tablet with a keyboard in overall word learning, relative to a tablet with stylus. However the tablet and keyboard may cause disadvantages in recognising letters, which the authors suggest may be related to embodied cognition, where the motion of drawing the letters leads to a stronger memory than typing the letters. The lack of statistically significant differences between T2 and T3 may have been due to the unknown level of literacy intervention during this time.

Taken together, these studies suggest that technology may be a useful tool in language acquisition, although traditional methods are still valuable. Technology which reduces the amount of physical engagement with the target knowledge (i.e. pressing a key rather than drawing a shape) may lead to detriments. This may be due to the amount of effort being expended reduces memorability, or it could be because the actions are no longer embodied (i.e. they no longer have a physical action associated with them to strengthen the memory).

2.4.4 Semantic Memory

Five papers were identified as investigating semantic memory, through tests of either recognition or recall. These papers examined various combinations of laptop and desktop computers, tablets, paper-based methods, and no device.

The first paper by Covaci *et al.* [32], investigated memory for content in a digital board game. One-hundred and seventeen Junior High School students (aged approximately 12 to 15 years) were asked to participate in an online game called 'Fragrance Channel' on either a desktop computer or a smartphone, and either with or without scents that matched the fragrances presented in the game. Following the game participants were asked sixteen questions – eight true or false, and eight multiple choice. The results suggested there were no significant differences in memory performance between the desktop and smartphone groups, and the device used did not have a significant interaction with the presence of scents.

Mangen *et al.* [90] conducted a study investigating whether writing word lists with a laptop, tablet, or paper impacted participants' memories of the words presented in the list. Tablet users interacted with a keyboard on a touchscreen interface. Thirty-six women took part in all three conditions, and their memory was assessed in terms of both recall and recognition. The results found that recall memory performance was significantly better in the hand-written condition than the tablet ($p = .050$), or laptop conditions ($p = .024$). Both had medium effect sizes ($r = .32$ and $.37$ respectively). This difference did not persist for the recognition memory test.

Wei *et al.* [147] investigated whether the device used to take notes during lectures impacted how well the lecture content was remembered. All 127 participants viewed the same ten-minute, pre-recorded lecture. The study compared no note-taking, paper-based note-taking, and 'computer-mediated' note-taking. While the device in the latter condition is not explicitly stated, it is implied to be a laptop. Participants were also assigned to either the chat or no-chat group, where those in the chat group took part in a simulation of an instant messenger chat application during the lecture. Memory was assessed via ten multiple choice questions about the content of the lecture, immediately after the lecture ended.

The results suggested there was no significant main effect of device used for note taking on memory performance, even when no notes were taken. However, there was a significant interaction between the note-taking method and the use of the chat application ($F(1,120) = .309$, $p < .05$, $\eta^2 = .090$). This interaction suggested that using chat applications was more detrimental to memory performance when no notes ($p < .001$) or paper notes ($p < .05$) were taken compared to the computer-mediated note-taking. Wei *et al.* suggest that this may be due to the greater effort required in switching from paper to keyboard relative to just changing to the chat on the same device, as well the divided attention caused generally by the chat, leading to decreased memory performance.

Kwok *et al.* [81] investigated 86 children's abilities to learn facts when they were taught in person with paper visual aids or via a tablet. Half of the children (aged 4 to 8 years) were in the paper-based group, where a female teacher instructed the class and showed paper visual aids. The other half of the children were instructed by a cartoon llama in a tablet-based application. This was also voiced by a female instructor and used the same visual aids as the paper condition. The children were then asked to answer eight multiple choice questions about the content of the lesson. The results suggested that while the tablet group recognised slightly less information, this difference was not significant, indicating that tablets can be used to educate young children as effectively as in person instruction with paper visual aids.

The final paper, by Ando and Ueno [6], compared the use of different devices in e-learning. The e-learning content was categorised as being text, textual diagrams, text with still images, still images, or video. Fifty undergraduate students took part in this study, and were asked to take notes on the content with either a pen and paper, a pen tablet, a tablet PC, or to not take notes. The pen tablet condition used a writing tablet that did not allow direct writing on the e-learning content. The tablet PC used a stylus input where the e-learning content could be annotated directly. There was also a fifth condition where a keyboard input was used, but it is not clear if this was with a desktop computer, laptop, or another device. Participants were given a memory test on the e-learning content directly after the study phase, and one week after the study phase. The no notes group were not included in the tests one week later, as participants were allowed to refer to their notes when answering these questions. Participants were also asked questions regarding how well they comprehended the study material. The exact nature of the comprehension tests is not clear, as the paper is very short and lacks sufficient detail to replicate the study.

The results suggest that directly after learning, the no notes group performed significantly worse than the tablet PC group on memory and comprehension for still images, video, and text with still images. The keyboard group also performed significantly worse than the tablet PC group for all memory and comprehension measures except for memory of textual diagrams, and memory of text with still images. When tested one week later, the keyboard group performed significantly worse than the tablet PC group for all measures except memory of text, and comprehension of text with still images. The paper group performed significantly worse than the tablet PC group for memory and comprehension of still images, while the pen tablet group performed significantly worse on memory for video content than the tablet PC group. All significant results were reported as $p < .05$, rather than giving the exact p-values. These results suggest that being able to annotate directly onto e-learning material has advantages for memory and comprehension, potentially due to the reduced need to switch between the material and input device, allowing for more focussed attention. It also suggests that technology which uses hand-written inputs leads to better memory performance than when keyboards are used. Furthermore, different types of content (e.g. videos, text) may be less memorable when encoded using particular devices.

Overall, the findings of these five papers suggest that the device used has a variable impact on semantic memory. In some cases, a technological device is just as effective as paper-based tools, but in others the device can lead to an advantage or disadvantage. In particular, the extent to which individuals are interacting with the content being learned, and the amount of effort expended in switching between tasks impacts performance more than the device itself. As such, devices that facilitate interaction with the material, and reduce the extent to which attention is divided, appear to lead to better memory performances.

2.4.5 Spatial Memory

Two papers were identified as relating to spatial memory. The first paper, by Antrobus *et al.* [7], investigated memory for driven routes, landmarks on routes, and memory for the area, which was assessed by asking participants to sketch a map. Participants took part in one of four conditions. They were either directed using a satellite navigation system (SatNav), by an informed passenger who knew the routes but was not known to the participant prior to the experiment, a collaborative passenger who was taught the routes and was known to the participant prior to the study, or a "Natural Language Interface" that provided directions and could be asked questions. While participants believed the latter condition was using a technological system, it was in fact an actor who used a video feed to give appropriate directions and answer questions.

Memory for driven routes was tested by asking participants to put photographs from the route in the order in which they occurred in the route. The results suggested that people in the collaborative driver condition performed significantly better in this task than participants in the SatNav group. Landmark recognition was measured by showing participants a series of landmarks and asking them to indicate if they were 'on-route' or 'off-route'. Participants in the collaborative passenger group and the Natural Language Interface group performed significantly better than participants in the SatNav group. Finally the participants' sketch maps were evaluated in terms of the number of landmarks, the landmark orientation, the number of route intersections and the number of pathways. The results showed that the number of landmarks identified was significantly lower for the SatNav group than the collaborative passenger group, and the Natural Language interface (both $p = .001$), however there was no significant difference in the accuracy of the landmark orientations. Participants in the Collaborative passenger group correctly identified significantly more route intersections than the SatNav group ($p = .01$) and the Natural Language Interface group ($p = .04$). Finally, the number of pathways identified by the informed passenger group was significantly higher than the number of pathways identified by the SatNav group ($p = .044$). The results suggest that overall using a SatNav is less effective in supporting memory formation than other navigation methods. Overall, different navigation systems led to the best spatial memory performance, depending on the aspect of spatial memory being assessed.

The second paper, by Ishikawa *et al.* [67], compared spatial memory for six routes in 66 participants who navigated the route using either a paper map, a GPS, or direct experience. Direct experience involved being walked from the start to the end point by the experimenter, before being taken back to the starting point via a different route and navigating to the end again by themselves. The GPS gave participants directions to the end point, while the map group were provided a paper map at the starting point which had the start and end clearly marked on it. Spatial memory was measured by asking the participants to draw maps of the six routes after having navigated all of them. The results showed that the GPS group were significantly worse at drawing maps of the routes than the direct experience group. While the difference was not significant between the map group and any other groups, their mean scores in the sketch map task was closer to the direct experience group than the GPS group. The study utilised an ANOVA to assess the differences in performance between the conditions. However, the nature of the post-hoc tests, the test statistics, and the p-values were not clearly reported. Due to this omission, while the difference between the GPS group and the direct experience group exists, it is not clear to what extent, nor the probability of that effect occurring by chance. The study results suggest that the use of a technological navigational aid is detrimental to spatial memory performance compared with direct experience.

Taken together, these studies suggest that while the effect is not consistently significant, there is a trend for spatial memory performance to be worse for those using SatNav and GPS systems compared to other navigational aids. This may be due to the devices causing participants to pay less attention to their surroundings when navigating with technology, and thus encoding less navigational information. However, in the case of Ishikawa *et al.*'s study, the difference may have been due to the novelty of the GPS device (as only one participant reported using a GPS before), although if this was the case the novel Natural Language interface used by Antrobus *et al.* might also be expected to cause spatial memory deficits, which it did not. As such, it appears to be the GPS/SatNav device itself that is impacting memory, rather than the novelty.

2.4.6 Other Types of Memory

Huang *et al.* [65] also investigated the effect of gaming device and gaming frequency on verbal learning (see Chapter 2.4.2 for a description of the study). Verbal learning refers to the memorisation of verbal stimuli. This was measured using the California Verbal Learning Test Second Edition [34]. There were no significant differences of gaming frequency on verbal learning, nor were there any significant effects of device on verbal learning.

2.5 Discussion

A thematic analysis of the findings described in Chapter 2.4 was conducted to identify potential underlying causes of the device differences on memory performance. In this section, the six identified themes are discussed with reference to the papers included within the review and, where possible, are grounded in relevant psychological theories.

2.5.1 Screen Size

As discussed in Chapter 2.2.2, there is mixed evidence regarding whether screen size has a significant effect on memory performance. In the present review, there was no clear trend of screen size impacting memory performance. For example Covaci *et al.* [32] who investigated semantic memory and Frangou *et al.* [42] who investigated narrative memory, both conducted comparisons between smartphones and larger devices (desktops and laptops respectively), and found no significant differences in memory performance. The exception to this was the study by Huang *et al.* [65] who compared the effect of gaming technology on WM. They found that mobile devices (which typically have smaller screen sizes) *benefitted* WM performance.

This latter finding aligns with the structural characteristics/information processing (SCIP) framework proposed by Arthur Jr. *et al.* [11], which suggests that larger screen sizes put smaller demands on WM than smaller screens. Considering Huang *et al.*'s [65] study in light of SCIP, the smaller screen sizes of mobile gaming devices require the user to use more of their own WM abilities and may thus aid in WM training. However, the extent to which games rely on WM may vary, depending on the nature of the game play. Therefore, genre differences in games should also be accounted for when examining device differences in this research area.

Overall though, the evidence in this review suggests that whilst screen size may have a beneficial effect on the training of WM, it has less of an effect on other forms of memory. Future work may wish to investigate the impact of screen size on other forms of memory, as the previous work reported in Chapter 2.2.2, focussed on news stories, and there maybe further effects of screen size on other memory types, for example how different sized SatNavs effect spatial memory.

2.5.2 Attention

A potential underlying cause of memory deficits is that of attention. The role of attention in memory is supported by previous research such as that by Craik *et al.* [33], who found that divided attention while learning information (the *encoding* phase) led to reduced memory performance in a later memory test. However, divided attention during *retrieval* (the recall of information) did not lead to memory detriments.

Papers investigating spatial memory found that when participants were using a GPS/

SatNav to navigate, they showed reduced memory performance [7, 67]. This is consistent with prior research from Gardony *et al.* [45, 46], who observe a reduction in spatial memory performance following use of aural aids when navigating a virtual environment. Through a series of experiments reported in 2013 and 2015, Gardony *et al.* demonstrated that these effects can be attributed to divided attention, rather than interference in the processes themselves. The present review indicates that this memory impairment appears not only in the case of aural aids, but also for screen-based devices – potentially due to visual attention switching between the environment and the screen. However, it is of note that in Antrobus *et al.*[7], the Natural Language Interface (which participants believed to be technological in nature) did not lead to reduced memory performance, suggesting that it's possible to design technological navigational aids in a manner that reduces divisions of attention. However, it may also be the case that a true technological natural language interface would lead to worse performance if an artificial intelligence (AI) responded to the user less naturally than the human acting as the interface. This unnatural interaction may then require more attention to be divided between the environment and the navigational aid, impacting memory performance.

Some of the papers examining semantic memory also provide evidence of the role of attention in memory tasks. The study by Ando and Ueno [6] found that participants performed better on a memory test when they were able to annotate the educational materials directly, potentially allowing for greater focused attention. Similarly, Wei *et al.* [147] found that the presence of a chat application during learning affected memory performance, and this effect was exacerbated in cases where the chat was on a different device to the note-taking, as this caused greater effort to be exerted to switch between tasks. This suggests that technology should be designed to facilitate necessary task switching and reduce distractions. Given this, future work may wish to investigate design principles to aid task switching to improve user interfaces, and thus facilitate memory processes by reducing divided attention.

2.5.3 Effort

Effort has been proposed to be closely related to attention by authors such as Kahneman [70], however more recently, this has been disputed [21]. Therefore, the results of the present review will consider these as two separate phenomena. However, there is the potential for these concepts to interact and overlap, such as in the work of Wei *et al.* [147] where negative memory effects caused by dividing attention through the use of a chat application were exacerbated when switching between note-taking and the chat required greater effort.

The role of effort is more explicitly indicated in isolation in other semantic memory papers such as that of Mangen *et al.* [90], as well as for narrative memory in the paper by Menkes [98]. These studies suggested that in cases where the device used at encoding required greater effort to use, memory performance decreased. In cases where the devices required similar levels of effort, memory performance did not appear impaired. This suggests that ensuring the device being used is accessible, ergonomic, and well-suited to the task at hand, will help to reduce effort levels, and thus prevent reduced memory performance associated with effortful devices.

However, the findings from the narrative memory papers suggest that too little effort can be detrimental to memory performance. This is particularly seen in the paper by Frangou *et al.* [42], where they suggested that typing on devices required less effort than hand-writing, leading to poorer memory performance. This reflects the Yerkes-Dodson law, which suggests that both too little, or too much pressure is detrimental to performance [154]. In terms of effort and memory, this would mean either too much or too little effort would lead to reduced memory performance. This may then explain the seemingly contradictory findings identified in the review, where effort is both beneficial and detrimental to memory processes. Future work may then wish to investigate methods of optimising the effort required to use technology and user interfaces to make technology appropriately stimulating to use.

2.5.4 Embodied Cognition

The exertion of effort to improve memory performance may also relate to the concept of embodied cognition. Embodied cognition is the theory that cognition is influenced by the ways in which a person's body interacts with the environment that they are in [150]. Mayer *et al.* [95] suggested that the better memory performance demonstrated by the hand-writing group in the letter recognition task may have been due to the act of drawing the letters, which grounded the semantic memory in the movements. A similar phenomenon may have been seen in the study by Mangen *et al.* [90], as writing word lists by hand led to better performance in a recall memory test, than typing the word lists on a laptop or tablet.

This has implications for cases where the use of technology is reducing the amount of physical input required. The associated reduction in embodiment may lead to weaker memory traces being formed, leading to poorer retention of information. This may also provide an indication of where the mid-point of the Yerkes-Dodson effect (i.e. the optimal amount of effort) may be found for cognitive processes, as it indicates a point where the cognitive process is grounded enough in physical stimuli to aid memory, without becoming too taxing on the individual. Future work may then wish to consider how technology may be developed to aid in embodied cognition processes, while still being more convenient for users than traditional methods.

2.5.5 Environmental Cues

The environment itself may also provide support to memory processes. While the majority of papers investigating narrative memory showed limited differences between devices, the exception was the study by Mangen *et al.* [91], where questions relating to the timeline of events were less accurately recalled when using e-readers. This is suggested to be due to the physical cues representing progress through a narrative (e.g. the depth of the pages left to read) being absent for an e-reader but not a paper alternative. Therefore, when technology is being used as a replacement for traditional mediums, it may come at the cost of potentially removing environmental features that subconsciously inform memories.

Environmental cues have been identified to influence retrieval performance in psychology. A common example of this effect – known as *environmental context dependency* – is by Godden and Baddeley [51] who found that divers who learnt information underwater and were tasked with recalling it on land performed worse in a memory test than those who encoded and retrieved underwater. This phenomenon is discussed in more detail in Chapter 3.4, but work in this area is indicative that the nature of the environment around people can influence their memory formations, and can act as a cue to prompt recall. This review only identifies evidence of technology replacing environmental cues in the context of narrative memory, however, there may be other use cases where this occurs. Therefore, future research may wish to investigate other scenarios where traditional mediums are replaced with technology to investigate whether memory performance is still affected. Furthermore, the device being used may also be acting as an environmental cue. Thus, future work may wish to consider conducting research where the device itself changes between encoding and retrieval, to see if this also elicits an environmental context dependency effect.

2.5.6 Use Case

Finally, the use case (i.e. the context a device is being used in) may also influence the effect a device has on memory. This is seen in the WM papers evaluated in the present review, where two different scenarios were investigated – gaming and cognitive offloading. In the context of gaming, the use of technology had a significant effect on WM, whereas the device used to offload information did not. This suggests that it is the context in which a device is being used rather than the use of devices themselves that can influence WM performance. A similar effect was also seen for semantic memory performance in the study by Ando and Ueno [6]. In this study, they found that different types of study material (e.g. video content or text with still images) were remembered better on different devices. Therefore, different devices may be beneficial or harmful to memory depending on the use case, rather than the device itself having a generalisable effect.

Furthermore, the results of this review indicate that not all types of memory are influenced in the same way by the presence of technology, suggesting that the issue of technology use on human memory does not have a ‘one-size-fits-all’ solution. Rather, the type of memory being investigated in research must be clearly defined, and care should be taken not to generalise the impact of technology from one memory type to others. Future research may wish to broaden the use cases, and test the effects of devices in a variety of ecologically valid studies. It would also be beneficial for future research in this field to carefully identify the memory process being investigated, as they should be treated as distinct from one another, rather than as a singular construct.

2.6 Limitations and Implications

This section will discuss the limitations of the papers included in the systematic review, as well as the limitations of the systematic review itself. This is followed by implications of the findings from this review.

2.6.1 Limitations of the Included Studies

The primary limitation of the papers included in the systematic review is in the reporting of the statistical analysis. While many of the included studies utilised post-hoc analysis, the nature of this analysis was not clearly reported. Furthermore, at times, statistical methods such as t-tests were used without corrections, instead of post-hoc tests such as Tukey's post-hoc. This increases the risk of a Type I error (where the null hypothesis is falsely rejected). There was also an inconsistency across the papers included in the review in terms of how the results were reported (e.g. some papers gave p-values to 3 decimal places while others only reported that p was less than .05). As such some of the studies give clearer insight into the probability of the findings occurring than others. Furthermore, some of the test statistics (e.g. the t-values) were not reported, meaning the magnitude of the effect was not clear. Future research would thus be benefited by precise and consistent reporting of significance levels, test statistics, and effect sizes.

A second limitation of some of the included works is that they had limited sample sizes. In psychological research, a sample size of 30 is considered large in regards to the central tendency theorem. As such, the studies that did not reach this sample size may be considered to be under-powered. Furthermore, smaller samples are less likely to be generalisable to the wider population, and the findings may then be lacking validity in wider contexts. The studies included in this review sourced their samples from across the globe (e.g. Korea, Germany). This means there are potential cultural differences between the samples, consequently meaning the results are not comparable between studies. However, it is overall a positive point that research in this area spanned multiple continents and cultures, as it means the results are not biased towards one culture or geographic area.

2.6.2 Limitations of the Systematic Review

The present review was limited in only considering articles written in English, and thus may have missed work in this area published in other languages. However this was necessary due to the language abilities of the researchers. Three articles due to be included in the full paper screening process were not freely accessible to the researchers, as they were not open access, and the researchers' host institution did not have paid access to the work. Therefore the full papers were unable to be screened, and were thus excluded. As such there may be further research in this area that was not considered. Furthermore, devices designed for specialist purposes and immersive technology were not considered within the scope of the review (as rationalised in Chapter 2.2.2). Devices fitting this criteria may also give insight into the effects of different devices on memory though, which future work may wish to consider. Finally, the sample of papers included in this review is relatively small, and therefore while the findings of the review indicate clear trends, there is room for further investigation within this domain, particularly that focussing on specific samples and use cases.

2.6.3 Implications

Given the array of positive, negative, and neutral effects of different devices identified in this review, no overall statement on the influence of technology on human memory can be given. Instead, the results of the present review suggest that other factors may be playing a role in the observed trends of different devices affecting memory performance. This means considerations that must be made to ensure technology is being designed and used in the most effective way possible to avoid cognitive harm. From this review the mechanisms influencing the impact of technology and memory appear to be other cognitive processes that underpin memory (such as attention), or the use cases and environments that technology is being used in.

Given this, to mitigate negative impacts of technology on memory, it is important to consider cognitive processes more holistically, and identify where processes intersect. At these intersections, interventions to reduce negative effects can be applied to aid memory, and cognition more widely. Furthermore, researchers should consider where the effects of technology were positive, and identify whether these positive effects can occur in other areas.

However, as discussed, the environment and use case can play a role in the effect of different devices on memory. Therefore, it is important that future research not only conducts lab-based studies to identify where memory is affected, but explores them in applied scenarios to see if the effects translate to real life. As the role of technology in daily living increases, it is important that research is conducted in this area to ensure the consequences of device use on memory are fully understood.

There are also implications regarding the way memory itself is considered in relation to technology. Different memory processes appear to interact differently with both the technology being used, and the underlying mechanisms identified in the thematic analysis. Therefore, it is important and necessary to clearly define the area of memory being investigated, to ensure the future research in this field is easily categorised by memory process. If this is not done, then erroneous assumptions of how technology influences memory processes may be made.

Overall the evidence discussed in this review suggests that there is no singular effect of technology on memory. There appear to be multiple factors influencing any significant effects, as well as different causes of said effects on memory performance. Given this, it is important that work in this area clearly define the memory processes being observed and the procedures used to study these phenomena. Doing so will better inform parties interested in the effects of technology on memory of the conditions under which effects occur, and allow for appropriate policies, guidance, and – potentially – interventions to be designed and implemented.

2.7 Conclusions

This systematic review investigated the existing literature comparing how different technological devices influence memory. Thematic analysis of the 17 papers included in the review identified six potential mechanisms – screen size, attention, effort, embodied cognition, environmental cues, and use cases – which may lead to memory processes being influenced. Future research directions related to each of these were discussed.

The findings suggest that the effects of technology on memory are complex, and can vary depending on the type of device being used, the type of memory being researched, and the context in which the device was being used. In many cases, there does not appear to be an effect of devices on memory, and in cases where there is, it is not inherently negative. For spatial memory, there does appear to be a trend towards devices reducing the amount of attention paid to surroundings, causing less information to be encoded, and subsequently remembered. There was also evidence suggesting that when devices reduce physical actions, the lack of embodiment can lead to weaker memories being formed.

While these are all potential avenues for future research, the present thesis will hone in on the influence of environmental factors on memory performance. In particular, the role of the device as an environmental cue will be explored. In the next chapter, further related work is reported, to explore what is known about devices as environmental cues, and how factors relating to devices can influence cognition. This will give context to the studies conducted in the later chapters of this thesis.

Chapter 3

Related Work

While the systematic review in the previous chapter suggested multiple processes which may be affecting device-mediated memory performance, this thesis aims to investigate one mechanism in greater detail- environmental cues. Specifically the potential for the device being used to play the role of an environmental cue, and how the associations and perceptions of the device may contribute to this effect. To gain a better understanding of the work in this area thus far, the present chapter aims to discuss the related work, in a broader context than that discussed in the previous chapter.

This chapter first explores the prevalence and use of various devices. There is evidence that while all devices have the capability to undertake a given task, individuals have a preference for using a certain device in certain circumstances. The next section of this chapter explores work that has been undertaken thus far to investigate the stereotypes and associations held about different devices. The evidence suggests that people hold different stereotypes about individual devices, even when they have similar interfaces and software capabilities. However, research in this area is limited in scope. The third section of this chapter concerns itself with how stereotypes held about devices can influence cognition. Again, research is limited in this area, but research considering non-technological objects has found that stereotypes held about objects can influence their memorability. As devices are pervasive in the environment around people, the next section considers how elements of the environment can serve as a cue in cognition. The evidence indicates that the strength of these environmental cues can vary depending on the type of memory being studied. The fifth section investigates evidence of technology specifically serving as this cue. Research in this area has also been limited, with mixed evidence regarding the capability of tangible device features such as screen size to influence cognition. The chapter is concluded with a summary of the related work and how it pertains to the aims of the thesis.

3.1 Device Use

The way individuals use their devices is an important underpinning for the attitudes and associations formed around those devices. Thus, understanding trends and patterns in device use is an important foundation for exploring potential device associations and the resulting effects on cognition. For example, a rarely used device may serve as a more salient cue due to it's novelty (an idea explored in Chapter 3.5). Furthermore, devices used for particular purposes may be associated with certain stimuli, strengthening the cue (for more discussion on this see Chapter 3.2). Devices and their uses is a rapidly changing field as, for example, the emergence of the personal computer in the 1970s was swiftly followed by innovations in mobile computing, which enabled the laptop, tablet and smartphone.

Ownership and use of such devices has steadily increased such that a majority of US adults now report using laptops, desktops, tablets and smartphones. Specifically, statistics from 2018 indicate that 77% report using a smartphone, 73% a desktop or laptop, and 53% a tablet [61]. The statistics for UK device ownership from Ofcom [103] (discussed in more detail in Chapter 1.1) also reflect this trend. The Ofcom statistics do suggest that desktop ownership is lower than the other ownership of other devices, however, a survey from 2017 [29] that asked individuals when they last used various devices found similar results for desktops and laptops, suggesting that while ownership of desktops may be low, they are still used regularly. While other personal devices are growing in prevalence, their adoption rates remain comparatively low when compared to the uptake of smartphones, laptops, tablets, and desktops. For example in the UK, adoption of smartwatches and fitness trackers was at 20% of the adult population in 2018 [103].

Similar patterns of device ownership and use are found in other Western, developed communities [61, 73], and similar rates of technology adoption are seen in developing countries when compared to developed countries. In 2012, Miah and Omar [99] looked at data from the ITU World Telecommunication/ICT Indicators database to investigate trends in technology adoption in developing countries. They observed a large growth in mobile phone subscriptions despite limited infrastructure. This study also identified a growth in the estimated number of Internet users. While such statistics are only indicative of the number of people utilising mobile devices rather than a direct measurement, they do suggest a universal growth in the use of connected, mobile technologies.

Prior studies have examined how devices are used in the context of a specific application such as messaging [96] or within specific locations such as at work [105]. Kawsar and Brush [73] investigated device use within the home by conducting surveys of 55 households, followed by semi-structured interviews with 19 of those households. They found that in households with at least four devices, the task often influenced device selection (e.g. using a desktop or laptop, rather than a mobile device, for work-related tasks), and that shorter tasks were more likely to be completed on physically smaller devices. Research regarding use trends has also focussed on under-served populations such as developing countries [96] or older adults [13, 118, 119]. Research has also been conducted investigating individual devices, particularly in relation to problematic technology use [26, 151].

Given the prior literature, limited work appears to have been conducted investigating the way devices are used in daily life, outside of any specific context. As such, further work is required to understand the overall ways devices are used, to establish if and how this may impact how devices influence cognition. While there is a wide variety of devices that an individual may have access to, the present thesis opted to investigate just a subset of these devices. Commonly used devices were chosen as the subject of this work as they are the devices people will come into contact with frequently, meaning any potential effects on cognition would be of larger concern than for less frequently used devices. In order for the comparisons between the devices to be fair, the devices investigated also needed to have similar uses and capabilities. Given this, despite games consoles and smart televisions being pervasive according to use trends seen in Figure 1.1 (page 17), they were not investigated in this project due to their limited range of uses. The four most common devices based on the trends identified by Ofcom [103] and excluding devices with limited uses are smartphones, laptops, tablets and desktops. As such, these are the devices investigated in the present thesis.

3.2 Device Stereotypes and Associations

Stereotypes are comprised of a belief, and the group a belief is held about [72]. However there is evidence that stereotypical beliefs can also be held about objects [97, 116]. Although often associated with negative (prejudiced) behaviour against individuals and groups, stereotypes are thought to serve an important role in cognition by reducing cognitive load and thus freeing up resources for other tasks [137]. When stereotype expectations are not met in a cognitive task, effects on response speed have been observed [101].

Stereotypes in human-computer interaction (HCI) have often been considered through the perspective of how different user stereotypes interact with technology. Cooper [31] posited that designing for personas stops designers creating technology for themselves, instead enabling them to create for a wider range of people. Such personas often embody stereotypes [141] which can deliver benefits to teams designing new technology [52, 93]. However, as yet there has been little research conducted to understand the stereotypes held about everyday computing devices.

This is despite the fact that studies such as that of Kawsar and Brush [73] (described in Chapter 3.1) have indicated that devices with similar computing capabilities and interfaces are preferred for different tasks. This suggests that another feature of the device may be related to how people choose to use their devices, and this may be an intangible feature. Given the aforementioned evidence of stereotypes influencing cognitive performance, research into the stereotypes held about devices may give insight into usage trends, and potentially describe a way in which the devices that are pervasive in society could impact cognitive abilities. This section therefore describes research that has been undertaken thus far to understand the stereotypes held about computing devices.

One study that has investigated the potential for stereotypes held about devices has centred on understanding the associations made by external onlookers with those using a given device. Schwind *et al.* [117] attempted to quantify these stereotypes using the Stereotype Content Model (SCM) [40] (see Chapter 4.2). The researchers investigated how different personas are perceived when using various mobile devices, finding that the type of person using a device affected where a device fell within the SCM. For example, a senior citizen tablet (computer) user was perceived as less competent than a tablet-using physician. However, they also observed that each device tended to inhabit a similar space within the model irrespective of the target persona (e.g. any user of a visual lifelogging device was considered to be less competent than the same user engaging with a tablet computer). A second experiment was therefore run to establish whether device stereotypes held when that device was being used by a non-stereotyped person. Devices were found to occupy the same general space within the SCM as they occupied in the first experiment, suggesting consistency in the stereotypes held for devices irrespective of the end user.

Some research has also indicated that technological devices can be seen as social agents. Carolus *et al.* [24] developed a measure called Positioning Others and Devices (POD) which assessed how emotionally important a person felt their acquaintances and devices were. The POD ratings of the participant's smartphone and other devices, as well as the POD ratings of people they knew were compared. The results showed that the mean score for a smartphone suggested that participants felt emotionally closer to their smartphones than their flatmates. This suggests that perceptions of devices can be seen as comparable to perceptions of humans. This therefore supports the use of the SCM to investigate how devices are stereotyped, independent of a specific user operating the device.

A further way of categorising objects is to consider them in terms of whether they are for hedonic or utilitarian purposes. A study by Liu and Wang [88] showed that hedonic and utilitarian device stereotypes can affect consumer decisions. When presented with websites for two online hotels, individuals planning for a hypothetical city visit were more likely to select the more practical (utilitarian) option when viewing the hotels on a desktop, and the more pleasure-based (hedonic) option when viewing on a tablet. This relationship was mediated by the extent to which the participants reported perceiving these devices as fitting the hedonic or utilitarian stereotype (i.e. if an individual felt a tablet was hedonic they were more likely to select the hedonic hotel when using a tablet). Explicit questioning with regard to hedonic or utilitarian perceptions may have influenced the observed results, and attempts have therefore been made to identify factors that make up the two constructs in order to enable more covert measurement (e.g. Batra and Ahtola [15]). Such scales have been applied in domains such as branding [1] and website design [64], but have not been used in further examination of everyday personal computing devices.

Insight into stereotypes held about technology has also been gained by utilising word association tasks. Specifically, free association has been used to investigate terms associated with computing technology. Contarello and Sarrica [30] found that terms related to the Internet fit into themes of how the Internet is used or experienced, the present and future of the Internet, and its relation to personal use or wider use. 'Mobile technologies' have also been investigated using free association [144], and common words generated pertained to the practicality of being mobile. However, this technique does not appear to have previously been used to compare different device types.

Outside of these studies, limited consideration has been given to the stereotypes and associations made with devices. As such, there is a gap in the research area regarding how devices are perceived, and what the most salient stereotypes and associations for each device are. Furthermore, as discussed in the next section, these associations and stereotypes may have an effect on cognitive abilities.

3.3 Device Associations and Cognition

Little work has been conducted in regards to the effect of technological device associations on cognition. The primary contribution in this area thus far was the work of Liu and Wang [88], described in Chapter 3.2. However, links between cognitive abilities and objects which people hold stereotypes about have been explored in contexts such as advertising.

A study by Youn *et al.* [155] investigated whether the relationship between liking a product and a person's memory of that product varies depending on the type of product being shown. Adverts shown during the Superbowl (a major sporting event televised in the USA), were categorised as being for either utilitarian products (items such as paper clips that are not largely enjoyed by their users), approach products (enjoyable products such as nice clothes or a new car), or avoidance products (products used to avoid a negative effect such as deodorant). Participants were asked via a telephone interview if they had watched any of the Superbowl and if so, which adverts they recalled seeing in an unaided recall test. This was followed by an aided recall phase where participants were given a list of brands and asked if they had seen adverts for any of those brands. For each brand they remembered seeing, participants were asked how much they liked the advert. The results suggested that approach products had the strongest correlation between liking the product and both types of recall. Avoidance products did not show any significant relationships between liking and memory performance. Utilitarian products did not show a significant correlation between aided recall and liking, but did show a significant correlation between unaided recall and liking. The authors suggest that this may be due to the style of adverts produced for utilitarian and avoidance products which may be more factual than for approach products. Nonetheless, this is indicative that stereotypes held about an object can lead to variations in cognitive performance.

Stimuli that are semantically associated with each other (e.g. rabbit and cat are semantically related by both being animals) can also influence memory. Montefinese *et al.* [100] conducted a study where participants were shown a series of words during the study phase. They were then shown more words in the response phase, some they had seen before (old), some new words which were highly semantically related to the old (HSS), and some new words which had a low semantic relation to the old (LSS). For each word in the response phase participants were asked to indicate if they were viewing an old word or a new word. The results found that HSS words were miscategorised as old words significantly more frequently that LSS words. This suggests that semantic relationships between stimuli and lure stimuli (an incorrect answer offered to the participant) can cause false recollections. In terms of the wider thesis, this indicates that associations made with an object (such as a technological device) can influence memory performance, in that associated items are more likely to be falsely remembered.

Semantic relationships between cues and target answers can also be used to enhance memory. Naveh-Benjamin *et al.* [102] conducted three experiments investigating this effect. In the experiments, participants were presented with images and a word. The word was either related to the image (e.g. someone standing at the bottom of a waterfall was shown with the word ‘shower’) or unrelated (e.g. the waterfall image presented with the word ‘bench’). The picture was then presented in a cued recall task, and the participant was asked to say the word which had previously been paired with that image. The results from all three experiments found a significant main effect whereby words semantically related to the picture were better recalled than the unrelated words (all $p < .01$). They also found that this significantly interacted with age, where younger adults benefited more from the pictures and words being related than older adults. The mean ages of the younger adults were 22.7 years or lower for each experiment while the mean ages of the older adults were 72.6 years or older for each experiment. These results show that semantically related items can serve as a stronger cue for retrieval than unrelated items, but this effect may vary by the participant's age.

Overall the prior work in this area is indicative that semantic relationships and associations between cues and target answers can benefit memory. However, when there is a semantic link between a stimulus and a lure, this can cause false recollections. There's also evidence that an object's memorability can vary depending on the broader category or stereotype it falls into. In terms of the present thesis, there is a lack of investigation into this effect in terms of devices being used, but existing literature supports that their may be a relationship between a device (serving as a cue) and related concepts, associations, and stereotypes. Given this, there is the potential for a device placed with in the environment to serve as a cue, and, as explored in the next section, there is evidence to support the importance of environmental cues on cognitive processes, but particularly memory.

3.4 Environmental Cues

As discussed in Chapter 2.5.5, elements in the environment can influence memory performance. In particular, the environmental context dependency effect indicates that differences within the environment between encoding and retrieval can be detrimental to memory performance. The study by Godden and Baddeley [51] identified the effect of environmental context dependency, and it has since been replicated with a variety of cues. As briefly described in Chapter 2.5.5, encoding and retrieving stimuli underwater led to a better memory performance than when asked to either encode or retrieve underwater. There was also a further effect where being asked to encode and retrieve on land led to a better memory performance than when changing environments between encoding and retrieval. This study shows that the nature of the environment around people can influence memory, and can act as a cue to prompt recall.

While the difference between being on land versus underwater is quite prominent, other studies have found environmental context dependency effects with subtler stimuli. For example, Herz [59] investigated the effect of different odours being present during encoding and retrieval on memory. In her first experiment, Herz used an incidental word learning task to present the word list which participants would be asked to recall in a surprise memory test two days later. In this task participants were presented with a word (e.g. pencil) and asked to tell a story from their life involving that word. The results found that those exposed to the scent of osmanthus at encoding and retrieval performed significantly better in the recall task than those in the no-odour group, the peppermint group, and the fresh pine group. Those in the peppermint group also recalled significantly more words than the no-odour group.

Herz then conducted a second experiment to investigate whether these findings were due to the consistent presence of the scents at encoding and retrieval. This experiment had four groups: no-odour, odour only at encoding, odour only at retrieval, odour at encoding and retrieval. The odour used was either osmanthus, peppermint or fresh pine, and the odours were never mixed (i.e. there were no conditions where one scent was used at encoding and a different scent was used at retrieval). The same procedure as experiment one was used except the number of stimuli was reduced in experiment two. The results suggested that for the scents osmanthus and peppermint, participants recalled significantly more words when the odour was present at encoding and retrieval than in the other three conditions. When fresh pine was used as the odour there were no significant differences between any of the groups.

These findings suggest that the presence of a smell at encoding and retrieval can serve as an environmental cue to improve memory performance, but not all scents are salient enough to induce this effect. However, it should be noted that Herz drew the participant's attention to the odour if it was present at encoding. Therefore, the effect of a subtle environmental cue such as odour may only be beneficial if participants are explicitly aware of it. Despite this, this research does support the idea that cues present in the environment can aid memory performance, if they are present at both encoding and retrieval. This study also shows that the more environmental cues there are, the better the memory performance, because while all the participants were in a consistent room across encoding and retrieval, the added consistent cue of scent led to better memory performance.

Both studies described thus far utilised recall memory tests in the retrieval phase, however research has been conducted investigating recognition memory. Smith, Glenberg and Bjork [125] conducted a series of experiments investigating environmental context cues, and in the fourth experiment of the series, they compared the effects of these cues on recall and recognition. They found that while there was a consistent significant effect in the recall condition, there was little effect of environmental context dependency on recognition. This lack of recognition effect was further found by Godden and Baddeley [50] who replicated their previous study comparing the underwater and on land contexts, but this time using recognition instead of recall at retrieval. They suggest that the limited effect of the environmental cues is found for recognition because having the target items present is a more powerful cue than the environment, thus causing the environmental cue to be inconsequential.

However, there has been evidence to suggest that the environmental context dependency effect on recall may not be a robust phenomenon. Eich [39] conducted a study looking at the effect of changing environment on participants who were told to memorise the stimuli in relation to something in the environment compared to those who were told to memorise the stimuli in isolation from the environment. The results showed that changing environment only had a detrimental effect on memory when the stimuli was memorised in relation to the environment. This suggests that changing context has a limited effect on memory when the environment is not emphasised during encoding. This may explain why the subtle environmental cue of odour led to an environmental context dependency effect in Herz's studies, as the presence of the odour was emphasised in the procedure.

Furthermore, there is evidence that environmental context dependency effects do not persist in ecologically valid scenarios. Saufley *et al.* [115] conducted a series of experiments over three years where the examinations from seven university courses served as the retrieval phase. Learning for each course occurred in large lecture theatres and the examinations were conducted in either the large lecture theatre or a small classroom. The results did not find any significant effect of changing venue on exam results, suggesting that within this context, there was no environmental context dependency effect.

To weigh up the evidence for and against environmental context dependency Smith and Vela [126] conducted a meta-analysis of the existing literature in 2001. They found that overall there was a reliable effect of environmental context dependency. From a human-computer interaction perspective, given the pervasive nature of technology within an environment (e.g. most cities have screen displays on the streets as well as within shopping centres) it is important to consider how these displays and devices can play roles as environmental cues. Furthermore, as indicated in the study by Eich [39], when the stimuli is memorised in relation to the environment, the environment plays a stronger role as a memory cue. Given this, the associations and stereotypes held about devices in the environment may lead them to have a greater influence on cognitive processes than previously thought, because utilising the device may draw greater attention to the device as a cue. Considering this evidence, the next section reviews research investigating aspects of technology as cues for cognition.

3.5 Technology as a Cue

There has been limited research into the role of technology as an environmental cue, and the predominant focus has been on a specific aspect of a device or its use, rather than the device itself. One example of such an aspect is the position of information on a screen at encoding compared to retrieval. Wright and Shea [152] conducted three experiments where participants were asked to memorise the order of button presses on a computer keyboard. The stimuli was presented either at the top of the screen, in the middle of the screen or at the bottom of the screen. If it was shown at the top, the information was in blue, the shape of the cues was a diamond and a tone of 2500Hz was played during presentation. For the middle and bottom it was red, square and 1000Hz, and yellow, circle and 300Hz respectively. The results suggested that changing the presentation of the information between encoding and retrieval led to poorer memory performance. While the variable of where the information was on the screen was not isolated, this study was indicative that elements related to the device could have an environmental context effect.

Further research into the effect of screen layout has been conducted by Wästlund *et al.* [145], who investigated reading comprehension and mental workload. Participants were shown text and questions arranged either over four pages, or as one continuous, scrollable page. During the experiment participants were also shown a pop up which asked the participant to complete a decision making task. Response times were used as a measure of mental workload, where a faster response indicated a lower workload. The results suggested that the scrolling layout led to a slower response time, suggesting that having the information in one continuous chunk led to greater demand on cognitive resources. This suggests that differences in presentation of information on a device can influence cognition in domains other than memory.

The effect of screen size on cognition has also been investigated. A brief overview of this was given in Chapter 2.2.2, but the research in this area will be discussed in more detail here. Detenber and Reeves [36] investigated the effects of image sizes on emotional arousal. They projected an image that was either 22 inches (diagonally) or 90 inches onto a large screen and asked participants to rate their emotional responses.

It was found that larger images led to greater emotional arousal than smaller images. The authors suggest that this increase in arousal from large images has implications for cognitive abilities (as heightened emotional responses have been linked to attention and memory performance). Within the context of screen sizes, this suggests devices with larger screens may cause differences in cognitive performance when compared to devices with smaller screens.

Subsequent research by Reeves *et al.* [108] investigated the effect of screen size on attention. In this study they used heart rate as a proxy measure for attention- where the larger the drop in heart rate (relative to baseline), the more attention an individual is paying to the stimuli. They found that there was a greater drop in heart rate for participants looking at a larger screen, however this was only significant if the scores were aggregated across participants, suggesting that the effect may not be robust.

Further research on the impact of screen size on attention was conducted by Heo [57]. They asked participants to watch television content which was categorised as being either ‘news’, ‘advertising’ or ‘entertainment’. The content was presented on either a 32 inch screen or a 150 inch screen, and all participants sat 150 inches away from the screen. The effect of screen size on short-term attention was measured using the same procedure as Reeves *et al.*. Heo also measured effects on long-term attention based on the changes to baseline heart rate over time. In terms of short-term attention, while there was a significant effect that showed the heart rates of all participants slowed after seeing the stimulus, there was no interaction with screen size, suggesting the size of the screen did not influence short-term attention. Similarly, for long-term attention, while there was a change in heart rate it did not vary by screen size, suggesting the size of a screen does not impact the amount of attention people pay to content. This detracts from the findings of Reeves *et al.*.

Heo also investigated memory performance as part of this procedure. Participants were first asked to recall as much visual and verbal information from the television content as they could. They then undertook two recognition memory tests- one for the visual content and one for the verbal content. The results showed a significant ($p < .001$) effect of screen size on the recognition of visual content, where the large screen group recognised significantly more items than the small screen group. There was no significant effect of screen size on verbal recognition memory. There was also no significant effect of screen size on recall. This suggests that screen size can have an effect on some memory processes, but it varies in terms of memory type (verbal or visual), and whether the test is recall or recognition based. There were also no significant interactions between screen size and content. This suggests that devices which vary in size alone do not have an effect of memory for specific types of content.

Memory for television content has also been investigated when participants used a 48 inch television, a 27 inch television, and a 2 inch screen attached to a camcorder [74]. The latter had audio delivered via headphones, while the televisions utilised cabinet mounted speakers. Participants watched five news stories which lasted an average of 90 seconds per story. After watching the content, participants were asked ten recognition questions which were about two of the news stories. The results found no significant effects of screen size on the recognition scores. The exact questions asked were not reported and as such it is unclear whether the questions pertained more to visual or verbal content in the news stories. Given the lack of significant findings and based on the findings of Heo, it is more likely that the memory test drew on verbal components than visual. Overall this study suggests that screen size does not influence memory performance, but the lack of specificity in reporting the study, means that conclusions on the type of memory being tested in this instance cannot be drawn.

A dissertation by Holdener [63] investigated memory for multimedia content on mobile phones compared to computers, where the mobile group used a 2.8 inch touchscreen, and the computer group used a 20 inch desktop with a keyboard and mouse. Participants were shown four types of content– a video, text with a table, a game with animations and text explanations, and an animation with a voiceover– about the physics of sailing. After each piece of content participants were asked five to thirteen questions, assessing their comprehension, visual recall and verbal recall. It should be noted that while the author reports the questions as recall, all the questions were multiple choice. As such this is a recognition test rather than a recall test. The results suggested that in the memory test, participants using a mobile device performed significantly better than those in the desktop group for all types of content ($p < .05$). This is contrary to the prior research discussed thus far, where typically larger screens led to better performance if there were any differences at all. This may be due to the novelty of the technology. The dissertation was written in 2008 and as noted in Chapter 2.3.2, the iPhone was first released in 2007. Prior to then smartphones were not common place. Therefore, it may be the case that using a touchscreen mobile phone would have been a new experience for many participants. As novel stimuli can lead to better memory performance than non-novel stimuli [77], this may have led to significant differences in performance, while today, due to the pervasiveness of smartphone technology, there may not be an effect.

In summary, the extent to which devices have been investigated as environmental cues is limited. The evidence regarding the impact of device features on cognition is also conflicting, particularly in terms of the impact of screen size. Furthermore, this research has been limited to investigating the effect of tangible device features. Consequently, little is known regarding how intangible features of technology may serve as a cue, despite evidence of semantic associations affecting memory in other scenarios (as discussed in Chapter 3.3). Given the pervasiveness of devices within the environment around us, understanding whether or not the presence of a device can serve as a cue and have an influence on cognition is an important area to study, to develop the understanding of how different devices impact cognition.

3.6 Conclusions

Overall, research in this field has identified that devices may be used in different scenarios despite having similar computing capabilities and interfaces. There has also been a large amount of research into how under-served populations may use their devices, but little research into how devices are used more generally. There is also evidence indicating that associations are made with technology, and stereotypical beliefs may be held about different devices. However, this research has been limited in scope, and at times been investigated using explicitly questioning, rather than scales designed to measure the stereotype. There is a further body of evidence indicating that associations and stereotypes related to a word or object can influence cognition, but only one study has explicitly investigated this in regards to technology. There is also evidence that the environment around someone, and particularly changes in that environment, can influence performance in cognitive tasks such as memory. However, again there is little research considering how devices that are pervasive in society may relate to this effect. Research that has considered devices has only focussed on the tangible features of the device (e.g. screen size) rather than intangible features such as associations and stereotypes held about devices, which the literature suggests do significantly differ between devices.

Given these gaps in the existing literature, the present thesis first aims to investigate the stereotypes held about everyday computing devices. In Chapter 4, this is done by utilising an online survey. Part of this study also investigates the general use trends of the devices, and whether these use trends relate to stereotypes held about the devices. This research will thus provide insight into the wider trends of how devices are used, and the associations held about them. This will in part replicate some of the previous work regarding device associations, to gain insight into how robust previous findings in this area are. The stereotypes identified in Chapter 4 will then be used as stimuli in chapters 5 and 6 to investigate the potential for technology to influence cognition, via associations between the device and the stimuli, and the device's ability to serve as a cue. This will address the current gaps in the research by specifically focussing on the intangible device features, which could affect cognitive abilities. As prior research has considered the role of stereotypes held about devices in decision making, this thesis will focus on two other cognitive domains: information processing; and memory.

Chapter 4

Understanding Device Use and Device Stereotypes

Given the prior work reviewed in the previous chapter, there is a need for further work investigating the associations and stereotypes held about everyday devices. Furthermore, there is evidence that different devices tend to be used for different tasks. Consequently, in investigating the stereotypes and associations held about devices, it is also important to understand the ways the devices are used, to give context to the findings. Therefore, an online survey in which participants (n=177) described how they use different devices and the associations they make with the devices was conducted. Four devices were investigated: laptops, desktops, smartphones, and tablets. Participants were asked to answer a series of questions about themselves, how they use their devices, how they would rate the devices in terms of utilitarian and hedonic properties, and how they would rate a person using the device in terms of the Stereotype Content Model. The participants also undertook two interactive tasks. One was a free association task where participants typed words they associated with either a desktop or a smartphone. The second task was a multiple choice association task where participants were shown a series of words and asked to indicate if they associate the word with a desktop, smartphone, both devices, or neither device.

The results suggested that there were trends in the ways devices were used. For example, laptops and desktops were used for work purposes significantly more than smartphones and tablets. This also reflected the results of the free association task where desktops were associated with terms such as 'work', while smartphones were associated to social terms such as 'calls'. There also appeared to be some effects of the device being used, whereby the device being used by the participant was rated as more hedonic and utilitarian than that device would be rated by an individual using a different device. Overall the results indicate that there are consistent trends in the associations made with devices, and the stereotypes held about those devices. These stereotypes may have the potential to influence user interactions, and in particular, have an effect on cognitive processes.

4.1 Introduction

In human-human interaction, stereotypes have a significant impact on how individuals engage with others [53]. The associations we make between an individual (or group) and external concepts produce a range of overt and subtle effects. Examples include jury decisions based on racial profiling [20], and what is known as the ingroup/outgroup bias, where individuals prefer those they perceive as being similar to them, even along arbitrary lines such as a preference for people who like the same artworks as them [135]. Similar stereotypes have been shown to exist in human-robot interaction with near-identical observable effects [109, 136] as well as ingroup/outgroup biases against artificial intelligences [107]. However, the presence and impact of stereotypes associated with technology is not limited to devices designed to mimic humans – prior research demonstrates that even inanimate objects produce consistent associations [97, 116].

Understanding the stereotypes held for any given device could allow greater understanding of existing observations, and has value in shaping design. As an example, Hurtienne *et al.* [66] asked participants to choose which object best represented a concept (e.g. do white bricks or black bricks represent brightness?). They suggest that a device interface which reflects this imagery will be more intuitive for users and thus facilitate interaction. However, prior work has typically focused on very specific aspects of user interfaces, rather than the device as a whole. Considering device stereotypes holistically may facilitate the whole task conducted on a device, rather than just aspects of a task. Existing research that has been concerned with stereotypes at a device level focuses not on stereotypes held for devices, but instead on the way in which their users are perceived [117], and on the impact of associations on decision-making [88] (see

Chapters 3.2 and 3.3).

This chapter aims to investigate the stereotypes held for desktops, smartphones, laptops, and tablets using a combination of pre-existing scales and free associations. As interactions can play an important role in forming and altering perceptions of a stereotyped group [120], current usage patterns for each of the four devices are also investigated. This research therefore seeks to *develop understanding of both the use and perceptions of four dominant everyday personal computing devices*, and *how usage patterns and stereotypes intersect*. The primary research questions are as follows:

RQ 1 *How do UK adults use their everyday personal computing devices?*

RQ 2 *What associations do UK adults make with these devices? How do they perceive them?*

RQ 3 *Is there a relationship between the way an individual uses devices and their perceptions of devices?*

An online survey was conducted, recruiting 177 participants using the `prolific.co` platform. Device use was investigated by asking people about the activities they undertook on devices, the time spent on those activities, and their confidence in using different devices. Perceptions of devices were investigated in terms of how utilitarian or hedonic the devices were thought to be, how society perceives people using the devices, and words associated with desktop computers and smartphones. The survey contained a combination of open and closed questions, as well as two interactive tasks designed to determine the associations that individuals make with different devices.

The results indicate that the device used to complete the study had an effect on how devices were perceived, suggesting there were intragroup similarities in perceptions of all devices. The results also suggest that confidence in using one device with a particular type of input control (e.g. touchscreen) does not necessarily transfer to other devices with the same input mechanism. Furthermore, the placement of tablet users and smartphone users within the stereotype content model is the reverse of that found in prior research. This may be due to differing samples and contexts of exploring stereotypes, suggesting a potential fluidity in how stereotypes are perceived and a need for further research in this area. There were also differences in how devices were perceived in terms of utilitarian and hedonic qualities, suggesting that some devices may be seen as more useful and as such may be preferred for productive tasks.

4.2 Techniques for Capturing Stereotypes and Associations

Given the related work reported in Chapter 3, in this study the stereotypes held about devices are investigated in terms of utilitarianism (*related to practical tasks and work-related activities*), hedonism (*related to pleasurable tasks and leisure-related activities*), and the Stereotype Content Model (warmth/competence), as well as free word association.

Batra and Ahtola [15] presented a two-axis model that measures the extent to which an object is considered hedonic and utilitarian. This scale utilises word pairs (semantic differential scales) to measure the terms, with one word representing very hedonic or utilitarian and the other representing not hedonic or utilitarian. Hedonism is measured with the terms pleasant/unpleasant, nice/awful, agreeable-disagreeable, and happy/sad. Utilitarianism is measured with the terms useful/useless, valuable/worthless, beneficial/harmful, and wise/foolish.

The Stereotype Content Model (SCM) [40] is a two-axis model that explores intersections of competence and warmth. Groups that are seen as high in competence are considered to be of a higher social status, while those rated highly for warmth are seen as not competitive. As such, people perceived as being warm and competent are admired while those who are not warm or competent are seen with contempt. The scale measures warmth by asking whether the participant considers a person to be sincere, tolerant, warm, and good-natured, on a scale of 1 to 5. Competence is measured by asking if a person is considered to be competent, confident, independent, competitive, and intelligent.

While the SCM is traditionally applied to people of different social groups, it has been applied to devices as described in Chapter 3.2. This was done by asking the participants to think about a person using a specific device, and rate that user in terms of the SCM. This allows an inference to be made about how the device itself fits into the model.

While these scales consider stereotypes within established constructs, it is possible that these are not the only associations held about devices, if they hold them at all. Another common technique used to identify views held about social groups is word association. For example, a study by Gardner and Taylor [44] asked nurses to write down words associated with either English Canadians, French Canadians, or Indian Canadians. The words generated reflected racial stereotypes held about these groups. A free association task allows participants to use their own terms to generate associations. This allows the capture of qualitative insights into associations, while the established scales allow quantitative insights.

4.3 Methodology

Given the literature summarised in the present chapter and Chapter 3, there are two important gaps. Firstly, there has been limited attention given to capturing generalised usage data (i.e. unconstrained with respect to context or target population) that compares patterns for commonly-used devices. Secondly, only one study to-date has attempted to capture the perceptions individuals hold about devices [117] – this capture was limited to just one stereotype model and considered only a subset of the commonly-used devices.

This chapter aims to address these gaps for four personal computing devices in everyday use: smartphones, tablets, laptops, and desktops. Focussing on the UK as an indicative example, the research questions are as follows:

RQ 1 *How do UK adults use their everyday personal computing devices?*

- **RQ 1a** With what frequency are devices used?
- **RQ 1b** What tasks are performed on each device?
- **RQ 1c** How do users rate their confidence with the devices?

RQ 2 *What associations do UK adults make with these devices? How do they perceive them?*

- **RQ 2a** To what extent are the devices perceived as utilitarian and hedonic?
- **RQ 2b** How are the devices rated in terms of the Stereotype Content Model?
- **RQ 2c** What associations are made with the devices in free association and multiple choice association tasks?

RQ 3 *Is there a relationship between the way an individual uses devices and their perceptions of devices?*

RQ 3a Is there a relationship between responses to questions about device use and those about device associations?

RQ 3b Is there a relationship between the device used to complete the survey and responses to questions about device use and associations?

An online (crowdsourced) survey was created, with two distinct methods of data collection. Analysis was conducted using demographic groups as naturally occurring conditions. The study was given ethical approval by the University of Manchester ref: 2019-6746-11456 (see Appendix A).

The majority of the data was sourced through a combination of open and closed questions (predominantly Likert-type questions). These questions were used to establish which of the four devices individuals used with high weekly frequency, how they used these devices, and how they perceived them. Specifically, *use* was measured in terms of the activities conducted on a device, the amount of time spent on four common activities, and how confident the participant felt using the device. *Perceptions* were measured in terms of the extent to which a device was thought to be hedonic or utilitarian and how people using the device were perceived within the SCM.

Additional data was captured through two word association tasks. These tasks only investigated two devices - desktop computers and smartphones. These two devices were selected due to their placement at opposite extremes in terms of mobility, and on the basis of prior work that suggested that desktops (a stationary device) and tablets (a mobile device) are perceived as utilitarian and hedonistic respectively [88]. Since current evidence suggests that smartphones are the more dominant mobile device [61], this device was chosen as the focus rather than tablets.

The study was pre-registered at aspredicted.org[1]. Two differences to the pre-registration occurred. Firstly due to a technical error, 310 people took part instead of the planned 200. However on application of the inclusion criteria (see Chapter 4.3.1) the total sample fell to 177. Secondly, the intent to exclude participants who categorised 60% or more of the neutral words as related to ‘desktop’, ‘smartphone’ or ‘both’ in the multiple choice association task was not feasible, as piloting of the free association task suggested that 35% of the neutral words could be associated with one or both of the devices. This would mean that one incorrectly categorised word or difference in stimuli perception

[1]See `https://aspredicted.org/2s3bi.pdf`.

could result in exclusion despite participants correctly attending to the task. Instead, participants' attention to the multiple choice association task was validated by filtering out participants with at least one response time of less than 700ms (threshold determined by pilot data).

4.3.1 Participants

Participants were recruited through the `prolific.co` online crowdsourcing platform and were compensated £2.88 (approximately $3.50) for taking part. The inclusion criteria required native English speakers, over the age of eighteen, who resided in the UK at the time of survey completion, and completed the survey to the end. The last three criteria were enforced by `prolific.co`, but 25 respondents reported languages other than English as their native language or did not report their native language and were therefore excluded. In total 177 participants took part in the survey and met the inclusion criteria. The participant demographics can be seen in Table 4.1.

4.3.2 Measures and Procedure

The survey consisted of eight sections and took approximately 25 minutes to complete. The full survey is archived online at Mendeley Data [131] and can be seen in Appendix B. The sections were as follows:

1. **Demographic information:** Participants were asked to provide their age, sex, gender, ethnicity, nationality, household income, native language, country of residence, highest level of education, and the device that they were using to complete the online survey (**RQs 1, 3b**).

2. **Device frequency and confidence:** Participants were asked how many desktop computers, laptops, tablets and smartphones they used *three or more times a week*. They were further asked the make, model and main uses of the device (of each type) that they used most frequently. Participants were asked how confident they felt using each device on a Likert-style scale (6-point scale: Very confident; Confident; Neither confident or unconfident; Unconfident; Very Unconfident; I have never used this device before)(**RQs 1a, 1c**).

3. **Free association**: Participants were asked to type in as many words as they could that related to the term *desktop computer* and to *smartphone*. The device prompts were displayed one at a time (**RQ 2c**).

4. **Multiple choice association:** Participants were shown forty words, one at a time, in a random order and asked to identify whether they most associated the presented words with a desktop computer, a smartphone, both, or neither. Based on pilot data, ten words were expected to be categorised as desktop computer, ten were expected to be categorised as smartphone-related and the other twenty words were selected for neutrality (i.e. expected response neither or both). Participants' response times to categorise each word were also measured (**RQ 2c**).

5. **Utilitarian or hedonic:** Participants were asked to answer eight questions on a semantic-differential scale (5-point scale where 1 and 5 are represented by two opposite words, e.g. Wise to Foolish) indicating the extent to which they felt that each of the four devices (desktop, laptop, tablet, smartphone) were utilitarian or hedonic. The word pairs used were *useful/useless, valuable/worthless, beneficial/harmful,* and *wise/foolish* for utilitarian, and *pleasant/unpleasant, nice/awful, agreeable-disagreeable,* and *happy/sad* for hedonic. These questions were taken from [15]. The reliabilities for this scale are not clearly reported (**RQs 2a, 3a,3b**).

6. **Stereotype Content Model:** Participants were asked to indicate on a set of Likert scales (5-point scale from Not at all to Extremely) how they believe society perceives people using tablets, laptops, smartphones and desktop computers in terms of competence and warmth. The questions in this section were taken from Study 2 of [117]. The question was phrased as 'As viewed by society, how ... are people with [Device]'. The items to measure competence were *intelligent, confident, competent, independent,* and *competitive*, and *sincere, good-natured, warm,* and *tolerant* for warmth. The items in this scale have an internal reliability (α) of .90 for warmth and .97 for competence [41] [2] (**RQ 2b**).

7. **Time spent on activities:** Participants were asked, for each of the four devices, to

[2] The alpha reported for warmth is including a fifth item of 'likeable'. It is unclear in the literature why this term was not used as part of the construct of warmth in [40]. As the present study is using the measure as used by [117] and thus [40], 'likeable' is not used.

indicate the number of hours a week (rounded to the nearest whole hour) that they spent on social media, working, gaming or watching television, and emailing (**RQ 1b**).

8. **Activities:** - Participants were presented with twenty-four activities thought to commonly take place on at least one of the target four devices. These activities were generated through discussions between members of the research team and included nine activities that were thought to be utilitarian (e.g. writing reports or essays), eleven that were considered to be hedonistic (e.g. video games), and four that were considered to be neutral or context-specific (e.g. checking the weather). The activities were generated by asking a small pilot sample for the activities they tend to do on their devices. These were then discussed by the researchers and categorised in hedonic, utilitarian or neutral/context-specific. The full list of activities and their categorisations can be seen in Chapter 4.4.2. Participants were asked to indicate which of the activities they undertook *three or more times a week* on each of the four target devices (**RQ 1b, 3a**).

Participants were asked to read the participant information sheet (Appendix C) and, if they had any questions, to contact the primary investigator. They were asked to complete a checkbox consent form prior to survey completion. Participants completed the survey questions and tasks in the order presented above. Upon completing the questionnaire a debriefing screen appeared, explaining the full intent of the study and reminding participants how to contact the experimenter should they have any questions or complaints. This page also had a link to the full debriefing information (Appendix D).

4.4 Findings

The results were analysed using ANOVAs, MANOVAs, correlations and regressions. Due to the non-parametric nature of the data, alternative tests were used where appropriate. The results were analysed using either the means of the groups or the raw scores depending on the test applied. For non-parametric tests these were the ranked means or scores. Means were used for MANOVAs, and ANOVAs. Scores were used for correlations.

In order to analyse the hedonic/utilitarian scale results, and the SCM results, MANOVAs were used. This approach was chosen in order to reflect the approach utilised by Schwind *et al.*, in their study of how devices are placed within the SCM.

The specific type of regression analysis used was a backwards stepwise regression. This was chosen in order to systematically reduce the number of predictors in the model, due to the large number of predictors entered initially.

4.4.1 Demographics

In total 177 participants completed the survey and met the inclusion criteria. A full demographic breakdown of participants (sex, gender, ethnicity, income, level of education) is given in Table 4.1, together with details of the device that was used to complete the survey (reported as “Current Device").

Sex			Gender			Current Device			Age	
Female	108	(61%)	Female	110	(62%)	Desktop	36	(20%)	Mean:	32.91
Male	68	(38%)	Male	66	(37%)	Laptop	70	(40%)	Std.:	10.29
Intersex	1	(< 1%)	Non-binary	1	(< 1%)	Tablet	9	(5%)	Median:	31
						Smartphone	60	(34%)	IQR:	26-38
						Not disclosed	2	(1%)	Range	18-68

Household Income, £			Ethnicity			Education		
< 10,000	13	(7%)	Asian	10	(6%)	None	1	(< 1%)
10,000 – 19,999	28	(16%)	Black/African/			Left education at 16	11	(6%)
20,000 – 29,999	21	(12%)	Caribbean	1	(< 1%)	Left education at 18	24	(14%)
30,000 – 39,999	22	(12%)	Mixed/Multiple			Vocational Qualification	23	(13%)
40,000 – 49,999	29	(16%)	Ethnicities	7	(4%)	Undergraduate Degree	75	(42%)
≥ 50,000	50	(28%)	White	158	(89%)	Postgraduate Degree	42	(24%)
Prefer not to say	14	(8%)	Prefer not to say	1	(< 1%)	Prefer not to say	1	(< 1%)

Table 4.1: Participant demographics (n=177).

4.4.2 Device Uses

Device Use and Confidence

The mean number of devices used by participants more than three times a week was 3.67 (SD = 3.35). Broken down by device, the percentage of participants reporting use of one or more desktops, laptops, smartphones, and tablets is 50.9%, 78.5%, 98.8%, and 49.1% respectively. When compared to the data from Ofcom [103] (reported in Figure 1.1 on page 17), the present study found slightly higher uptake of all devices. This may reflect a change in market penetration for these devices between the time the data was collected in the Ofcom study and this one. However, it may also be due to the fact that the participants in this study were recruited via an online crowdsourcing platform, and are thus likely to be more inclined to utilise technology than the wider population. As such, the present study may not be representative of the whole population, as it over represents technology users. Three participants reported only having access to one type of device, and in all three cases, this device was a smartphone. This means that the participants who did not have access to a smartphone had access to two or more of the other devices. A Spearman's rank correlation reveals a significant positive relationship between age and tablet use (r_s = 0.17, p = .021) and total number of devices used (r_s = 0.16, p = .038) only. These findings indicate that *(i) smartphones and laptops have higher adoption than desktop computers and tablets, and (ii) both tablet adoption and use of multiple devices increases with age*.

Fourteen participants were excluded from the analysis of household income and device use due to not reporting their household incomes (resulting $n = 163$). Results show a significant correlation between household income and total number of devices used three or more times a week (r_s = 0.26, p = .001) as well as between income and the number of laptops (r_s = 0.26, p = .001) and smartphones (r_s = 0.21, p = .008). These findings suggest that people in *low income households have less access to laptops and smartphones, as well as to devices in general*, and is consistent with prior research indicating that low income families are more likely to share or have fewer devices [138, 153]. Income was not significantly correlated with age, indicating that the prior finding that older adults using more different devices is not due to them being in wealthier households.

For each of the four device types, between one and seven participants reported never using that device and as such the number of participants included in the analysis of confidence using desktops, laptops, smartphones and tablets was 176, 174, 173, and 170 respectively. A Spearman's rank correlation was conducted to investigate the relationship between age and confidence using devices. There was a significant negative relationship between confidence in using a smartphone and age (r_s = -0.27, p $<$ .001). There were no significant relationships between age and confidence using the other three devices.

Analysis of the relationship between number of devices and confidence using a particular device found significant relationships for desktops (r_s = 0.23, p = .002), laptops (r_s = 0.27, p $<$ .001) and tablets (r_s = 0.29, p $<$ .001). Despite similarities in input mechanisms of desktops and laptops, there was no significant relationship between the number of desktops frequently used and confidence in using a laptop (r_s = 0.01, p = .858) or vice versa (r_s = 0.06, p = .452). Similarly, despite both smartphones and tablets being operated via touchscreen, there was no significant relationship between the number of smartphones used frequently and confidence in using a tablet (r_s = -0.03, p = .696) or vice versa (r_s = -0.04, p = .641). These results suggest that confidence using a smartphone decreases with age, and that the frequency with which laptops, desktops and tablets are used is related to increased confidence using them. *Using a device with a certain interaction method does not equate to confidence using a different device with the same interaction method* suggesting that other factors influence an individual's confidence using a device. However, it should be taken into account that these correlations have small effect sizes.

Time Spent on Activities

A Scheirer-Ray-Hare test was conducted to investigate the effect of device and age on the amount of time spent on activities. Significant results were then further examined using a Tukey post-hoc test. Three participants gave negative time values and were thus excluded from the analysis (resulting $n = 174$). Participants were split into four age groups based on the median and interquartile range: Group 1 (*n=43*), age $18-25$ years; Group 2 (*n=44*), age $26-30$ years; Group 3 (*n=45*), age $31-38$ years; Group 4 (*n=42*), age 39 years and above.

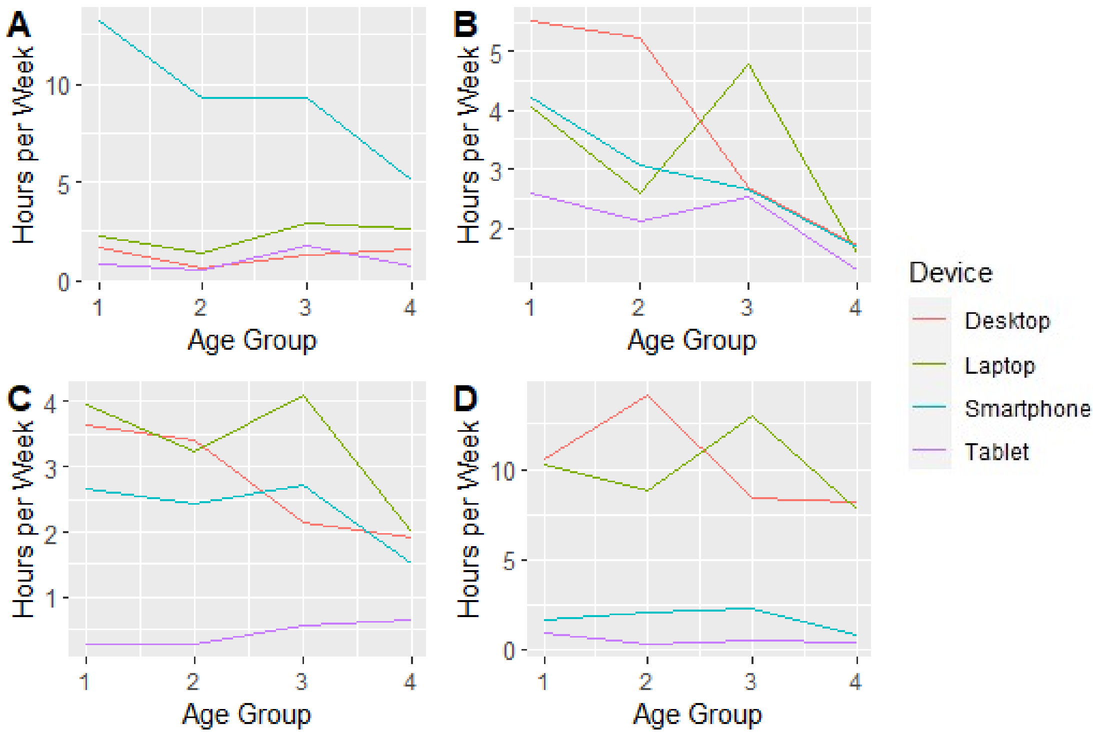


Figure 4.1: Interaction plots showing the mean number of hours each age group spent per week using their device for A) social media, B) gaming or watching television, C) emailing, D) working. The device being used is represented by the colours of the lines.

Main effects of device were found for time spent on social media ($H(3,695)$ = 244.53, $p < .001$), time spent working ($H(3,695)$ = 137.06, $p < .001$), and time spent emailing ($H(3,695)$ = 124.53, $p < .001$). The post hoc tests indicated that *less time is spent on tablets, laptops and desktops for social media than smartphones* (p = .001), *less time is spent working on tablets and smartphones than desktops and laptops* (p = .001) *and less time is spent emailing on tablets than the other three devices* (p = .001). While time spent gaming or watching television showed a significant main effect with device ($H(3,695)$ = 26.77, $p < .001$), the post-hoc analysis did not find any significant differences between the devices. This may be a false negative due to Tukey's post-hoc over-correcting, however it may also suggest that the differences between the devices in time spent gaming or watching television are negligible despite the initial ANOVA suggesting there is a main effect.

Main effects of age were found for time spent gaming or watching television ($H(3,695)$ = 13.01, p = .005). The post-hoc analysis suggested that *the oldest age group spent significantly less time watching television or gaming than the youngest age group* (p = .002). Time spent on other activities did not have a significant main effect of age.

Time spent on all activities showed significant interactions between device and age group. For social media the interaction was $H(9,695)$ = 27.25, p = .001. The interaction plot suggests that *time spent on social media on a smartphone decreases as age increases*. Despite the significant interactions between age and other activities, no clear trends emerged after reviewing the interaction plots (see Figure 4.1). This suggests that while there may be an effect of age and device on how much time is spent on different activities, the exact nature of this effect is not clear and requires more investigation.

Activities

Participants were asked how many different activities they undertook on each device. The results in Figure 4.2 suggest that there are large variations in the number of different activities that people use their devices for.

There was a significant positive Spearman's r correlation between the number of activities undertaken on a tablet and age (r_s = 0.19, p = .015), supporting the earlier findings regarding overall use. A Kruskal-Wallis analysis found a significant effect of

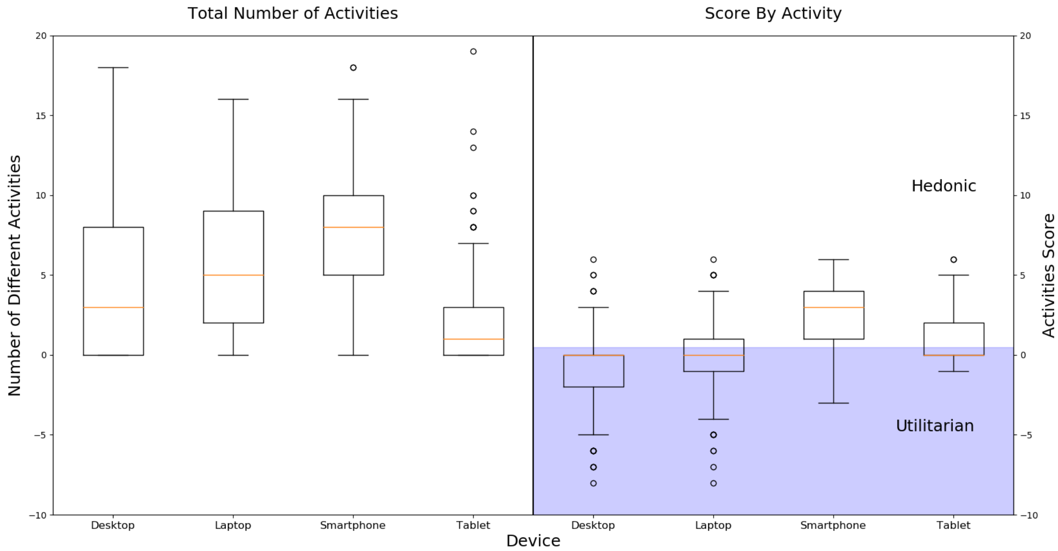


Figure 4.2: Boxplots showing amount of activities undertaken on each device in total and when utilitarian activities are scored as -1 and hedonic activities are scored as 1. The whiskers represent 1.5 times the interquartile range.

device on number of activities undertaken ($\chi^2(3) = 39.61$, p $<$ .001). A Dunn post-hoc test with a Bonferroni correction suggested that *more activities are undertaken on smartphones than laptops* (p $<$ .001), *on laptops than desktops* (p = .005) *and on desktops than tablets* (p $<$ .001).

Each activity was then categorised using the definition given in Chapter 4.2 as 1, -1 or 0 according to whether the activity was hedonic, utilitarian or neutral respectively[3]. This resulted in an activity score of between -9 and 11 for each participant. Using a Spearman's r correlation, it was found that activity scores for smartphones were negatively correlated with age ($r_s = -0.18$, p = .019). This negative correlation suggests that *younger adults utilise their smartphones for hedonic activities to a greater extent than older adults*.

[3]Utilitarian activities: emailing; meetings; reading for work; making notes; banking; reading/watching the news; report writing; spreadsheets; presentations.
Hedonic activities: social media; messaging; gaming; shopping; writing for pleasure; reading for pleasure; watching television; listening to music/podcasts; browsing; online dating; taking/viewing photos.
Neutral/Context-specific: fitness or health; journey planning or navigation; checking the weather; arranging travel.
These are generalisations and may not be true for all participants.

A Kruskal-Wallis analysis found a significant effect of device on the activity score ($\chi^2(3) = 275.39$, $p < .001$). A Dunn post-hoc test with a Bonferroni correction suggested that more hedonistic activities are undertaken on smartphones than tablets ($p < .001$), on tablets than laptops ($p = .002$) and on laptops than desktops ($p = .008$). Together, these results suggest that *smartphones are most heavily used for hedonic activities, whilst desktops and laptops are more frequently used for utilitarian activities*.

4.4.3 Device Perceptions

Free Association

167 participants provided data for the desktop free association task. Some data did not save correctly and was not included in the analysis. In total 585 different words related to desktops were generated (mean = 10.14, SD = 8.07). For the smartphone association task, 163 participants provided data. In total 550 different words related to phones were generated (mean = 9.65 SD = 7.77).

Using the Natural Language Tool Kit library [17] the words were categorised into word classes (e.g. nouns, adjectives). Due to the lack of context in the dataset (i.e. the words were not generated as part of a sentence) some words were incorrectly classified and these words were reviewed and re-categorised appropriately. The most frequently generated words and their word class can be seen for each device in Table 4.2.

For desktops, the word ‘work’ was generated by 45 participants (27%), indicating that this device is associated with use in workplaces and employment. Similarly, ‘desk’ was generated 29 times. Smartphones seemed to be associated with brands and branded products frequently. ‘Apple’ was generated 20 times (as well as the specific make of device ‘iPhone’ being generated 19 times) for smartphone compared with 8 times for desktop. The brand ‘Facebook’ was associated with smartphones more than desktops (19 and 4 generations respectively). The association of smartphones with brands may suggest that the type of phone a person has is more important than the type of desktop someone has. Relatedly, ‘social media’ was also a frequent association with smartphones (19 generations) suggesting that *concepts that are considered social are associated with smartphones*.

For both desktops and smartphones nouns were the most common word class (327 and 281 respectively) followed by adjectives (154 and 154) and verbs (77 and 83). The top nouns associated with desktops tended to be objects that are used with the device – such as ‘keyboard’ and ‘mouse’ – while for phone the top nouns were ‘camera’ – a part of the device – and ‘text’, which may refer to sending a text message. However, ‘text’ could also be considered a verb as it is common for people to say phrases such as ‘I will text you’.

More verbs were reported for smartphones than desktops. The top verbs for smartphones tended to relate to communication such as ‘calls’ (28) and ‘calling’ (3), as well as words such as ‘messaging’/‘messenger’ (4 and 4 generations). ‘Email(s)’ was identified by 16 participants for smartphone compared with 11 for desktops. The most frequently generated desktop verbs were ‘typing’ (8) and ‘learning’ (4). As such, *verbs relating to smartphones appear to be communication orientated.*

Many adjectives related to desktops can be considered negative when referring to technology (e.g. bulky, slow). Conversely the *adjectives pertaining to smartphones are more positive*, suggesting this device is seen more favourably by users. In addition to the adjectives relating to smartphones in Table 4.2, ‘convenient’ (18) and ‘modern’ (15) were frequently generated.

%	**Desktop** *($n = 167$)*	**Smartphone** *($n = 163$)*
41+	*keyboard(85)*; *mouse(81)*; *monitor(51)*	
21-40	*screen(46)*; *work(45)*; *internet(37)*	*apps(44)*; small(36)
14-20	large(29); *desk(29)*; *windows(27)*; bulky(25)	*camera(32)*; **calls(28)**; portable(26); *text(26)*; easy(25); *game(25)*; useful(24); fast(23)

Table 4.2: Words generated by ≥14% of participants, for desktop computers and smartphones, in the free association task, grouped by percentage of the participants generating the word. The number in brackets indicates the actual number of participants generating the given word. Words in bold face are verbs, italicised words are nouns and other words are adjectives.

Multiple Choice Association

As described in Chapter 4.3, participants whose response times were less than 700ms were excluded from the analysis (resulting n = 149). Not all participants categorised all

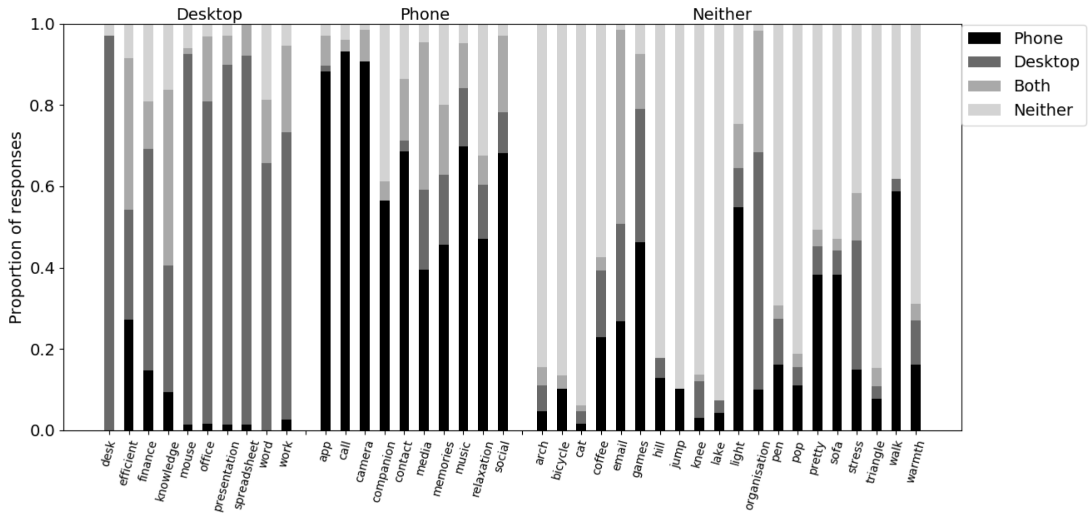


Figure 4.3: The proportion of times each stimulus word was categorised as being related to a desktop, smartphone, both or neither. The words are grouped by the expected categorisation, while the colours represent the actual categorisations by participants ($n = 149$).

the words. The categorisations can be seen in Figure 4.3. Participant responses largely aligned with the device-word classifications that emerged from pilot data, however there were some exceptions.

‘Organisation’ was expected to be neutral but participants largely related the word to desktops. This does, however, fit the trend that words relating to work (including the word ‘work’ itself) are associated with desktops. While ‘efficient’ was expected to relate to desktops, participants categorised it frequently as related to both desktops and smartphones. Where participants did indicate ‘efficient’ as being associated with a single device rather than ‘both’, smartphone and desktop were selected with similar frequencies. The words generated from the free association task in Table 4.2 show that smartphones were associated with being ‘useful’ and ‘convenient’. This combination of perceived utility and convenience is likely to have contributed to the resulting associations with efficiency, despite the device’s inherent limitations (e.g. screen size, lack of physical keyboard).

Another unexpected result was the word ‘walk’ being related to smartphones. This may be due to phones being ‘portable’ (a finding from the free association task), suggesting the way devices are used impacts how they are stereotyped.

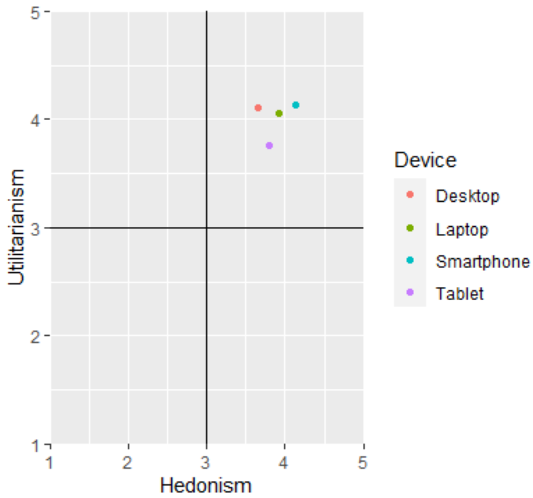


Figure 4.4: The mean, hedonic and utilitarian ratings for desktops, laptops, smartphones and tablets ($n = 175$).

Utilitarian or Hedonic

A 4(current device) by 4(device being rated) MANOVA was conducted to investigate perceptions of hedonism and utilitarianism. Two participants did not disclose which device they completed the survey on and were not included in this part of the analysis ($n = 175$). The data was non-parametric, however due to a lack of suitable alternative analysis methods, the sample size meant it was appropriate to use parametric tests. There was a significant effect of device on perception of hedonism and utilitarianism (F(8, 709) = 14.95, p $<$ .001; Pillai's trace = 0.14).

A follow-up two-way ANOVA found that there was a significant main effect of device on hedonism (F(3.00, 513.00) = 16.27, p $<$ .001, η_p^2 = .087). A Tukey post-hoc comparison found that for hedonism, *smartphones were considered significantly more hedonic than tablets* (p = .001), *and desktops* (p = .001). *Laptops were also significantly more hedonic than desktops* (p = .008). A significant effect of device on utilitarianism was also found (F(3.00, 513.00) = 14.30, p $<$ .001, η_p^2 = .077). Post-hoc comparisons showed that *tablets were perceived as significantly less utilitarian than the other three devices* (p = .001). The mean scores for each device are plotted in Figure 4.4. These utilitarian/hedonic device associations interacted with the device used to complete the survey (see Chapter 4.4.4).

Together, the results from the two word association tasks and the hedonism/utilitarianism questionnaires suggest that *desktops have stronger associations with some work-related terms but smartphones, laptops and desktops all have comparable levels of utilitarian association*. By comparison, there appears to be a greater variety in how hedonic devices are perceived to be. The free association task generating social- and leisure-related words for smartphones is consistent with this device being the most hedonic. Tablets appear to have different associations to all other devices, scoring lowest on utilitarianism and second lowest on hedonism.

Stereotype Content Model

A MANOVA was used to investigate how desktops, smartphones, tablets and laptops were perceived within the SCM. The SCM was also investigated in terms of the current device, and as such the sample size was 175. The data was non-parametric but the sample size meant it was appropriate to use parametric tests. There was a significant effect of the device being rated on perception of competence and warmth ($F(9, 689) = 9.63$, $p < .001$; Pillai's trace = 0.11). A follow up ANOVA suggested that there was a significant effect of device on competence ($F(3, 522) = 12.98$, $p < .001$). Tukey's post-hoc tests suggested that laptop users were perceived as significantly more competent than desktop (p=.002) and tablet (p=.001) users, and smartphone users were perceived as significantly more competent than tablet (p =.029) users. A further ANOVA suggested that there was a significant effect of device on warmth ($F(3, 522) = 3.78$, $p = .010$), however, post-hoc analysis did indicate any significant differences between the individual devices. The placement of the devices within the SCM can be seen in Figure 4.5. While Schwind *et al.* [117] also found a significant difference between smartphone and tablet users with regard to competence, this *prior work indicated a relationship in the opposing direction*, whereby tablet users were seen as the more competent group.

The differences between the present study and that of Schwind *et al.* were 1) their population was German while the present sample were UK residents; 2) they presented pictures of the devices while the present study did not; and 3) participants in this study answered more prior questions than those of Schwind *et al.* and the SCM question set was the only set of questions relating to devices in society, rather than to the individual's perceptions of the devices themselves. These factors may have either influenced how

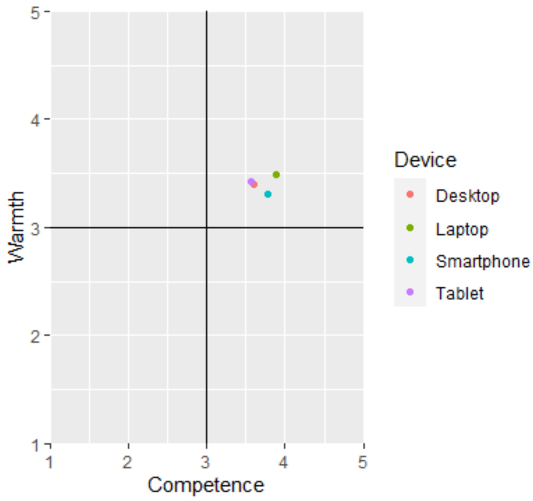


Figure 4.5: The mean ratings for desktops, laptops, smartphones, and tablets within the SCM ($n = 175$).

participants rated the devices or, in the case of difference three, fatigued participants to the point where they did not attend to the questions fully. Overall though, this suggests that the perceptions of devices in regards to the SCM may vary, and more research may be required to understand why these differences occur.

4.4.4 Use and Perception Interactions

A backwards stepwise regression was conducted to investigate the effects of activities undertaken on a device, and the device being used, on utilitarian or hedonic ratings for the device the activities were undertaken on. The activities defined as utilitarian or hedonic can be seen in Chapter 4.4.2.

Utilitarian ratings were found to be predicted by the device being used to complete the survey ($\beta = -.105$, p = .004), whether emailing was done on the device ($\beta = .118$, p = .004) and whether the device was used for banking ($\beta = .187$, $p < .001$). The overall model fit was R^2 = .074. Hedonic ratings were found to be predicted by the device being used to complete the survey ($\beta = -.088$, p = .015), and whether the device was used for social media ($\beta = .195$, $p < .001$), gaming ($\beta = .080$, p = .038), watching television ($\beta = .092$, p = .030) and browsing ($\beta = .093$, p = .028). The overall R^2 for the model was .123.

For both models the values of the device being used to complete the survey were desktop = 1, laptop = 2, smartphone = 3, and tablet = 4. Therefore the negative beta for current device indicates that higher utilitarian and hedonic scores were given when using desktops and laptops and the scores were lower when using smartphones and tablets. Positive beta values for the activities indicate the utilitarian or hedonic ratings of a device were higher when that activity was undertaken on the device.

These findings suggest that the way a device is used can be predictive of how that device is perceived. However, the overall predictive power of the models was fairly low, suggesting that more research is needed to fully understand the impact of how devices are used on how they are perceived.

Effect of the Current Device

A MANOVA found a significant effect of device being used to complete the questionnaire for utilitarian and hedonic perceptions (F(24, 2133) = 2.05, p = .002; Pillai's trace = 0.07).

A follow-up two-way ANOVA showed a significant main effect of the current device on hedonic perceptions (F(3.00, 171.00) = 3.42, $p < .001$, $\eta_p^2 = .056$). Post-hoc tests demonstrated a significant difference between participants using tablets and participants using desktops (p = .004), laptops (p = .001), and smartphones (p = .014). *Tablet users generally rated all devices as less hedonic than other users, smartphone users rated devices second least hedonic, and laptop users rated them most hedonic.* An ANOVA investigating the effect of current device on perceptions of utilitarianism also found a significant effect (F(3.00, 171.00) = 3.07, $p < .001$, $\eta_p^2 = .051$). Post-hoc analysis showed that *laptop users rated devices as significantly more utilitarian than smartphone users* (p = .020) *and tablet users* (p = .003) *but desktop users did not rate devices significantly differently from any other users*.

Significant interactions between the device being rated and the device being used were also found for hedonism (F(9.00, 513.00) = 6.35, $p < .001$, $\eta_p^2 = .100$) and utilitarianism (F(9.00, 513.00) = 5.08, $p < .001$, $\eta_p^2 = .082$). The interaction plot indicates that desktop users rated desktops as the most hedonic, as did smartphone users for smartphones, and tablet users for tablets. Laptop users rated laptops the second most hedonic after smartphones. The same trend was seen for utilitarianism except for laptop users who rated laptops as the most utilitarian.

Overall, *participants rated the device they used higher on both scales* suggesting that they may have a bias towards the device they use, particularly given that in the present study the participants chose which device to use. However, there were also significant differences in their ratings of the other devices, indicating a trend in these users perceptions, and supporting the findings regarding current device in the regression analysis. This may be due to people with similar opinions opting to complete the study on the same device but could also be an indication that *the device used to complete a task may impact the task performance*. However, the findings regarding the effect of current device were not replicated with regard to perceptions within the SCM (the effect of current device on competence and warmth in the MANOVA was not significant). This may be because participants were focusing on a fictional user as well as the device, thus reducing the impact of the current device.

4.4.5 Summary

Key findings of the reported data can be summarised as follows:

- Using a device correlates with confidence in using that device but not confidence in similar devices.
- The amount of time spent undertaking activities on a device varies with age, the device being used or both.
- Activities and words with social connotations are associated with smartphones.
- Desktops have stronger associations with some work-related terms but smartphones, laptops and desktops all have comparable levels of utilitarian association.
- In the SCM, tablet users were seen as less competent than smartphone users, contrary to prior findings.
- The device being used might influence performance of the tasks being executed on them.

4.5 Implications

This section discusses the implications of these results for the research questions outlined in Chapter 4.3.

4.5.1 Research Question 1: How do UK adults use their everyday personal computing devices?

The device usage data is consistent with several existing findings suggesting that the present sample is valid and lending credence to the additional findings presented in this chapter. Evidence of reduced access to devices based on household income indicates that observations made in prior research persist [153], suggesting a continuing need to address the challenges and impact of a digital divide. While time spent on social media was found to largely occur on phones, there was a further finding that younger adults used their phones for social media more than older adults. Likewise, while all devices were used similarly to play games or watch videos, it was found that people over 39 did so to a lesser extent than those under 26. This may be a reflection of general trends for increasing social media usage and gaming in younger age groups [5, 47].

There is some conflict here with prior findings that indicate smartphones are preferred for shorter tasks [73], but this may be explained by the recent emergence of streaming platforms (e.g. Netflix) potentially influencing younger participants' media consumption, encouraging smartphone-based TV-viewing whilst on the go. The finding that older adults have a preference for more devices contradicts existing research [138]. It is possible that in the decade since this prior research was undertaken, the attitudes of older participants towards technology uptake have changed. However, it should be noted that 'older' participants is relative to the rest of the sample, and the oldest age group is younger than that which would be considered an older population in other research. Furthermore, the participants were recruited via an online platform, and may therefore be more open to using technology than others.

While it has been suggested that there is no overall effect of age on smartphone usage [87], the results suggest the activities that different age groups undertake vary. Younger adults appear to undertake hedonic activities more than older adults, who undertake more utilitarian activities. Thus, hedonic apps such as games are best tailored towards younger audiences, while older adults may prefer more practical apps. However, determination of exactly what constitutes hedonic and utilitarian is likely to vary with individual interests and personality traits.

This research shows that confidence using one device does not translate to confidence using another device that has the same interaction methods. For example, smartphones and tablets both use touch-screens and provide access to the same software. While use of a tablet correlated with confidence in tablet use, it did not correlate with confidence in smartphone use. Although this may be due to screen-size — a smaller touchscreen may be more difficult to use — a difference was also found between desktops and laptops. This suggests then that these pairs of devices are not seen as interchangeable despite similarities in input mechanisms and applications. Other factors must therefore impact individuals' confidence when interacting with these devices. The findings suggest that devices with similar interaction methods do carry different stereotypes, which could be a contributing factor in the observed failure for confidence in one device to translate to others. For example, the hedonistic associations of a smartphone that are absent when using a tablet may override any similarities in interface design, leading to a very different user experience, thus reducing an individual's confidence using a device, despite the lack of change in input mechanism. As such designers of new technology should not assume that the prevalence of an interaction method in society erases the learning curve that occurs when adopting new devices.

4.5.2 Research Question 2: What associations do UK adults make with these devices? How do they perceive them?

Laptops, desktops and smartphones were considered to be similarly utilitarian, while tablets were seen as significantly less utilitarian than other devices. All four devices differed in the degree to which they were considered hedonic, suggesting that this may be a more discriminatory measure. Considered together, these results show that smartphones and tablets have markedly different profiles to each other, whilst laptops and desktops are more similar to each other (Figure 4.4). This finding potentially validates current approaches to software development, which often involves the production of a common interaction experience for laptops and desktops and two further separate interfaces for tablets and smartphones (assuming all four are targeted by the same application). The placement of smartphones as the most hedonic device is consistent with the wealth of 'casual games' targeted at this platform [75]. This therefore suggests that user interface designers should continue to differentiate based on device (e.g. responsive design: [49]), particularly when developing applications for tablet devices or those with strong hedonic characteristics.

The results further suggest that smartphone users were seen as more competent than tablet users, despite previous research by Schwind *et al.* [117] suggesting the opposite to be true. According to Fiske *et al.* [40] low competence is often associated with groups such as older adults. Given that the present results suggest the use of tablets is more prominent in older adults, the findings of tablet users being less competent may be a reflection of the users generally being considered older. As noted in Chapter 4.4.3, there were some differences between the present study and Schwind *et al.*, which may account for the difference in findings. In particular, the context in which the questions were asked was different. This may have led participants to not actually consider the devices in the context of other people, although the similar competence score between tablets in the present work and older adults in prior work suggests other people may have been considered.

An alternative explanation is that differences may have occurred due to stereotypes being variable across populations, or being context dependent. However, the lack of replication may also suggest that the findings may not be as robust as previously thought. This highlights the importance of replication within HCI, as in wider psychology [94].

In the free association task, the verbs generated in relation to smartphones (e.g. 'calls', 'messaging') tended to be related to communication. Smartphones were also related to social media frequently, suggesting that smartphones are associated with social concepts. This reflects the patterns of use reported in the present study, where time spent on social media was highest on smartphones. 'Work' was largely related to desktops in the free association task and participants reported spending more time working on desktops and laptops. Therefore, the free association terms appear to broadly reflect the trends in how devices are used. This indicates that the associations people make with devices are congruent with the way they use their devices. It may then be beneficial for users to use devices in ways that are congruent with their associations, to reduce cognitive load. However, the current results show correlations and further research would be required to understand whether there is a causal relationship between use and stereotypes.

Given the evidence that devices have distinct stereotypes held about them, it may be beneficial for designers to consider the use of 'device personas'. Personas are traditionally used in HCI to elicit an end user for whom the device could be optimised. Including considerations of the device's persona in this process may aid in the design of a given device's software, and allow for optimisations to be made to suit the device's stereotypical purpose.

Of the four subscales used to measure utilitarianism (see Chapter 4.3.2), only one term (*useful*) was generated in the free association task, and only for smartphones not desktops when the expectation would be for these terms to be generated fairly equally for the devices, given that these devices were not significantly different in this measure. Subscale terms for hedonism were not generated in the free association task. It is possible that terms used on the subscales would emerge for all devices in a larger free association dataset. However, the present findings suggest that subscale terms may not reflect natural language use, and as such, may not be fully capturing what makes a device utilitarian or hedonic to the user. Therefore, it may be appropriate to review scales in reference to associations made in free association tasks, so the measurements used reflects natural language.

However it should be noted that the analysis performed on the results for the SCM and hedonic/utilitarian scale was a MANOVA. This was chosen to replicate the analysis method used by Schwind *et al.*. While the rest of the analysis was conducted using non-parametric measures, there was no suitable non-parametric alternative to the MANOVA. Given this, further investigation is required to fully understand the ways everyday devices are perceived in terms of the SCM and their hedonic/utilitarian qualities.

4.5.3 Research Question 3: Is there a relationship between the way an individual uses devices and their perceptions of devices?

The regression models, while lacking strong predictive power, do give an initial indication that the way a device is used relates to how the device is perceived, supporting the observed trends between device use and the free association tasks. The device used to complete the study was a predictive factor in the model. It was also associated with the way that devices were rated in terms of hedonism and utilitarianism. Specifically, three clear patterns of behaviour were demonstrated.

1. The device used to complete the survey was rated as both more hedonic, and more utilitarian, than other devices.

2. Tablet users rated devices as less utilitarian and hedonic than other device users.

3. Participants using a desktop give significantly higher ratings of hedonism than tablet users.

Whilst the results agree with Liu and Wang [88] that behaviour differs according to the device used, Liu and Wang also suggest that the ratings of the hotels in their study may have been impacted by the device being used. Specifically, they suggest that the decision made is mediated by the extent to which the participant considers the device to be hedonic or utilitarian. The participants of Liu and Wang were assigned the device to use, which may elicit different effects to those seen in the present study where participants had free choice over their device. Understanding whether trends observed in this study exist in the circumstances of studies like Liu and Wang is an important step, as it may be the case that the device a study is conducted on is a confounding variable in the research.

There appears to be a relationship between an individual's choice of device when participating in research and how they perceive other devices. This has implications for research comparing devices as the results may be skewed in favour of the device being used. There is also the potential that questions about a device that are answered on said device will be more positive than if asked on a 'neutral' device. This may then lead to false positives which are not representative of opinions when this bias is removed. Overall, the results suggest that using a device can prime people to give particular responses and as such change interaction. Unlike measures of utilitarianism/hedonism, the device being used did not appear to impact the placement of devices within the SCM. Therefore, the device used to complete a survey may not universally affect all measures, and more work should be conducted to clarify when and why this effect occurs.

The higher ratings of the user's own device for hedonism and utilitarianism could be an indication of an ingroup/outgroup effect being elicited. As noted in Chapter 4.1, this bias can be formed along arbitrary lines such as a preference for people who use the same device as you. However, the lack of consistency with this higher rating for other measures such as the SCM suggests that this may not be a robust effect. Despite this, given the prevalence of brand associations in the free association task for smartphones, there may be implications for the potential of device ingroups and outgroups, whether that be over the device itself or the brand of device, and in particular, smartphone brand. This may be important in contexts such as schools, where individuals may be subject to bullying by their peers if they are perceived to not have the same technology or the most 'popular' brand. This issue may also be exacerbated by the digital divide which the present findings suggest is still prevalent.

However, the findings from the backwards stepwise regression should be considered critically. This method of regression removes predictors from the regression, until only the most significant predictors are left. As this process is done without input or monitoring of the process by the researcher, it can be unreliable [124]. This method was chosen in order to reduce the large number of predictors. However, future work should explore the use of more reliable regression approaches, such as linear or logistic regressions.

4.6 Limitations and Future Work

The reported study uses a variety of methods for capturing device use and perception data. Due to the inclusion of the word association task, all participants were required to be native English speakers, to ensure the meanings of words generated were not influenced by linguistic differences. As such, these findings may not be representative of those outside of the sampled population, although they may hold for populations with similar linguistic and/or cultural background (e.g. native English speakers in the United States of America – [25]). Replication of this study in other languages and countries would establish the degree to which behaviour patterns and perceptions transfer across cultures and societies, and may reveal new phenomena specific to other settings. This may also aid understanding of the differences found between the present work and Schwind *et al.*, and give insight into the robustness of device stereotypes across populations.

The present study uses online crowdsourcing for data collection, and prior literature indicates this to be a valid method of executing questionnaire-based research [22]. However, online study execution does have the potential to introduce bias, and it is possible that the sample may have been more frequent and confident users of technology than a sample recruited offline. The ratings of elements, such as confidence using devices, could be higher than those in the general population. That said, any bias is likely to be consistent across groups within the sample (e.g. between older and younger adults) meaning that the reported relative differences should be unaffected.

A potential limitation is also the categorisation of the activities reported in Chapter 4.4.2. There are potential individual differences in which activities are considered hedonic or utilitarian. As such the present work should be considered a starting point for work in this area. Future research may wish to elaborate on this work by asking participants to rate the activities themselves, as well as reporting which they undertake to better control for individual differences.

The present study largely reports correlational and naturally-occurring trends. As such, understanding of the impact of device stereotypes and the relationship between use and device associations is at an early stage. Prior work indicates the potential for technology stereotypes to impact decision making [88], but the role of these stereotypes has yet to be considered for a wider set of user interactions and cognitive domains. Future work may therefore investigate the role of technology stereotypes both on interaction itself, on task performance, and on the underlying cognitive processes. For example, future studies might consider whether more hedonic iconography is preferred, and leads to better performance, on a device with stronger hedonic associations (and likewise for utilitarianism). Other studies could consider whether exposure to hedonistic stimuli on a device with stronger hedonistic associations improves response times, attention, or memory for the stimuli.

The idea of 'device personas' is also discussed based on these findings. Future research may wish to expand on this idea by collaborating with designers and exploring how these personas may be used in conjunction with traditional personas as well as alone, and whether this lends itself to improved experiences for users.

The results suggest that the device being used to complete our questionnaire had an impact on some measures. Given the wealth of research being conducted on both desktop/laptop computers and on smaller mobile devices, further understanding of this phenomenon could be beneficial. In clinical domains in particular, there have already been attempts to validate technology-delivery of paper-based scales (e.g. [37]). However, these typically target a single device type (e.g. through use of a smartphone app), or overlook differences that may occur as a result of device-agnostic delivery (i.e. a web form). Existing literature is therefore unable to quantify when and why the device used to respond to a questionnaire will shape the captured responses. Future work may benefit from running a battery of measures across multiple devices and investigating what types of measure are impacted (e.g. Likert-style measures), as well as the measurements themselves.

Finally, the collected free association data may itself be of value for future work and is therefore archived on Mendeley Data [131]. Research that captures free association data for the two studied devices would further enhance this dataset and current understanding of individuals' device associations. The methodology could also be used for other devices and across other populations. The findings suggest some differences between the terms used in pre-existing scales and the similar concepts emerging from the word association. Further comparative studies with free association data could be of value in the revision of existing scales for hedonism/utilitarianism and competence/warmth. In particular, mapping these scales to naturally-occurring language may improve the usability of the scales for research participants, and may also improve accuracy of the scales in capturing a concept. Other concepts emerging from the free association dataset are not captured at all within these scales (e.g. mobility, association with brands) and thus a more comprehensive exploration of free associations may lead to the emergence of new scales.

4.7 Conclusions

The study reported in this chapter suggests that associations made with devices reflect broader trends in use. Specifically, desktops were frequently associated with work, which mirrors participants reporting that they spend more time working on desktops and laptops than smartphones and tablets. Similarly, smartphones were associated with words that could be considered social, reflecting the longer time spent on social media on smartphones compared with other devices. The results also indicate that the device used to complete the research itself impacted utilitarian and hedonic perceptions, but this effect was not observed for the SCM, suggesting not all stereotypes are influenced by the same factors. Finally, the present study failed to replicate previous findings that placed smartphones and tablets within the SCM – this inconsistency could suggest an instability in stereotypes across populations and contexts.

Given these observations, there is a need for further research to explore the robustness of stereotypes associated with technology and technology use, as well as research into how these stereotypes impact behaviour in device users. Further research into how device stereotypes may be leveraged in the design of future interfaces and devices would also be beneficial. Additional research that develops understanding of the individual properties of devices that appear to contribute to these stereotypes could aid in the development of novel devices that are targeted towards specific tasks. This may in turn reduce cognitive load for users, thus improving their efficiency at tasks congruent with their device.

Given the findings in this chapter regarding the potential impact of stereotypes held about the device being used influencing task performance, Chapter 5 aims to investigate whether these associations can influence processing speed. This is done by utilising the free association data generated in this chapter as stimuli in a Stroop colour naming task, to investigate whether stimuli associated with the device being used influences the speed at which information is processed.

Chapter 5

Information Processing

This chapter investigates the effect of device stereotypes on information processing speeds, with the overall aim of establishing whether the stereotypes and associations related to the device being used have an impact on information processing. The results of the free association task in Chapter 4 generated two sets of words– one set associated with desktop computers, and the other associated with smartphones. This chapter investigates the use of these as stimuli in an information processing task. The hypotheses suggested that the time taken to react to the stimuli would be affected when the word shown was related to the device being used.

Participants used either a desktop/laptop or smartphone to view stimuli, and were asked to identify the colour of the font that these stimuli were presented in, and response time was used as an indicator of effects on information processing (a paradigm known as the stroop test [133]). This was tested across two experiments because the pre-registered study used directional hypotheses. The first experiment hypothesised that stimuli related to the device being used would elicit a faster response time (RT), and the second that this would lead to a slower RT. The results suggested that words related to the device being used led to longer response times, indicating that stimuli associated with the device being used can slow down information processing.

5.1 Measuring Information Processing

The stroop test, named after the author J. Ridley Stroop, was published in 1935 [133], and gives insight into the extent to which interference has occurred during information processing, which Stroop tested across three experiments. Interference in this context refers to the delay in information processing, when a stimulus is processed slower because it is related to the task, and requires more attention to process. In their first experiment, Stroop asked students to read aloud a list of colours printed either in black, or printed in a range of colours other than black or the colour of the word (i.e. red was never printed in red ink). They found that the list printed only in black was read faster than the colour list. In the second experiment, participants were asked to say the colour the that word was printed in rather than the word that was written. Participants' speed of response was compared to the speed taken for them taken to name the colour of squares. Stroop again found that it took longer for participants to say the colour of the word, than to say the colour of the square. The final experiment in Stroop's series examined the effect of practice on the stroop test, where practising over consecutive days improved response times (RTs).

In each experiment, longer RTs were indicative of interference, caused by the written word interfering with the processing of the ink colour, or the ink colour interfering with the processing of the written word. In subsequent and current research, the term "stroop test" is typically used to describe a study using a methodology similar to that of Stroop's experiment two, where people must identify the colour of the ink. Stroop's experiment one, is now referred to as an 'inverse stroop test', where people identify the written word not the ink colour.

Following the advent of the stroop test, further research was conducted into what can cause interference. Klein [78] found that words that were not colour names, but that implied a colour, led to a delay in RT. For example, the word 'sky', while not a colour in itself, is closely related to the colour blue, and thus causes interference when used as a stimulus in the stroop colour-naming task. Similarly, linguistic features of the stimuli can cause interference. For example, longer words can cause delays in processing as words with more letters take longer to read [83]. The frequency with which a word is used in everyday language can also cause interference, as less common and thus unexpected words take longer to process [78, 83]. When selecting stimuli for a stroop test, researchers must therefore take care to control for unintended associations and unwanted differences in word length, complexity or familiarity. In the following sections, three factors which can impact RTs in a stroop task are explored.

5.1.1 Emotional Stroop

The emotional stroop task suggests that when a word is emotionally salient, the response time in a stroop task increases. This was demonstrated by Watts *et al.* [146], where emotional words such as 'fail' and 'death' led to slower RTs for naming the colour ink compared to neutral words such as 'clock' and 'gate'. They further found that people with a fear of spiders showed similar delayed responses for words semantically related to spiders such as 'creepy' and the word 'spider' itself, whereas those not afraid of spiders did not show a delay in RT. Furthermore, Siegrist [122] found that when showing people taboo words such as 'naked', there were delays in response times, suggesting they elicited a stronger emotional response than neutral stimuli. This emotional response then leads to longer response times, as it takes the participant longer to process the information.

Emotional stroop effects have also been found when the stimuli pertains to specific clinical groups. Subsequent research investigated emotional stroop effects for many mental health conditions including anxiety, depression, and post traumatic stress disorder, the findings of which were summarised in a review by Williams *et al.* [149]. The review showed that there were attentional biases in clinical populations for words pertaining to their condition, which led to longer response times in the stroop task.

5.1.2 Stereotype Stroop

In a stereotype stroop test, there is an effect where *congruent* stimuli leads to faster response times. For example, in a standard stroop test, when the word red is shown in a red font, this facilitates processing and causes a fast RT. A similar effect has been seen in relation to gender stereotyping. A study by Most *et al.* [101] conducted an auditory stroop task, where participants were asked to identify the gender of the voice speaking (male or female). The voice said either a traditionally male name or traditionally female name. The results found that when the voice said a name which matched the voice's gender, the RT was faster. Similar effects have been reported when participants were shown occupations in blue and pink ink, and asked to identify the gender stereotypically associated with that career [84]. Stereotype stroops can therefore give insights into the stereotypes held by individuals, and is illustrative of the way stereotypes can facilitate cognition, as mentioned in Chapter 3.3.

5.1.3 Priming

Stroop effects can also be elicited through priming, which is where a participant is exposed to a stimulus prior to the task, and this exposure later facilitates or inhibits performance [8]. For example, there is evidence that showing a threatening image such as a snake [19] or someone holding a gun [54] before a stroop trial, can prime the participant to respond slower. This effect can also be elicited by showing the stimuli prior to the stroop task, rather than interspersed between trials. This was evidenced in a study by Johansson *et al.* [69] found that priming women with high body satisfaction levels with images of thin models *facilitated* response times in a stroop task for performance-related words such as 'incapable'. Johansson *et al.* suggest that this facilitation effect occurred as a result of the thin priming leading to positive emotional responses for women already satisfied with their body image, thus causing faster response times. This illustrates that priming can lead to facilitatory or inhibitory affects, depending on the stimulus. Research by La Heij *et al.* [82] found whether the priming stimulus facilitated or interfered with the RT varied depending on whether the stimulus was associated with the stroop stimulus (i.e. the word written) or the desired response (i.e. the ink colour). If the prime was related to the stroop stimulus, this led to a delay in RT, but when it was related to the desired response it led to a faster RT.

5.1.4 The Effect of Device

Over time, the method of presenting stroop stimuli has evolved from hand written cards to computer run tests, but little research has been conducted on the effect of different technological devices on stroop response times. One study did compare the effect of using a desktop/laptop in a neutral university room compared to a smartphone in a home environment on an alcohol-related stroop task, and found that the results from each device/environment were comparable [127]. Furthermore, it has been found that conducting a stroop task via the internet (i.e. in a browser) yields similar results to a stroop test conducted using installed software, although the internet version leads to a delay in RT in general [86]. A further study has also been conducted using computer-related words as stimuli. Sparrow, Liu and Wegner [128] used a stroop paradigm to investigate response times to computer-related words (e.g. Google- a prominent computer search engine) compared to neutral words (e.g. Target- a US superstore) after being primed by answering either easy or difficult questions. They found that being asked difficult questions led to a slower response time to the computer-related words. Sparrow *et al.* suggested this effect may have occurred as being asked difficult questions primed the participants to think about the need to use a search engine to answer the questions. However, as yet no studies have been conducted investigating response times to words related to computing devices on different devices. Given this gap in the literature, the experiments presented in this chapter aimed to investigate the effect of words related to the device being used on processing speed.

5.2 Method

This study was conducted across two experiments which both used the same methodology and materials. The experiments only differed in terms of hypothesis, in that Experiment One hypothesised that stimuli related to the device being used would lead to faster response times, and Experiment Two hypothesised that the response times would be slower. Therefore, this section describes the materials and procedure for both experiments one and two. Sections 5.3 and 5.4 will describe the results of each experiment with its hypotheses, participant demographics, and the implications of the results. Section 5.5 then considers the results of Experiments One and Two in conjunction with each other, as well as in the context of other chapters within the thesis, and the wider research domain. The experiments received ethical approval from The University of Manchester, ref: 2020-10117-16292 (see Appendix E).

5.2.1 Materials

A custom online stroop software was developed in HTML and Javascript, and hosted at `iam-research.manchester.ac.uk/stroop`. Participants accessed the software from their own device (they were free to choose which device they used) using their preferred web browser. To account for different device screen sizes and resolutions, the stroop test scaled to the appropriate size based on the device being used (as indicated by the User-Agent header received from the participant's web browser). Likewise, interaction was tailored to best-suit the utilised device: if the participant was using a desktop or laptop, they were asked to indicate their responses using the arrow keys on their keyboard. If they were using a smartphone, arrow keys appeared on the touch screen for participants to use to indicate their responses. On all devices participants used the left arrow to indicate the word was written in red, the down arrow to indicate the word was written in green, and the right arrow to indicate the word was written in blue. The stroop application recorded the presented stimuli, correct and incorrect responses, and response time. The dependent variable of the study was the amount of time taken for the participant to identify the colour the stimuli was presented in, measured in milliseconds (ms).

5.2.2 Participant Inclusion Criteria

Due to colour recognition being an important part of the study, only those with normal or corrected to normal colour vision could participate. Participants were also required to be aged 18 or over, and be a fluent English speaker. This language requirement was included because if the hypothesised effects of stereotypes caused by the stimuli were to occur, it would require the stimuli used to be comprehended fully by the participant. The sample sizes for the experiments were informed by sample sizes used in previous studies (e.g. [60, 101])

5.2.3 Stimuli

The practice stimuli (words shown in the practice trials) were 'cat', 'dog' and 'rabbit'. These were presented in a random order.

The experimental stimuli (words shown in the experimental trials) came from three categories:

1. Words generated in the free association task (see Chapter 4.4.3) called desktop words and smartphone words.

2. Words from the Hedonic/Utilitarian Scale (see Chapter 4.3.2) called hedonic words and utilitarian words.

3. Words from the Stereotype Content Model (see Chapter 4.3.2) called warmth words and competence words.

A total of 45 experimental stimuli were used in the study. As emotional stimuli can lead to differences in response time (see Chapter 5.1.1), three participants were asked to rate each potential stimulus in terms of how positive they felt it was (Likert scale of one to five where one was neutral and five was very positive) and how negative they felt it was (also a scale of one to five but this time five was very negative). The mean positive and negative scores for each word were then compared for each stimuli type (i.e. the desktop words were compared to smartphone words). These were compared using repeated measures t-tests. The results of these tests can be seen in Table 5.1. Only one comparison led to significant differences, and the potential effect of this will be considered in the analysis.

As the words in the hedonic/utilitarian scale were pairs, these were compared internally as well. This means that the positive and negative scores for a word pair were compared, and if the positive score for the positive word was more than a mean score of .333 away from the negative score of the negative word, it was not included. For example in the word pair wise and foolish, if the mean positive score for wise was more than .333 higher than the negative score for foolish, the word pair was not included in the study. This led to two word pairs measuring utilitarianism (*beneficial/harmful* and *wise/foolish* to not be included in the study. However, this did lead to more stimuli being in the hedonic category than the utilitarian category.

T-tests were also run to determine whether words in each category had significantly more letters than their counterparts, as longer words can also lead to delays in RT. The results showed that there were no significant differences in word length for the stimuli (desktop and smartphone words: $t = -1.97$, $p = .062$; hedonic and utilitarian words: $t = 0.24$, $p = .817$; warmth and competence words: $t = 1.58$, $p = .159$).

While each participant was meant to see each word once, there was a consistent issue where words were missed or seen multiple times due to errors with the server. These errors were not deemed to be detrimental to the results, and thus no data was excluded

Word Category	Positive score		Negative score	
	t-value	p-value	t-value	p-value
Desktop and smartphone	0.00	1	1.31	.321
Hedonic and Utilitarian	0.58	.622	8.54	.013*
Warm and Competent	1.12	.383	-1.51	.270

Table 5.1: A table reporting the results of six repeated measures t-tests investigating how positive and negative the stimuli reported in the study were considered to be. There was a significant difference in how negatively rated the hedonic and utilitarian stimuli were, whereby the hedonic stimuli were rated as significantly more negative than the utilitarian stimuli. * denotes a significant effect.

due to this error[1]. The full word lists used for the experimental stimuli can be seen in Appendix F.

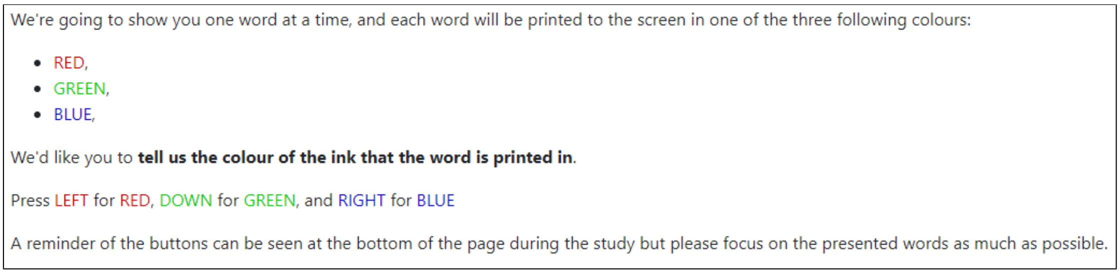
5.2.4 Procedure

Participants were asked to read the participant information sheet (Appendix G) and were told to email the researcher if they had any questions. If they were happy to participate, the participants were asked to complete a check box consent form to proceed.

The participants were presented with instructions which explained the stroop task and controls for the device the participant was using. The instructions can be seen in Figure 5.1. After clicking continue, 10 practice trials commenced. The practice stimuli was written in either red, blue, or green (with the colour being randomly assigned) and participants had to indicate the colour the word was written in as fast as they could. Each word was shown for up to 2 seconds followed by a fixation cross for 1 second. When participants responded to a word, it was replaced by the fixation cross.

Following the practice trials, the participants were shown the instructions again, before commencing 45 experimental trials. This had the same procedure as the practice trials but used the experimental stimuli. After the experimental trials, participants were shown the debriefing sheet (Appendix H) which explained the purpose of the study and reminded the participant of how to contact the researcher if they had any questions or

[1] The missing words are potentially due to the participant not responding to the stimuli in time, and thus the word being displayed was not recorded by the website. It is unclear why some words appeared multiple times. It was decided that no trials would be excluded on the basis of these errors, and thus the only exclusion criteria applied to the data was that described in Chapter 5.2.5


We're going to show you one word at a time, and each word will be printed to the screen in one of the three following colours:

- RED,
- GREEN,
- BLUE,

We'd like you to **tell us the colour of the ink that the word is printed in**.

Press LEFT for RED, DOWN for GREEN, and RIGHT for BLUE

A reminder of the buttons can be seen at the bottom of the page during the study but please focus on the presented words as much as possible.

Figure 5.1: The instructions given before the practice trials and the stroop trials in Experiments One and Two.

wished to withdraw their data.

5.2.5 Data Processing

The mean scores for each of the six stimuli types (desktop words, smartphone words, hedonic words, utilitarian words, competence words, and warmth words) per participant were calculated. These mean scores were then analysed using the appropriate statistical tests. The data in both the experiments was not normally distributed for any of the dependent variables. As such, Scheirer-Ray-Hare tests were used to analyse the results. To evaluate the nature of any significant interactions, Wilcoxon signed-rank tests or Mann-Whitney U tests were conducted. These were chosen due to a lack of suitable post-hoc test for the data. To reduce the number of type-1 errors, these were only conducted for the data points of interest rather than for all possible combinations of data, and Bonferroni corrections were applied. Due to the directional nature of the hypotheses, the post-hoc tests were one-tailed.

For all participants and data points, the data analysis was only conducted in cases where the colour the word was printed in was correctly identified. RTs faster than 200ms and slower than 2500ms were excluded. This upper threshold allowed answers given as the stimulus changed to the fixation cross to be included in the analysis. The lower threshold was chosen based on the reaction time of an average person. Specifically, Bellis [16] asked people of all ages to take part in a reaction time test. When grouped by decade (e.g. 10-19, 20-29, etc), the mean reaction time for all age groups fell between 220 and 320ms. This study did not require any decision making regarding which button to press. Therefore, while the fastest reaction times in Bellis' study did fall beneath

200ms, when taking into account that the participants in the present study must decide which button to press, the minimum threshold of 200ms was deemed appropriate.

As the Hedonic/Utilitarian scale uses semantic differentials, the RTs for those stimuli needed to be standardised so that a fast RT to a negative hedonic or utilitarian word was not treated the same as a fast RT for a positive hedonic or utilitarian word. For example, the words *useful* and *useless* should elicit opposite responses to each other if the word relating to the device impacts RT. To allow this differentiation, the following formula was used to standardise the mean RT:

$$RT_{std} = \frac{\sum_{i=1}^{\frac{n}{2}}(NW_i - PW_i)}{n} \tag{5.1}$$

where:

RT_{std} = The mean standardised response time
n = The total number of stimuli
NW_i = The response time of the i^{th} non-hedonic/utilitarian word(e.g. useless)
PW_i = The response time of the i^{th} hedonic/utilitarian word (e.g. useful)

Given this, if RT_{std} is less than zero, the PW stimuli have an overall slower response time than the NW stimuli. If RT_{std} is greater than 0, then the reverse is true. The closer RT_{std} is to zero, the more similar the RTs are for each of the stimuli. For example, if we have two stimuli *useful* and *useless*, where the RTs are 500ms and 800ms respectively, then the equation becomes:

$PW_1 = 500$
$NW_1 = 800$
$n = 2$
$RT_{std} = \frac{\sum_{i=1}^{\frac{2}{2}}(800-500)}{2}$

$RT_{std} = \frac{300}{2}$

$RT_{std} = 150$

In this case, we would then conclude that NW_{RT}s are responded to slower than PW_{RT}s. The standardisation of the hedonic/utilitarian stimuli was conducted after the previously described exclusion criteria was applied.

5.3 Experiment One

In this experiment it was hypothesised that words related to the device being used would elicit a stereotype stroop effect (see Chapter 5.1.2), in that a word related to the device would lead to a shorter RT. For the three types of experimental stimuli, the hypotheses were as follows:

H1 Desktop words would be responded to faster when seen on a desktop than a smartphone, and smartphone words would be responded to faster when seen on a smartphone than a desktop.

H2 Smartphone users would respond faster to hedonic and utilitarian words than desktop users.

H3 Smartphone users would respond faster to competence words than desktop users and desktop users would respond faster to warmth words than smartphone users.

H2 and H3 were generated based on the differences in ratings observed in Chapter 4. This experiment was pre-registered at `https://aspredicted.org/J9J_TQX`.

5.3.1 Participants

Thirty-six people took part in Experiment One. While the initial plan for the study was to recruit approximately 50 participants, the data analysis indicated that the one-directional hypotheses were in the opposite direction to the trends in results. As such, this experiment was stopped at 36 participants, and Experiment Two was commenced.

Of the participants in Experiment One, 26 participated on their smartphone[2] and 10 participated using a desktop or laptop. The desktop group was made up of 7 men and 3 women, and the smartphone group consisted of 15 women, and 11 men. The participants were from a range of occupations, and the student participants studied an array of degrees. As such it is unlikely the sample was biased towards a particular type of individual.

[2] Or tablet device which used a mobile browser.

Age	Desktop (*n* = 10)	Smartphone (*n* = 26)
Mean	33.60	29.58
Standard Deviation	7.92	8.88
Range	20 - 45	18 - 57
Median	35.00	27.00
Interquartile Range	28.00 - 37.75	25.00 - 33.50

Table 5.2: A table describing the ages of the participants in the desktop group and smartphone group for Experiment One. Ages are broken down into the mean, standard deviation, range, median, and interquartile range.

Participants' ages ranged from 18 to 57. A description of the participants ages by group can be seen in Table 5.2. As age can have an effect on RT, where older adults tend to have slower RTs, the two groups were compared to determine whether the ages were significantly different. A Shapiro-Wilk test of normality indicated the data was not normally distributed (W(36) = .907, p = .005). As such, a Mann-Whitney-U test was conducted, which suggested there were no significant differences in age between the groups (U = 175.5, p = .111). Therefore, any differences in RT between groups are not likely to be because of age.

5.3.2 Results

Desktop and Smartphone Stimuli

When looking at the stimuli related to either a desktop or smartphone, the results suggest that there is a significant main effect of the device being used on RT (H(1,71) = 4.43, p = .035). The mean RTs indicate that smartphone users (mean RT = 863ms) took longer to respond than desktop users (mean RT = 785ms). There was no significant effect of word type on RT (H(1,71) = 0.75, p = .386). There was a significant interaction between the device being used and the type of word being shown (H(1,71) = 31.15, $p < .001$).

Two Wilcoxon signed-rank tests were conducted to investigate the nature of this interaction. The first compared desktop users' RTs for phone words and desktop words and did not find a significant effect (W = 34, p = 1). The second compared smartphone users' response times for desktop words and phone words. No significant effect was found (W

= 70, p = 1) for the hypothesis of smartphone users reacting faster to smartphone related words than desktop words. The mean RTs for each group can be seen in Figure 5.2A.

Hedonic and Utilitarian Stimuli

For the stimuli relating to hedonic and utilitarian words, the results suggested there were no significant main effects of the device being used (H(1, 71) = 0.99, p = .319) or the type of word (H(1, 71) = 0.02, p = .875). There was however a significant interaction between the type of word and the device being used (H(1, 71) = 38.84, $p < .001$).

To investigate this relationship, two Mann-Whitney U tests were conducted for hedonic words seen by desktop users compared to smartphone users, and utilitarian words seen by desktop users compared to smartphone users. However neither of these tests were significant (U = 128, p = .958 and U = 162, p = 1 respectively). The mean RT_{std} for the desktop and smartphone groups for each stimulus type can be seen in Figure 5.2B.

Warmth and Competence Stimuli

For the SCM words, the results suggested that there were no significant main effects of the device being used (H(1, 70) = 3.57, p = .059) or of the type of word being shown (H(1, 70) = 3.19, p = .075). There was however a significant interaction between the independent variables where H(1, 71) = 30.78 and $p < .001$.

Follow up Mann-Whitney U tests investigated the difference for competence words seen by desktop users compared to smartphone users, and for warmth words seen by desktop users compared to smartphone users. For the warmth words, there was no significant difference in RT between the two groups (U = 124, p = .986). There was also no significant effect of device being used on RT for competence words (U = 65, p = 1). The mean scores for each stimulus type and group can be seen in Figure 5.2C.

5.3.3 Findings

The results suggest that while there were significant interactions between the groups and the stimulus types, the post-hoc tests indicated no significant differences were found in the directions hypothesised. However, when looking at the graphs in Figure 5.2 there

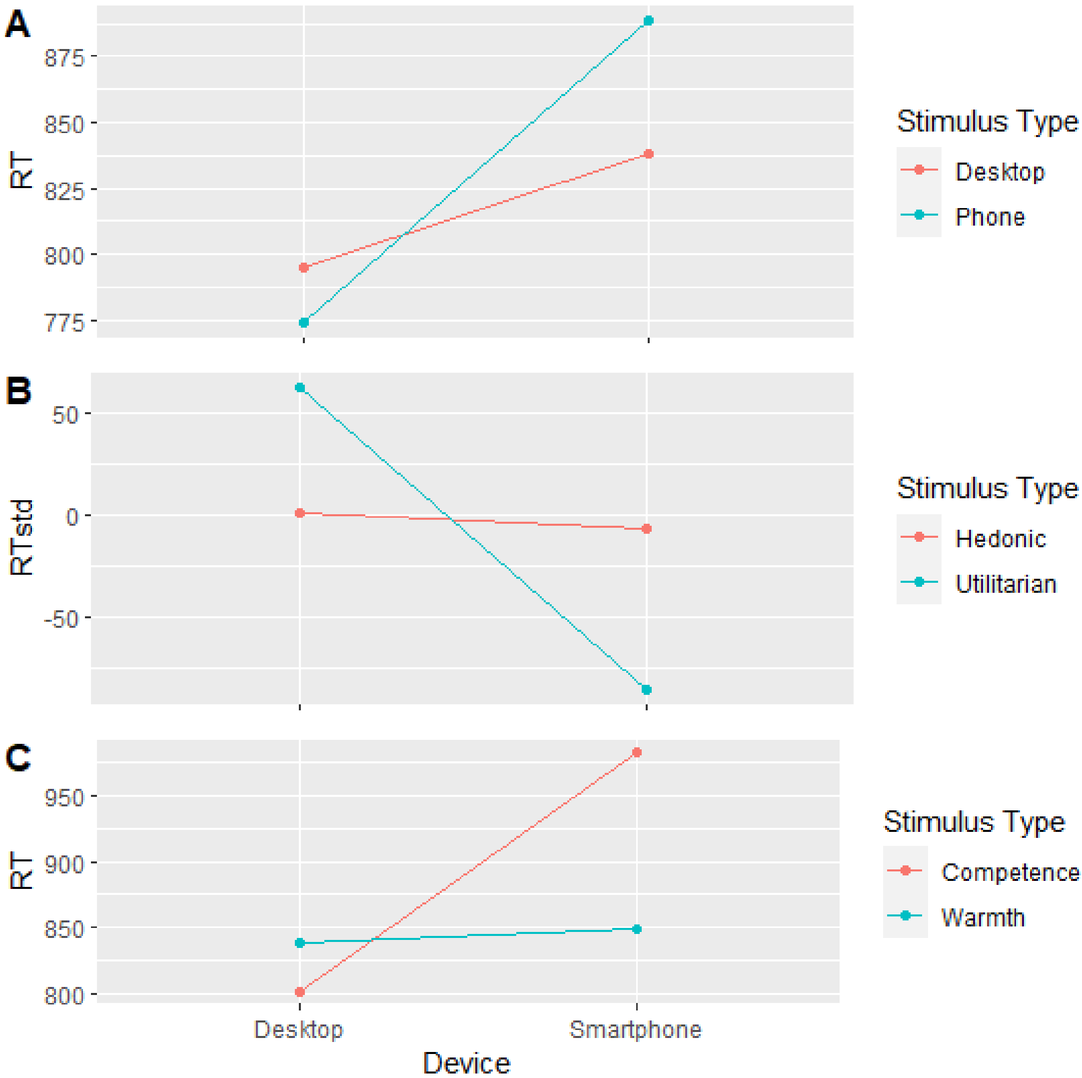


Figure 5.2: The results for each stimulus type by device for Experiment 1. A and C show the mean RT in ms, while B shows the RT_{std} (calculated as per Equation 5.1). The different coloured lines indicate the stimulus type.

appears to be some differences between the groups in the opposite direction to that which was hypothesised.

In regards to H1, Figure 5.2A indicates that instead of smartphone users responding faster to smartphone words, they responded slower in comparison to desktop words. Desktop users also appeared to respond slower to desktop words than smartphone words, which again goes against the hypothesis. However, the difference in response time to desktop and smartphone words for desktop users does not appear to be as large as the difference seen for smartphone users. It is of note though, that the number of participants in the desktop group was lower than that in the smartphone group. This was due to participants having free choice of the device being used, and data collection ending early before any efforts were made to increase the number of participants in the desktop group. As such the effect of each stimuli type on desktop users may become clearer if there was a larger sample.

H2 predicted that smartphone users would respond faster to hedonic words than desktop users, however the results suggested there were no significant differences. Although, Figure 5.2B suggests that smartphone users were taking longer to respond to utilitarian words than desktop users (as the graph indicates a slower response to utilitarian PW_{RT} words than utilitarian NW_{RT} words). A similar trend can be seen for hedonic words, although the difference between groups is not as large. This experiment suggests that there is no evidence to support H2 despite the significant interaction between indicated by the Sheirer-Ray-Hare test. Thus further research is required to investigate why this interaction may have occurred.

For H3, it was hypothesised that smartphone users would respond significantly faster to competence words than desktop users, however this was not supported. Instead, as can be seen in Figure 5.2C, smartphone users reacted slower to competence words than desktop users. Given the lack of significant main effect of device being used, this suggests that the type of word influenced RT. Furthermore, the hypothesis that desktop users would respond faster to warmth words was not supported, and Figure 5.2C indicates that there is very little difference between the groups.

Overall, these results suggest that the type of word being shown does not influence the RT of the two device groups differently, but Figure 5.2 suggests that there may be an effect in the opposite direction to that suggested in the current hypotheses. The hypotheses were formulated based on the assumption that words congruent with the stereotype of the device would elicit a *stereotype* stroop effect. Instead, the results suggest an *emotional* stroop effect, where words related to the device are causing interference resulting in delayed processing speeds. However, as the hypotheses were directional, further work was needed to verify this effect, thus leading to Experiment 2.

5.4 Experiment Two

In Experiment 1, the observed effects appeared to be in opposition to the predicted stereotype stroop effect. For this reason, in this experiment, the hypotheses are based on the potential presence of an emotional stroop effect (see Chapter 5.1.1), and the trends identified in Chapter 4. For the three types of experimental stimuli, the hypotheses are as follows:

- H4 Desktop words will be responded to slower when seen on a desktop than smartphone words, and smartphone words will be responded to slower when seen on a smartphone than desktop words.
- H5 Smartphone users will respond slower to hedonic and utilitarian words than desktop users.
- H6 Smartphone users will respond slower to competence words than desktop users and desktop users will respond slower to warmth words than smartphone users.

This experiment was pre-registered at `https://aspredicted.org/Y96_2CF`.

5.4.1 Participants

Fifty-three people took part in Experiment Two. The smartphone group was made up of 28 participants and 25 participated on a desktop or laptop. The smartphone group consisted of 17 women, and 11 men. The desktop group consisted of 13 women, nine men, two non-binary people, and one person who did not wish to specify their gender. The participants were from a range of academic disciplines and occupations, suggesting the sample is not biased towards a particular type of person.

Age	Desktop (n = 25)	Smartphone (n = 28)
Mean	28.44	32.89
Standard Deviation	9.02	12.35
Range	18 - 50	20 - 69
Median	27.00	30.50
Interquartile Range	22.00 - 33.00	25.75 - 34.50

Table 5.3: A table describing the ages of the participants in the desktop group and smartphone group for Experiment Two. Ages are broken down into the mean, standard deviation, range, median, and interquartile range.

Participants' ages ranged from 18 to 69 years. The full breakdown of participants' ages by group can be seen in Table 5.3. To ensure the age of the participants was not influencing any differences that may be found between the groups, a statistical analysis comparing the ages of each group was conducted. A Shapiro-Wilk test for normality suggested the age data was not normally distributed (W(53) = .853, $p < .001$). This meant a Mann-Whitney-U test was used to determine if the ages between the two groups were significantly different. The analysis showed that there were no significant differences in age between the participants (U = 267.5, p = .143). This suggests that any differences in RT were not due to differences in ages between the desktop and smartphone groups.

5.4.2 Results

Desktop and Smartphone Stimuli

For the stimuli related to either a desktop or smartphone, the results suggest that there was a significant main effect of the device being used on RT for words relating to desktops and laptops (H(1,105) = 37.57, $p < .001$). The mean RTs indicate that smartphone users (mean RT = 955.49ms) took longer to respond than desktop users (mean RT = 748.80). There was no significant main effect of word type on RTs (H(1,105) = 0.02, p = .902). There was a significant interaction between the device being used and the type of word being shown (H(1,105) = 4.21, p = .045). The mean RTs can be seen in Figure 5.3A.

To evaluate the nature of this interaction, two Wilcoxen signed-rank tests were conducted as per Experiment One. The first compared desktop users' RTs for smartphone words and desktop words and found a significant effect (W = 88, p = .045) where the mean RT for desktop words was slower than for smartphone words (741 and 710ms

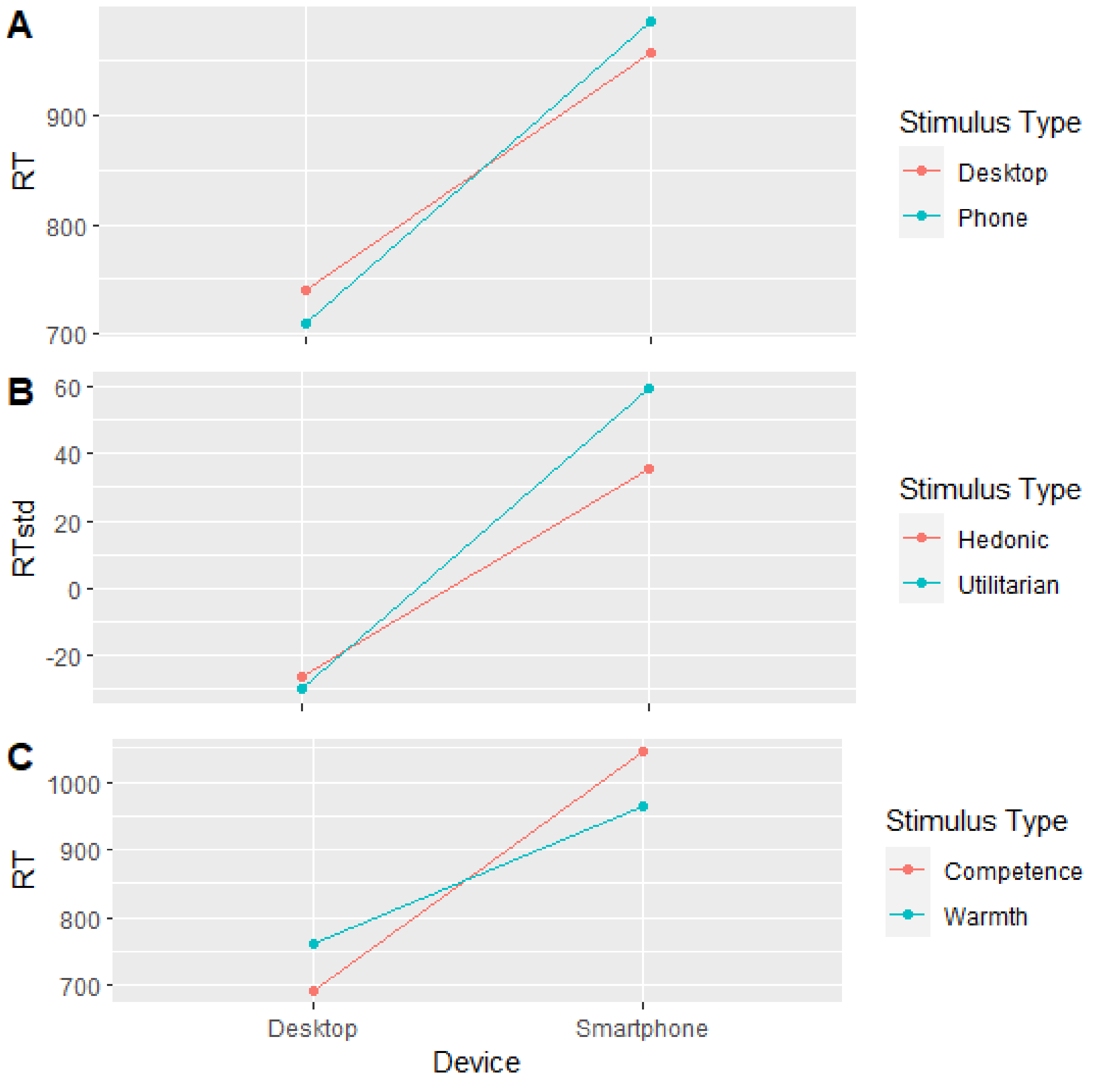


Figure 5.3: The results for each stimulus type by device for Experiment Two. A and C show the mean RT in ms, while B shows the RT_{std} (calculated as per Equation 5.1). The different coloured lines indicate the stimulus type.

respectively). There were no significant differences for the smartphone group for smartphone words compared to desktop words (W = 150, p = .232).

Hedonic and Utilitarian Stimuli

In terms of the hedonic and utilitarian stimuli, there were no significant main effects of the device being used (H(1,103) = 0.77, p = .379), or the type of word being shown (H(1,103) = 0.25, p = .614). There was a significant interaction between the independent variables (H(1,103) = 5.92, p = .015). The mean RT_{std} for each stimuli type can be seen in Figure 5.3B.

Follow up Mann-Whitney U tests were conducted to investigate the difference in desktop users' RT_{std}s for hedonic words compared to smartphone users' RT_{std}s, and the difference in desktop users' RT_{std}s and smartphone users' RT_{std}s for utilitarian words. The results suggested that there were no significant differences between the groups for either word type (Hedonic words: U = 314, p = 1; Utilitarian words: U = 290, p = 1).

Warmth and Competence Stimuli

For the SCM words, there was a significant main effect of the device being used (H(1,104) = 30.23, $p < .001$), where smartphone users reacted to the stimuli significantly slower than desktop users. There was no significant main effect of whether the word shown was related to competence or warmth (H(1,104) = 0.02, p = .875). There was a significant interaction between the device being used and the type of word being shown (H(1,104) = 5.72, p = .017). The mean RTs can be seen in Figure 5.3C.

Mann-Whitney U tests were conducted to investigate whether there were significant differences for the desktop group compared to the smartphone group when shown a competence word, and for the desktop group compared to the smartphone group when shown a warmth word. The former test found a significant difference (U = 78, $p < .001$), where the desktop group responded faster to competence words (mean RT = 692ms) than the smartphone group (mean RT = 1045ms). There was no significant difference between the groups for warmth words (U = 185, p = 1), where again desktop users were faster than smartphone users (RT = 762ms and 965ms respectively) contrary to the hypothesis.

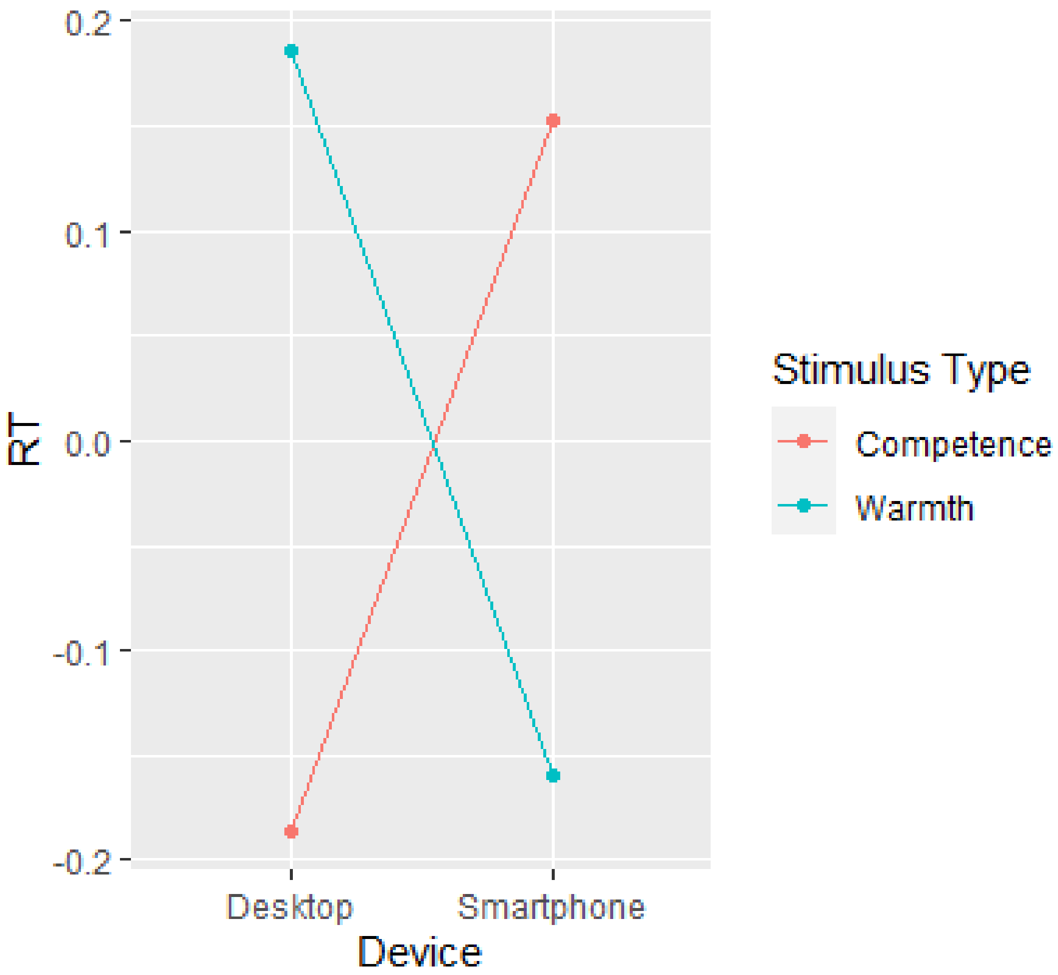


Figure 5.4: The standardised results for the warmth and competence stimuli.

Standardised results

Given the significant main effect of device for the warmth and competence words, z-score transformations were used to standardise the RTs for desktop and smartphone users. Doing this allowed for a comparison of the RTs for the word type without the effect of device being a factor, as the mean RT for each device becomes the same. Mann-Whitney-U tests were run on this transformed data, with a Bonferroni correction applied. A visualisation of the Z-scores can be seen in Figure 5.4.

For competence words, the results were not significant (U = 290, p = .292), but the standardised scores indicated that desktop users responded faster to competence words (z = -0.19) compared to smartphone users (z = 0.15). The results for warmth words were also not significant (U = 395, p = .299), but smartphone users responded faster than desktop users (z = -0.16 and 0.19 respectively). While these results were not significant, the results trend in the direction of the hypothesis.

5.4.3 Findings

The results partially support H4, as there was a significant effect whereby desktop users responded significantly slower to desktop words than smartphone words. This suggests that words related to the device being used elicited an emotional stroop effect. There was no significant effect for smartphone users when being shown smartphone words compared to desktop words. However the RTs for smartphone users did trend in the direction of the hypothesis, with smartphone words leading to a mean RT of 985ms compared to 956ms for desktop words.

H5 predicted that smartphone users would respond slower to hedonic words than desktop users, however the results did not support this. The findings instead suggest that there is no significant difference in the speed at which different device users respond to hedonic words. There was also no difference in RT_{std} between the groups for utilitarian words. This suggests that words from the hedonic/utilitarian scale do not affect information processing speeds.

H6 predicted that competence words would be processed slower by smartphone users than desktop users and this was supported by the unstandardised results. However, the hypothesis that warmth words would be processed faster by smartphone users was not supported. Further analysis was conducted to control for the significant main effect of the device being used. The standardised results found that there was no significant difference for either warmth or competence words between groups. Thus, there is insufficient evidence to support H6, despite standardised scores trending in the hypothesised directions for both warmth words and competence words. The interference caused by the type of device suggests that the effect of word type was masked by the difference in interaction speed based on the device.

Overall, the results suggest that words related to devices can lead to an emotional stroop effect, where the RT is slower when a stimulus is related to the device being used. However, words from the hedonic/utilitarian scale do not appear to elicit such an effect, and confounding variables such as differences in the interface being used should be considered when interpreting the results.

5.5 Discussion

This section jointly considers the findings of Experiments One and Two, along with their implications.

Both Experiments One and Two suggest that words known to have an association with a particular device (desktop/smartphone) impact participant RTs when presented on that device. Specifically, both experiments show a trend for slower RTs when the word presented is associated with the device being used. However, the differences in RT in testing H1 were larger when examining smartphone users' responses than desktop users' responses (see Figure 5.2), whereas the significant difference when testing H4 was in relation to desktop users. This suggests that while stimuli related to the device being used can influence information processing speeds, this is not necessarily consistent. In particular, individual differences are a confounding factor when it comes to the associations people make. While the desktop and smartphone stimuli used in these experiments were among the most frequently generated in the free association task in Chapter 4.4.3, they are not necessarily words that the participants in the experiments reported in this chapter would generate if they participated in the free association task. It is possible that the smartphone users in Experiment One more closely associated the smartphone words with their device than the smartphone users in Experiment Two, and vice versa for desktop users, causing variations in the amount of interference observed in the stroop task.

In Chapter 4, despite visible differences in the placement of devices on the utilitarian/hedonism scales (Figure 4.4), the results did not show a statistically significant difference between desktops and smartphones. Results from the experiments in this chapter are similarly inconclusive — both Experiment One and Experiment Two demonstrated significant interactions between word type (hedonic/utilitarian) and participant device, but post-hoc tests did not identify the nature of these interactions. This may be explained by the point made in Chapter 4.5.2, that the words used in the hedonic/utilitarian scale are not naturally generated in association with the devices, suggesting that they may not be appropriate stimuli to elicit an effect. Overall the testing of H2 and H5 suggest there was a lack of significant effect of hedonic and utilitarian traits on information processing. However, the hedonic stimuli were generally rated more negative than the utilitarian words, when the emotional salience of the words was being assessed prior to the study. As such, this significant emotional effect may have led to the results not truly reflecting the impact of the associations. As such, further replication of this experiment may be required to better control for the emotional salience of the hedonic/utilitarian scale.

The raw results of Experiments One and Two found that competence words elicited slower RTs on smartphones than desktops although this was not a significant effect for the standardised results of Experiment Two. Overall, the direction of this trend was consistent with H6. While this suggests that the effect seen in the raw scores was due to differences in the devices (for example, the use of touchscreens on smartphone devices leading to slower interactions overall), this main effect of device was not found consistently throughout Experiments One and Two, indicating interaction modality and device was not the underlying cause. Rather, it may be the case that smartphone users in Experiment Two associated some stimuli with their device to a greater extent, leading to slower RTs. Conversely, it may be the case that the desktop group had weaker associations between their device and the stimuli (potentially due to the use of laptops rather than desktops). However, more research is needed to fully understand why these effects occurred. Overall it appears that the significant effects did not persist once the effects of device were controlled for, suggesting H6 was not supported. The standardised results of Experiment Two did however trend in the direction that was hypothesised, suggesting that a study with higher power might support the hypothesis.

Hypotheses for Experiment 1 assumed a stereotype stroop effect. However, the results of that experiment clearly trended in the opposite direction to a stereotype stroop.

This opposing trend suggested that an emotional stroop may instead be present. This is perhaps not surprising given the findings of La Heij *et al.* [82] which suggest that the stereotype stroop facilitation effect only occurs when the association is with the response, not the stimulus. This also makes sense in the context of previous stereotype stroop studies (see Chapter 5.1.2), as again the response the participant was required to generate was related to the stimulus. In the context of this study, it is also logical that there would be an emotional stroop effect, as if it is the device influencing RT, then that would be related to the stimuli, not the desired response (the colour of the word). Given this, the findings of the present study are consistent with prior literature.

One further consideration to make is that the lack of significant results for the SCM and hedonic/utilitarian terms may be due to the devices being investigated. The choice to investigate desktops and smartphones in this thesis was due to their prevalence in society as 'everyday devices' as discussed in Chapter 1.1. However, in belonging to this category, these devices may be perceived more similarly than a desktop would be when compared with a pair of smartglasses. When Fiske *et al.* [40] mapped how different social groups were perceived onto the SCM, while there were four clear quadrants, multiple groups were contained within those quadrants. Inferences on how groups within the same quadrants were seen compared to each other could still be made despite there not being statistically significant differences. In the case of the devices considered in this study, it may be the case that if mapped they would be within the same quadrant, but within that, smartphone users are seen as slightly more competent than desktop users. This in turn still gives insights into how the technology is perceived despite the lack of significant difference.

5.5.1 Implications

Overall the results suggest that when a stimulus relates to the device being used, it can cause some level of interference in processing. This in turn suggests that the way an individual perceives their device can have unconscious effects on information processing. The delays seen in this study are small (measured in 100s of ms), but the effects seen in this study nonetheless have implications for cognitive processes.

As noted in Chapter 2.5.2, attention can influence memory processes by influencing the amount of information encoded. This is evidenced in a study by MacKay *et al.* [89], who conducted a taboo stroop task followed by a surprise memory test. The results found that taboo words took longer to respond to in the stroop test and were then better recalled than the neutral words in the memory test. In terms of the present study, this suggests that the words associated with the device, which caused a longer RT, would likely be more easily recalled than the words which were processed faster, as the longer response delay leads to those words being attended to more.

The differences in RT for hedonic/utilitarian and SCM stimuli are consistent with the responses reported in Chapter 4.4.3. Specifically, in cases where a device rated higher on a construct in the hedonic/utilitarian scale or SCM, there was similarly a higher RT for users of that device when shown those stimuli. Given this, the use of stroop tasks may allow the implicit assessment of stereotypes in place of traditional scales. As people are asked to respond quickly to the stroop task, there is less opportunity for participants to try to give socially desirable responses. Therefore, should a stroop task with terms from the SCM be conducted where before being shown stimuli the participants are shown the name of a social group, this may allow an inference of where that group would be placed on the SCM. However, to verify this is a viable method, research would need to be conducted to investigate the similarity of results from a stroop task and the SCM. This has been done to an extent in a study by White and Gardner [148] who conducted a stroop task using synonyms and antonyms of the words ‘competence’ and ‘warmth’. They found that being asked to think about women prior to the stroop task led to longer RTs for words related to warmth. Given this, there is evidence that this may be a viable approach to investigate device stereotypes. However, as discussed, some devices may not be different enough to elicit significantly different response times to each other, and as such the results may be more indicative than conclusive.

This may further be applicable to user experience (UX) research. A study by Dell *et al.* [35] found that participants would express a preference for what they believed to be the interviewer’s product, even when it was identical to the control product. Given this, explicitly asking for opinions on artifacts in UX research may lead to participants saying what they believe the researcher wishes to hear. Instead this stroop method may be a viable way to gain insight into beliefs about products implicitly. For example, if the participant were asked to explore an interface (which would serve as a priming stimulus) before completing a stroop task where the stimuli related to usability traits, the results of this study suggest that a longer response time would be seen if the stimulus related to the interface. However, this would need to be verified by further research, investigating the use of an interface as a priming stimulus. If verified by future research, this could benefit UX researchers as for example, if after viewing an interface, a slow RT was seen for the term ‘intuitive’, this would suggest the interface performs well in terms of this usability criteria.

A final implication is that the device used to conduct stroop studies may influence the results. While using technology to conduct a stroop task rather than conducting it using paper materials allows for more accurate measurements of response times and reduces the impact of human error on the part of the experimenter, care needs to be taken that the device is not influencing the results. In particular, if the stimuli for the stroop task has a relation to the technology being used this could be a confounding variable. For example, research utilising a stroop task to investigate stereotypes held about older adults may find the terms ‘old’ or ‘slow’ are also associated with desktops, thus potentially confounding the results. As such, researchers should carefully consider the stimuli and technology used to conduct a stroop study. Given the previous findings of human stereotypes from the SCM relating to devices (see Chapter 4 and [117]), the stimuli which could be affected by device may not be immediately apparent. This is particularly important in studies investigating the internal reliability of stroop tests on different devices (e.g. [127]) as the stimuli may interact with the device, influencing the results. However, in cases where the stimuli is not interacting with the device being used, the repeated finding of a main effect of device indicates that there are device differences on RT. In particular, it appears that touchscreen interfaces lead to RT delays relative to keyboard inputs. Therefore work measuring RT in tasks available across multiple devices should take care to account and control for delays caused by the interaction modality, so as not to mistakenly attribute slow RTs to other variables.

5.6 Limitations

This study was impacted by two notable limitations: assumptions regarding the associations held about the devices, and a semi-controlled design. As mentioned in Chapter 5.5, a potential limitation of the study is that individual associations made by participants with their devices were not accounted for. This could be resolved by asking participants to generate words they relate to the devices, and utilise those words as stimuli. However, there are three key issues with this which led to this proposed methodology not being used in the present study. First, it is difficult to ensure all participants generate the same amount of stimuli as seen in Chapter 4.4.3, where the standard deviation for the number of words generated was approximately eight words. This would have then meant the length of the procedure and number of stimuli shown varied between participants. Second, as described in Chapter 5.1.3, there is evidence to suggest that priming individuals before a stroop task can influence their response times. Having participants generate the stimuli prior to the study may then lead to a priming effect and thus influence the RTs. Third, being asked to explicitly generate stimuli related to the devices may have led to participants responding in the way they thought the researcher would want if they recognised the stimuli, thus manipulating the results. In the present study participants were naive to the nature of the words they were shown, with this only being revealed after the study was complete. However, future work may wish to utilise this method as a further condition, to clarify the extent to which it influences the results.

A further limitation of this study is that participants were asked to participate from home due to the COVID-19 pandemic, which limited the ability to conduct in-person research for safety reasons. By participating in an uncontrolled environment, there may have been confounding variables such as noise or interruptions which could have influenced response time. Further, this study hypothesised that the device would serve as a cue to elicit a stroop effect (whether stereotypical as per Experiment 1 or emotional as per Experiment 2). The lack of control over the participants could have meant that both devices were present during the task (e.g. participating on a smartphone while sitting at a desktop computer), which may have confounded the results. This may have particularly impacted the desktop group, as it is rare for individuals to be apart from their smartphones. Participating at home also meant that there was limited ability for the device being used to be controlled, as participants were required to use their own devices. Therefore, while the findings of Chapter 4 suggests that devices with similar interaction methods should not be used interchangeably, this experiment allowed participants to use a desktop or laptop to participate. It is also possible that tablets were used in place of smartphones. This may then have impacted the results, as the stimuli pertained specifically to desktops and smartphones, and using other similar devices may not have elicited the same level of interference.

These factors may account for the lack of consistent replication between Experiment

One and Two. Given these limitations, further work is necessary to fully understand the impact of device stereotypes on information processing.

5.7 Conclusions

The experiments reported in this chapter indicate that stimuli relating to the device being used to conduct a stroop task can lead to slower RTs than unrelated stimuli. The RTs found in response to stimuli from the hedonic/utilitarian scale and SCM also reflected the ratings of the device on those scales in Chapter 4. The results of the stroop task therefore indicate that the device being used can impact the cognitive ability of information processing, where stimuli related to the device are attended to more, and thus responded to slower. There are potential applications for this (subject to future research), whereby the stroop paradigm may be used to investigate implicit opinions, aiding in the avoidance of demand characteristic effects. Furthermore, the increased attention observed may have implications for memory performance, whereby when more attention is paid to stimuli related to the device, making that information more easily remembered. In Chapter 6, the specific case of encoding and retrieving desktop and smartphone related stimuli on those devices is investigated.

Chapter 6

Memory

The findings from Chapter 5 indicate that stimuli associated with the device being used can cause interference in information processing tasks, as the related stimuli are attended to more than unrelated stimuli. Given this, the present chapter aims to investigate whether stimuli associated with the device being used are more or less easily remembered than the unrelated stimuli.

This was investigated in an online experiment, where participants were asked to memorise twelve words- six of which were related to desktops, and six of which were related to smartphones. Following a distractor task, the participants were asked to recall as many words as they could in two minutes before participating in a recognition memory test.

The results found a general trend whereby smartphone users remembered fewer stimuli than the desktop users, however this was only significant for the recognition of smartphone stimuli. The smartphone group also tended to falsely recognise more stimuli than the desktop group, although this was not significant. Overall, the results suggest a trend of smartphone users exhibiting worse performance in a memory task than desktop users.

6.1 Memories and False Memories

Memory in itself is a broad field which can encompass many tasks undertaken in everyday life, as described in Chapter 2.2.1. While there are many forms of memory, this chapter predominantly concerns itself with the recognition and free recall of semantic memory[1].

Free recall is a test of memory, where the participants are asked to generate the stimuli presented previously without any prompts. This differs to another form of recall called category-cued recall, where the prompt of the overall category name is given, and participants generate the previously seen stimuli within that category. Recognition memory is where previously seen stimuli are presented alongside previously unseen lure stimuli, and participants are asked to indicate which they have seen before and which they have not.

Recognition memory and recall memory are susceptible to interference to differing extents. As discussed in Chapter 3.4, the effect of environmental cues on memory performance vary depending on if the memory task is recall or recognition based. Specifically, Chapter 3.4 describes the phenomenon of recall memory processes being more susceptible to interference from outside factors, as the viewing of stimuli in the recognition tasks serves as a more powerful cue than the external factors [50].

Recognition memory tests can also give insight into false memories by utilising lures. When lure stimuli are semantically related to the original stimuli, the related but previously unseen word can be falsely recalled. For example, lures semantically related to stimuli were falsely identified more frequently than control lures which were not semantically related [142] (for further discussion on this see Chapter 3.3). This has also been found in recall memory. A study by Roediger and McDermott [113] found that when participants were presented with words such as 'bed' and 'awake', participants tended to falsely recall the word 'sleep'. This suggests that semantically related words are more likely to be falsely generated in recall tasks.

There is also evidence that memories are not recalled exactly as they were encoded but rather are reconstructed. A study by Bartlett [14] asked participants to read a story twice. The following day they were asked to recall the story by retelling it. Bartlett found that in the retelling, participants would replace aspects of the story that were culturally unfamiliar to them, with synonyms that were more familiar. For example the story recounted events involving a canoe, but the participants tended to replace the term with 'boat', a term more commonly used in their lives. Because of this, memories can be prone to false recollections, as details can be substituted for familiar, semantically

[1] With the caveat that this could potentially be considered episodic memory, if the memories are not independent of the context they are learnt in (see Chapter 2.2.1 for more details).

related terms due to memories being reconstructed each time they are recalled.

6.2 The present study

The present study aims to investigate whether the device used to encode and retrieve stimuli related to smartphones or desktops leads to differences in memory performance. Memory performance is assessed across three domains: recall, recognition, and false recognition.

The results of the stroop colour naming task in Chapter 5 found that stimuli related to the device being used took longer to respond to than unrelated stimuli. This is indicative of a greater level of attention being paid to the related words, thus delaying the speed of information processing. This increased attention to some stimuli may influence memory performance. As discussed in Chapter 2.5.2, not paying enough attention during the encoding phase can lead to poorer memory performance, as the information is not taken in. Furthermore, there is evidence that the increased attention paid to salient stimuli in the stroop task can lead to those stimuli being better recalled in a subsequent memory task [89] (see Chapter 5.5.1 for more details).

As such, the stimuli used in this study is from the desktop word and smartphone word categories, utilised in Chapter 5. The present study hypothesises that:

- H1 There will be a difference between the desktop group and the smartphone group in the number of stimuli recalled.
- H2 There will be a difference between the desktop group and smartphone group in the number of stimuli recognised.
- H3 There will be a difference between the desktop and smartphone groups in the number of incorrectly recognised stimuli.

This study was pre-registered at `https://aspredicted.org/TFF_FQX` and received ethical approval from the University of Manchester (ref: 2021-11329-18169; see Appendix I).

The following sections are structured as follows. First, the method and procedure of conducting the study along with information about the participants is presented. Then the results for each dependent variable are described, followed by a discussion of the findings.

This is followed by a discussion of the implications of the results, the limitations of the work, and directions for future work. The overall findings of the study are then presented in the conclusions.

6.3 Method

6.3.1 Materials

The software for the memory study was developed in HTML and Javascript. The study was hosted at `iam-research.manchester.ac.uk/memory`. Participants ran the study on the browser of their choice on either a desktop or laptop, or a smartphone. Participants were free to choose the device they used. However, part way through the data collection process, the site stopped running on 'Google Chrome' for desktops and laptops, and so participants in this group could only participate through alternative browsers.

The study scaled to the appropriate size for the device they were using based on the User-Agent header received from the participant's web browser. The stimuli presented at encoding and the text box to generate stimuli for the recall task were centred in the middle of the screen, and surrounded by a rectangular, black border. The recognition stimuli was presented in the centre of the screen just above the midpoint and surrounded by the same black boarder. The buttons to indicate if the stimuli had been seen before were directly under the stimuli. The presentation of the stimuli and its placement on the screen was kept as consistent as possible between encoding and retrieval to control for any effects changing layouts may have had on memory performance, which prior research indicates may occur [152].

6.3.2 Stimuli

The stimuli in this study were words generated from the free association task (see Chapter 4.4.3) and were the same words as the desktop and smartphone words used in Chapter 5. The list of words can be seen in Appendix F. Participants were shown 12 words (selected at random) during the encoding phase, 6 of the words were desktop words and 6 were smartphone words. During the recognition phase, all 24 words were displayed to the participants. Twelve was chosen as the number of presented stimuli as it has been used in prior research (e.g. [9]) and shown to be enough stimuli to allow memory effects to be detected, but not so many as to overwhelm the participant.

6.3.3 Procedure

Participants were asked to read the information sheet (Appendix J) and told to contact the researcher if they had any questions. They were then asked to complete the consent form if they were happy to participate. This was followed by demographic questions, where the participant was asked to provide their age, gender, and occupation. Participants were also asked to generate their own unique participant ID code so they could withdraw from the study if they wished.

Participants were then taken to the instructions for the encoding phase, where they were told they would be shown a series of words and they were asked to memorise them. Each word was displayed one at a time for five seconds, followed by a fixation cross for two seconds. After viewing the words, participants took part in a two-minute distractor task. In this, the participants were shown two numbers and asked to either add or subtract them. The first number was an integer between 51 and 100 (inclusive) while the second number was an integer between 1 and 50 (inclusive).

Following this, participants were shown the instructions for the recall task, which asked them to type as many of the words as they remember previously being shown in the study. They were told to press ‘enter’ before typing the next word, so each word was entered individually. After two minutes they were automatically taken to the page giving the instructions for the recognition test. The instructions asked participants to look at a series of words and click the ‘yes’ button if they remembered seeing the word earlier in the study or the ‘no’ button if they did not remember seeing the word previously. The words were presented one at a time. After the recognition task participants were given the participant debriefing sheet (Appendix K) which explained the purpose of the study in more detail.

6.3.4 Data Processing

To avoid the data being biased by participants saying they recognised every stimulus presented in the recognition test, the data was normalised. This was done by subtracting the number of false positives from the number of true positives, using the following formula:

$$Recognition = Pos_{true} - Pos_{false} \tag{6.1}$$

where:

$Recognition$ = The score given on the recognition test of memory
Pos_{true} = The number of stimuli correctly recognised
Pos_{false} = The number of stimuli falsely recognised

For example if the participant correctly recognised 4 words but falsely recognised 2 words their recognition score would be 2, as $Pos_{true} = 4$ and $Pos_{false} = 2$. By applying this formula, any participant who said they recognised all 24 of the stimuli would have a recognition score of 0, allowing those who cheated on the task by saying ‘yes’ for every stimulus to be identified.

False positives (Pos_{false}) were also analysed in isolation as a measure of false memories. For this the raw Pos_{false} score for each participant was used.

For the recall data, only exact matches irrespective of capitalisation, were considered correct. This means that if a participant recalled the word ‘call’, this was marked as incorrect, as the presented stimulus was ‘calls’. If a correct word was entered twice, then only the first occurrence of the word was marked as correct.

6.3.5 Participants

The inclusion criteria for this study required participants to be aged 18 or older, and have normal or corrected to normal vision. As the study was investigating the effect of words associated with the device being used, participants were required to be fluent English speakers, so the stimuli were more likely to be fully understood, and interpreted in similar ways.

Data for 24 participants was collected. However, one participant took part twice and thus their second attempt was excluded, as they may have recalled the stimuli from participating previously. A further participant was excluded because they answered less than 1% of questions in the distractor task correctly. As this task was undertaken to ensure that participants were not rehearsing the stimuli, and that the stimuli were

Age	Desktop (n = 7)	Smartphone (n = 13)
Mean	31.29	25.69
Standard Deviation	6.26	3.2
Median	34	26
Inter-Quartile Range	27. 00 - 34.50	23.00 - 26.00
Range	22 - 40	23 - 35

Table 6.1: A table describing the ages of participants in the desktop and smartphone groups.

stored in the long term memory, poor performance in this task was indicative that they may have been rehearsing, hence their data was excluded. The remaining participants correctly answered between 41.67% and 100% on the distractor task, with a mean score of 84.40%.

Finally, as per the pre-registration, a comparison of the ages between the desktop and smartphone groups was conducted, as older adults tend to exhibit poorer memory performance. Of the 22 participants remaining, the mean ages of the desktop and smartphone groups were 31.29 years and 25.07 years respectively. This was found to be significantly different between the groups (U = 81, p = .047)[2]. Thus, the two youngest participants from the study- both aged 21- were excluded from the analysis, consequently eliminating the significant difference between the ages of the two groups (U = 67, p = .093).

This led to a final sample of 20 participants to be included in the analysis. This was below the target sample size of 50 which was identified in the pre-registration, which was determined based on previous studies (e.g. [59]). Due to these issues in recruitment, and the application of the previously described exclusion criteria. The final sample was below the intended target, and as such the study may be underpowered.

The gender distribution of the desktop group was six males and one female, while the smartphone group was composed of seven males, three females, two non-binary people, and one gender-fluid person. The ages of the participants can be seen in Table 6.1.

The occupations of the participants were also collected. One participant in the desktop group said they were a student, and the rest were from a range of professions. The smartphone group contained four students, with the remaining nine coming from a

[2] A Mann-Whitney U test was run as the ages of the smartphone group were not normally distributed (W = 0.82, p = .007)

Memory Type	Stimuli	Desktop Group	Smartphone Group
Recall	Total	8.43 (2.44)	7.85 (1.63)
	Desktop	4.86 (1.68)	4.15 (0.90)
	Smartphone	3.57 (1.90)	3.69 (1.38)
Recognition	Total	10.71 (1.70)	9.23 (2.05)
	Desktop	5.00 (1.41)	4.62 (1.12)
	Smartphone	5.71 (0.49)	4.62 (1.33)*
False Positives	Total	0.43 (0.53)	1.36 (1.91)
	Desktop	0.43 (0.53)	0.71 (0.91)
	Smartphone	0.00 (0.00)	0.64 (1.08)

Table 6.2: The mean number of words recalled, recognised, and incorrectly recognised (false positives) for the total stimuli, desktop stimuli, and smartphone stimuli by group. The standard deviations are given in brackets. Significant differences between groups are indicated with * ($p < .05$)

diverse range of occupations. As such, the results of this study are unlikely to be biased towards a specific participant group, although the four students in the smartphone group may lead to higher scores as they may be more used to memorising information.

6.4 Results

Recall

The data from the recall task was normally distributed for the total number of words recalled and the number of smartphone words recalled (all $W > 0.90$, $p > .100$). Two independent groups t-tests were run to investigate if there were significant differences between the smartphone and desktop groups for these dependent variables. The desktop stimuli data was not normally distributed ($p < .05$) and this dependent variable was therefore analysed using a Mann-Whitney U test.

For the total number of stimuli recalled, there were no significant differences between the desktop group and the smartphone group ($t = 0.64$, $p = .529$). There were also no significant differences between the desktop and smartphone groups for desktop stimuli ($U = 64$, $p = .140$), or for smartphone stimuli ($t = -0.16$, $p = .871$). The mean scores and standard deviations for each stimuli type by group can be seen in Table 6.2.

Recognition

The data for the recognition test was not normally distributed (Shapiro-Wilks test of normality, $p < .05$ for at least one of the groups per dependent variable). Mann-Whitney U tests were used to analyse the performance of the two groups in terms of total corrected recognition score, corrected desktop recognition score, and corrected smartphone recognition score. These scores were corrected as per Equation 6.1.

There were no significant differences in performance for the total recognition score (U = 68, p = .076) or for the desktop recognition performance (U = 59, p = .266). There was however a significant difference in recognition of smartphone words (U = 71.5, p = .030), whereby the desktop group had a mean corrected recognition score of 5.71, while the smartphone group has a mean recognition score of 4.62 for smartphone words. The means and standard deviations for each recognition measure can be seen in Table 6.2.

False Recognitions

In the recognition task, some participants responded that they had seen words which had not been previously shown in the task. This data was not normally distributed (Shapiro-Wilks test of normality, $p < .05$ for at least one of the groups per dependent variable). Therefore, the rate of false positives for the total words, desktop words and smartphone words was analysed using Mann-Whitney U tests.

There were no significant differences between the number of total false positives (U = 34.5, p = .262) or false positives for desktop words (U = 42.5, p = .617). There was also no significant difference in the number of false positives for smartphone words (U = 28, p = .053). However, this was close to being significant, and the smartphone group had a mean error rate of 0.64 words, while the desktop group had a mean error rate of 0.00 (no errors were made). All the means and standard deviations by group for false recognitions can be seen in Table 6.2.

6.5 Discussion

6.5.1 Findings

The results of the recall memory test suggest that there were no significant differences between the groups in memory performance. There was however a trend for the desktop group to perform better overall compared to the smartphone group, as well as performing better for desktop words. This suggests that although not significant, those learning and retrieving on a desktop did better than those on a smartphone. The performance of both groups for smartphone words was very similar, suggesting that those stimuli were recalled equally well across devices, although the smartphone group did perform slightly better than the desktop group. While not included in the analysis, it is also of note that for the desktop group, the mean number of desktop words recalled was 1.29 words more than smartphone words, while for the smartphone group, this difference was only 0.61 words. This indicated that the desktop group found it easier to recall desktop words than smartphone words, compared to the smartphone group.

The results of the recognition memory test supported the results of the recall test in regards to performance, as the desktop group correctly recognised more of the desktop stimuli and stimuli overall than the smartphone group. There was also a significant effect in that the corrected recognition score for smartphone words was significantly higher for the desktop group than the smartphone group. This then suggests that recognition memory is worse on a smartphone than a desktop, particularly for words relating to smartphones.

The significant difference in the corrected recognition score for smartphone words may also be due to the trend seen in the false positive results. For all stimuli and stimuli subgroups, there was a trend of smartphone users falsely recognising stimuli that had not been presented before to a greater extent than desktop users. This would have then led to the recognition score for the smartphone group being smaller, as the recognition scores were calculated by taking the number of false positives away from the number of true positives (see Equation 6.1).

6.5.2 Implications

Overall the findings of this study suggest that those using a smartphone to encode and retrieve stimuli under perform when compared to their desktop user counterparts, although this is not consistently significant. The finding which was significant was in regards to the recognition of smartphone words, where the desktop group performed better than the smartphone group. This is contrary to the results of Chapter 5, where words relating to smartphones were attended to more by smartphone users than desktop users, leading to delayed response times. In this chapter, it was suggested that this increased attention would lead to better memory performance when the stimulus was related to the device being used, based on the findings of prior research [89]. However, the results of this study contradict this for smartphone users, who demonstrated poorer recognition of smartphone words. Nevertheless, the hypothesis that words related to the device being used did appear to hold true in the recall test for desktop users, where words related to the device were recalled more frequently by the participants using said device. Smartphone users also recalled slightly more smartphone words than desktop users, although this difference was very small (0.12 words). None of the results relating to recall were statistically significant so this evidence suggests that this is just due to chance.

Furthermore, these patterns in desktop users' recall of desktop words may just be a coincidence in the overall trend of desktop users performing better than smartphone users overall. For total and desktop recall and all measures of recognition, the desktop group remembered more items correctly than the smartphone group (although again this was only significant in the recognition of smartphone-related stimuli). This suggests that desktops may be more conducive to encoding and retrieving stimuli than smartphones. This has real world connotations. As reported in Chapter 4.4.2, one or more smartphone devices were used three or more times a week by 98.8% of participants, making it the most commonly used everyday device. As such, smartphones are more likely to be used for tasks and consequently, this means smartphones may be used even when it is not the most beneficial device. The results of this study indicate a trend whereby smartphone use may be detrimental to memory performance, and such tasks are not best undertaken on smartphones.

The issue of smartphones being detrimental to memory performance also raises concerns in relation to the digital divide. Research by Tsetsi and Rains [139] found that people in the US from ethnic minority groups, and lower income groups were more likely to only have internet access via a smartphone than white people, and higher income groups. This study suggests that the groups dependent on smartphones for internet access are potentially at a disadvantage compared to those not dependent on smartphones when it comes to undertaking memory tasks on a device. While statistical significance of this is not demonstrated in this study, the sample was less than half the desired size, and so it is highly likely the study was underpowered. Therefore, replication of this work would be beneficial to truly establish whether or not the hypothesised effects are present, and the potential impact of the digital divide on cognition.

The results regarding false positives in the recognition task also have implications. The results, while not significant, do trend towards the idea that smartphone users tended to make more errors in the recognition task than the desktop group. None of the participants scored zero on the recognition test, which indicates they were not just selecting ‘yes’ or ‘no’ repeatedly. This indicates that any false recognitions were the result of false recollections, rather than lack of effort. Thus, the results of this study suggest that smartphone users may be more prone to false memories in a memory task than desktop users. This again could have widespread consequences, given the prevalence of smartphones in society.

The ways in which technology can distort memories and lead to false recollections has been discussed in terms of existing technologies, and more speculatively in relation to future technological developments designed to augment cognition [27]. The present study indicates that devices may vary in terms of the extent their users may be impacted by this phenomenon, as smartphone users appeared to be more susceptible to false recognitions. However, there is a further potential effect that the nature of the false memory may vary in terms of device, due to the stereotypes associated with it. A study by Araya *et al.* [9] found that stereotypical priming presented before word learning can influence false recognition of stereotype consistent lure stimuli. However, this effect was only significant for participants who were told they could forget the words after they learned them (directed forgetting). In real life, most people do not actively try to remember everything and there can be occasions where it is necessary to recollect something unintentionally encoded. As memories are reconstructions of past events, being primed with a stereotype can lead to false memories in line with that stereotype.

In terms of the current study, the stereotypes associated with the device may serve as a stereotypical cue, thus influencing false memories in line with the device being used at the time. While future work would be required to fully explore this possibility, it is an important consideration to make given the pervasiveness of devices within the world.

Stereotypes held about smartphones may also provide an explanation for the trend in the results whereby the smartphone group performed worse than the desktop group in all measures bar one (smartphone word recall). Smartphones, as suggested in the findings of Chapter 4, were associated with arguably casual concepts such as ‘game’ and ‘easy’. As such this may lead to use interactions being more casual than desktop interactions which were associated with terms such as ‘work’. This may have then led to a more laisez-faire attitude in smartphone users, resulting in less care being taken during the task, and thus poorer performance in the memory tasks. The recall task in particular required precision in typing because only the words recalled exactly as presented (as described in Chapter 6.3.4) were marked as correct. Therefore, if there was a larger degree of carelessness in the smartphone group due to the stereotypes associated with the device, this may have led to poorer scores in the recall task.

It is however, of interest that the only significant effect was in regards to recognition memory, given prior research (see Chapter 3.4) has indicated that recall is the more susceptible form of memory to interference. Given the mean Pos_{false} scores, it seems likely that a driving factor in the significant recognition effect was due to the smartphone group incorrectly identifying smartphone stimuli, while none of the desktop group falsely identified smartphone stimuli. As such it may be the case that rather than smartphones being detrimental to memory, it is instead the case that they have a stronger influence in promoting false memories. However, further research would be required to discriminate between these concepts, and understand the nature of the effect.

6.6 Limitations

A limitation of this study was the sample size. Twenty participants were recruited, which was less than half the target recruitment number identified in the pre-registration (25 participants per group). Consequently, the study was likely to be underpowered. This means that the results of this study are only indicative, and there is a need for further investigation to fully understand this phenomena.

Another limitation of this study is that while it was designed to be conducted under laboratory conditions, this was not possible due to public safety restrictions. This led to two major limitations of this work. First, there was little control over the device the participant used. This may have resulted in the desktop group using laptops, and the smartphone group using tablets. As discussed in previous chapters, these devices– while having similar interaction modalities– are not interchangeable in terms of the associations people make with them. Therefore, the stimuli may have not have been related as strongly to the devices being used, limiting the amount they were attended to. Second, participants may have been subject to extraneous variables that impacted their performance, such as other devices, associated with different stereotypes and concepts, being present. In particular, the smartphone users may have been subject to push notifications during the task, which would have served as a distractor, potentially contributing to their poor performance. However, while the task was not designed to be ecologically valid, under real world conditions, a person using their phone to encode and retrieve stimuli would be subject to such notifications unless they opted to turn them off. As such, while this may have contributed to the smartphone groups poorer performance, it is reflective of real world interactions with smartphone devices.

This lack of ecological validity was further reflected in the nature of the task, where participants were presented the stimuli in isolation, and retrieved them without context. Utilising a context driven task would have improved the ecological validity, and would potentially strengthened the salience of the device as a cue. For example, if the participants had been presented with a narrative that incorporated the words, this may have strengthened the link between the devices and the words. Drawing greater attention to the link between the stimuli and the device may have led to more notable differences in memory performance. Utilising a narrative would also have allowed for a comprehension test to be given, which would have allowed an insight into whether or not a device linked to social or relaxing contexts impacted how well an individual comprehends information on that device. This would then give insight into important issues such as the spread of misinformation, by investigating a potential cause of incorrectly recalled and comprehended news. However, taking this approach would have also introduced more confounding variables. For example, when utilising narratives, multiple passages of text would be needed in order to counterbalance the stimuli. However, if one narrative is notably more engaging or emotive than the other, this may lead to the content being more memorable, and influencing performance.

The original design of this study aimed to explore the effect of changing devices between encoding and retrieval to investigate environmental context dependency effects, but this was not possible to carry out as part of a remote study. As such, while there appears to be a trend of smartphone users performing worse than desktop users, it is unclear whether this is due to effects at encoding, retrieval or both. Therefore, the current study gives limited insight into the nature of the trends seen in this study, and does not give insight into the effects of device switching. Further research utilising the original design would thus be beneficial in order to gain insight into whether changing devices between encoding and retrieval impacts memory performance, and whether the effects of the device are specifically influencing encoding or retrieval.

Finally, the stimuli used in this experiment were generated by a different sample. As such, it is not necessarily the case that the stimuli presented were strongly associated with desktops and smartphones by the participants included in this study. However, asking participants to generate the stimuli prior to the study would have run the risk of potentially biasing their responses as they may have become aware of the independent variable in the study. Furthermore, participants may generate too few or too many stimuli, leading to a ceiling effect (where too few stimuli causes participants remember everything correctly) or a floor effect (where too many stimuli cause the participant to reach their maximum memory capacity and there is an effect of people under-performing). As such to mitigate these potential issues, the present study used previously generated stimuli, despite the limitations.

6.6.1 Future Work

Replication of the present study with a larger sample size that is split evenly between the groups would be beneficial to further understand the extent of the trends observed in the present study. Furthermore, it would be beneficial to carry out such a replication under controlled conditions as originally intended, so as to account for any extraneous variables. It may also be useful to carry out a version of this study using an incidental memory test, where participants are asked to generate stimuli and after the generation told they are to recall the previously generated words. However such an approach carries limitations, as outlined when discussing the limitations of the present study.

This work could also be expanded by adding in additional conditions where participants change the device they use between encoding and retrieval. This would allow more insight into the effects of changing devices on memory processes, and also identify if the poorer performance in the smartphone group seen in the present work is due to difficulties in the encoding or retrieval stage. This would have real world implications, as it is common for people to use multiple devices throughout the day, and people may complete tasks across multiple devices, influencing cognitive performance.

Finally, further research would also benefit from investigating what underlying mechanisms are influencing memory performance in this task. As suggested by the results of the systematic review (Chapter 2), many mechanisms appear to underpin the effects of devices on memory performance, rather than devices being detrimental to memory itself. While the present study outlines the potential role of stereotypes, the study itself is not proof of this. This study, when taken in conjunction with the results of the stroop task, appears to suggest that increased attention given to a stimulus related to the device does not lead to benefits in memory performance. As such further investigation is required to understand the mechanisms underpinning the results of the present study, where there were differences in memory performance, that appear unrelated to stimulus type.

6.7 Conclusions

The results of this study suggest that smartphone users may under perform in memory tasks compared to their desktop user counterparts, although this was only significant in the case of recognition memory of smartphone words. While this study is limited in terms of it's sample size, it does suggest that smartphones may have a negative effect on memory performance. This may be due to the associations held of phones as they tend to be more casual in nature, while desktops have a stronger relation to work, and utilitarian tasks. Poorer memory when using smartphones has implications for people of lower socio-economic statuses, as they are less likely to have multiple devices, and if they have a single internet enabled device, it tends to be a smartphone as it is portable, and able to make phone calls. This means they would be more affected by the negative effects found for smartphone users, which may exacerbate existing issues seen in regards to the digital divide.

Chapter 7

Discussion and Conclusions

This thesis has explored the effect of everyday devices and the stereotypes held about those devices on two cognitive abilities: information processing and memory. Despite the pervasiveness of technology within the environment, there has been little research investigating the cognitive effects of using these devices, despite concern over this topic in academia [48] and the media [79]. There has also been limited research differentiating between common devices (such as desktops and smartphones), and how these devices may differentially impact cognition.

Research within psychology (described in Chapter 3.3) has indicated that when a cue is related to the target answer in a cognitive test, performance is improved. In terms of technology, there is some evidence (see Chapter 3.2) that devices can hold consistent associations and stereotypes, which may then lead to them serving as a cue within the environment when they are present. However, as shown in Chapter 3.5, there is limited work tying these concepts together. Consequently, the present thesis aimed to conduct a novel investigation into whether devices and the stereotypes held about them could impact cognitive performance.

To do this, the following research questions were explored:

RQ1: What is currently known about how memory is affected by the device being used?

RQ2: How are everyday devices used, and what associations or stereotypes does the user hold about these devices?

RQ3: What effect do devices have on information processing, and is this influenced by whether the stimuli are related to the device being used?

RQ4: How is memory for word lists affected by devices, and is this influenced by whether the words are associated with the device being used?

Multiple methods were used to answer these questions. In response to question one, a systematic review of the literature in both computer science and psychology was conducted, and the results were analysed thematically to identify the trends in the research findings. Question two was addressed using an online survey, comprised of a mixture of open and closed questions and two interactive tasks. Questions three and four were answered using experimental procedures.

This range of methodologies produced an array of findings, which overall suggest a trend in the stereotypes held about each individual device. There is further evidence suggesting that these differences may play a role alongside the tangible device features, in impacting cognitive performance, however, this varied depending on the cognitive process being investigated.

The next section of this chapter will discuss the main findings of each research question. These findings will then be synthesised in relation to each other. The limitations of the work will then be discussed, followed by the potential directions for future work, with reference to specific application areas. The overall conclusions of this thesis will then be presented.

7.1 Main Findings

The effects of technology on memory are not uniform. The findings from the systematic review showed that the effects of technology on memory could be positive, negative or neutral. In terms of device, devices that had a negative effect in one study could have a positive effect in another. Similarly, the effect of technology on memory varied with the type of memory being studied. For example, spatial memory was consistently negatively affected by the use of SatNavs, but for semantic memory, the nature of the effect of technology on memory varied. The impact on spatial memory may be best explained by the impact of having the device display, dividing attention from the environment. However, the impact on semantic memory may be harder to explain in terms of tangible device features, due to its varying nature. Furthermore, the results from this review suggested that it is oftentimes other processes such as effort or attention that are having a subsequent effect on memory, rather than the device directly impacting memory in isolation.

This finding highlights the need for specificity in reporting the type of memory being assessed rather than memory being treated as a single concept. It is instead a construct made up of multiple facets. This is important in the reporting of future work, as not clearly defining the aspect of memory being study may lead to the nature of the findings not being clear. For example, the research by Holdener [63], reported in Chapter 3.5 reported the use of a recall paradigm, when the description suggested a recognition study. As described in Chapter 3.4, recognition and recall performance can be influenced by the same cues differently. As such, accurately reporting the form of memory being tested ensures a clarity in the scientific understanding of phenomena. Furthermore, the findings highlight that there is a need to differentiate between the types of device being used when considering the way they impact cognition. While the same types of device were used in multiple studies included in the review, the majority did not have a consistently negative or positive effect on memory processes, and different devices within the same study led to different effects. This carries the implication that research in this area should be as specific as possible regarding the device being used, and that it is not necessarily sufficient to compare the effect of a device to an analogue option, as there may be variable effects between technological devices.

Everyday devices are used and perceived differently to each other. In the online survey comparing use trends of desktops, laptops, smartphones, and tablets, there were clear trends of devices being used differently to each other. This persisted between devices with similar qualities. For example smartphones and tablets– despite both being portable, touchscreen devices– were used significantly different amounts in terms of time spent on social media and time spent emailing. Further to this, the devices also varied in terms of how they were perceived, both in existing scales used to measure stereotypical properties and in the free association and multiple choice association tasks.

This finding supports the first in that everyday devices are not stereotyped the same way and therefore should not be treated as a single entity. In particular, devices that may be similar in terms of interfaces and hardware properties are still perceived and used in unique ways. A prominent example of this within the present research is the difference in stereotypes held about smartphones and tablets, whereby the devices were used differently, and stereotyped differently across multiple scales, despite sharing the same interaction method and software. This then supports the idea that the technology used to undertake a cognitive task may bias an individual's performance in that task, due to unconscious associations made with the device.

Words related to the device being used can slow down information processing. This finding, stemming from the experiments detailed in Chapter 5, is indicative that performance in tasks can vary based on the device being used. In particular, this occurs when the stimuli being presented is associated with the device being used. The association causes the individual participating in the study to take more time processing the information.

This delayed processing speed indicates that the amount of attention given to stimuli related to the device being used increases, which has implications in terms of task performance and processes influenced by attentional resources. While the effect in this study was modest, in cases where the user holds strong stereotypical views of their device, this effect may increase. Furthermore, this evidences that the way an individual perceives their device can influence cognition, and that considerations should be made regarding which devices are used, as this may influence the outcomes of a given task.

Using a smartphone to complete a memory task leads to variations in performance when compared to using a desktop. This finding from Chapter 6 suggests that individuals using a smartphone to encode and retrieve stimuli tended to remember less than those using desktops at encoding and retrieval. While the data generally trended in this direction this was only significant for recognition memory of smartphone related words, where the desktop group recognised significantly more smartphone related words than the smartphone group.

This may suggest that those who use smartphones to encode and retrieve information may perform worse in memory tasks than those who use desktops. This may be due to the associations held about the devices, in that they are more casual, while desktops are considered more work orientated. This has implications for those who cannot afford multiple devices, as they are more likely to rely on a smartphone for all their technological needs, and thus may be more frequently impacted by this detrimental effect on memory. However, this study was limited by being underpowered in terms of sample size. Consequently, further research is required to understand whether the non-significant results were a true reflection of reality.

7.2 Synthesis

The findings from this thesis can be synthesised into two main themes: the effect of the device being used, and the effect of the associations and stereotypes held about the device.

The effect of the device being used was evidenced in Chapter 4 in the analysis regarding the ‘current device’, whereby participants using a given device would rate that device as more hedonic and utilitarian than the other devices. Further evidence of this effect was presented in Chapter 6 where the results showed an overall trend of smartphone users remembering fewer stimuli than the desktop users. This suggests that people reliant on smartphones may be at a disadvantage when undertaking such tasks. The use-trend data from Chapter 4, suggests that nearly all participants in the study used a smartphone three or more times a week, making it the most commonly used device. The data also showed that younger people were more likely to have access to fewer devices, as were people from lower income groups. With smartphones being the most common device, young people and people of lower incomes are more likely to have a smartphone as their one device, and this was supported by the results of Chapter 4, where of the three participants that reported having frequent access to one device, that device was a smartphone for all of them. Consequently, they may be at a cognitive disadvantage compared to their older, or higher income counterparts, as they are reliant on a single device for all tasks, rather than being able to choose the most appropriate device for a given task.

A potential connotation of this finding is the effect of device on memory for news. Many people now read the news via social media sites which are largely accessed via smartphones, as indicated in the results of Chapter 4. However, given the trend identified in Chapter 6 that smartphone users seem to recall less information than desktop users, and are more likely to falsely recognise information than desktop users, this may mean that consuming news in this way may lead to improperly recalled memories, and the potential spread of misinformation based off of false recollections. Accessing news via technological devices and the potential this has to distort memory has been summarised in a book chapter by Clinch *et al.* [27], however none of the work highlighted in their chapter considered the role of different devices on these memory distortions. As such, the present thesis highlights that the device used can influence performance, and consequently the device used to view news articles may influence the accuracy of remembering these articles. If smartphones are detrimental to memory performance, this may then contribute to the unintentional spread of misinformation. However, there is a need for further research to investigate if such an effect exists, as the majority of research in this domain focusses on screen size rather than the nature of the device, as reported in Chapter 3.5. While Chapter 6 begins to contribute to this area, it is not in an ecologically valid manner, and the small sample size limits the contribution of the research.

While the evidence from Chapter 5 suggested stimuli associated with the device being used were attended to more (and thus processed slower) and the evidence from Chapter 2 suggested that increased attention should aid memory performance, Chapter 6 did not support the notion that stimuli associated with the device being used is better remembered. Instead Chapter 6 found there were no significant differences in memory performance for words associated with the device compared to unassociated words. Similarly this detracts from the idea that environmental cues that relate to the stimuli should aid memory, although the environmental context dependency effect was not investigated fully in Chapter 6. However, it is possible that the stereotypes held about smartphones may play a role in the reduced memory performance of smartphone users. Smartphones were found to be associated with casual concepts and social stereotypes in Chapter 4. These associations may have influenced task performance, as less care may have been taken by the smartphone group whilst undertaking the memory task. Evidence from Chapter 4 supports this, as there were significant effects of the device being used to complete the study on how devices were perceived. This suggests that the device used may bias a person's behaviour, thus leading to differences in performance.

The thesis also suggested wider findings regarding the effect of stereotypes held about devices. The results of Chapter 4 suggested that individuals can hold stereotypes relating to the device's hedonic and utilitarian properties, or how the device's user is perceived in terms of the stereotype content model. There were further associations with the device beyond the scope of these existing scales in that smartphone devices seem to be considered more social, and casual than desktop devices. There was also a trend of desktop devices being related to negative words such as 'heavy' and 'slow' while smartphones were associated with brands. There was a further trend of desktops being related to 'work' while smartphones were associated with social activities such as 'messaging'. These associations and stereotypes indicate that some devices may be more strongly associated with particular tasks, for example desktops being more strongly work related, and this may cause the devices to be more appropriate for certain tasks.

This is supported by evidence from the experiments reported in Chapter 5, where more attention was paid to the stimuli associated with the device being used, thus delaying information processing speeds. For example the word 'work' took longer to respond to when using a desktop than a smartphone. This is evidence that the device being used can impact cognitive processes in different ways, depending on the device being used. This is particularly evidenced in that there were no consistent significant main effects of device on response time, and there were significant interactions between the type of device and the type of stimulus. This has potential ramifications for research utilising stroop paradigms, as words such as 'slow' may be used to investigate stereotypes held about older adults but participant responses may be influenced if the task is undertaken on a desktop. As such, care should be taken to control for device stereotypes, particularly in studies where multiple devices are being used interchangeably.

The stereotypes held about devices may also be of relevance to research in human-computer interaction (HCI). As stated in Chapter 3.2, the use of stereotypes in HCI have largely pertained to the human user, rather than the device. These have particularly taken the form of personas of device users, which designers use to imagine different consumer bases. However, the evidence that stereotypes are held about devices, and that these stereotypes can be activated when a given device is used has implications for future methods of design. For example, laptops and desktops may be given the 'device persona' of being utilitarian and for work. Therefore, software designed for these devices may wish to lean into this persona, and be optimised for these types of tasks, which may then facilitate the user's interactions.

However, in synthesising the results from this thesis, it is hard to completely separate out which findings are due to tangible features, and which are due to intangible features. This is particularly due to the fact that there are limited devices that carry the same interaction methods and different stereotypes and vice versa. Consequently, while the present work gives preliminary evidence that stereotypes held about devices and associations made with devices can effect cognitive processes, this research is at an early stage. Therefore, further research is required to understand the areas in which the cognitive effects of devices are due to the device's features, and which are due to the stereotypes that are held about devices by the end user.

7.3 Limitations

While the work presented in this thesis has implications for multiple areas, it is not without limitations. One major limitation of this work is that the experiments in Chapters 5 and 6 were designed to be run in highly controlled laboratory conditions to reduce the number of confounding variables. However, due to restraints to face-to-face research in 2020/21 this was not possible. Consequently, the studies were conducted remotely, with individuals participating in their own homes where there may have been more background noise or other technologies present. The latter is particularly problematic as the hypotheses were based on the idea that the associations made with technology present within the environment would influence performance in cognitive tasks. As such the results of these experiments may have been influenced to cause an effect where one wouldn't have been or conversely led to an effect being masked. Therefore there is a need to replicate this work in controlled environments to see if the effect persists. Nonetheless, the experiments reported in these chapters do give initial evidence in favour of the idea that stereotypes associated with a device can influence cognition, and indicate this is an important research area to pursue.

A further limitation is that the research was limited in scope in terms of the devices being used. In Chapter 4, only laptops, desktops, smartphones and tablets were investigated, and this reduced to smartphones and laptops/desktops (which were treated as a singular device) in Chapters 5 and 6. Treating laptops and desktops as a singular device also goes against the evidence presented in Chapter 4 of this thesis. However, due to the lack of face-to-face research, this factor was difficult to control for. Furthermore, while these devices were chosen due to their pervasiveness within society and similar computing capabilities (as discussed in Chapters 1.1 and 3.1), there are a plethora of other devices used frequently that were omitted from this work. As such, this work has a relatively small scope within the technological landscape which means that other common devices, used for different purposes, may elicit different effects on cognition. Furthermore, the treatment of laptops and desktops as a single construct in Chapters 5 and 6 means that the stimuli may not be as strongly associated with the device used as they were originally elicited for desktops specifically. Nonetheless the findings did find some significant effects and these may have stronger effect sizes should future research better tailor the stimuli to the device being used.

Sample size is also a limitation, particularly in regards to Chapter 6. As recruitment was conducted online, there were issues in recruiting participants to the study, particularly as there was no incentive to participate. Consequently the research may not be sufficiently powered and thus the results may not be conclusive. Given this, it is important that future work is undertaken to replicate and verify the findings, to investigate the robustness of the present work. Additionally, the work largely sampled from the UK (although it was advertised online and thus may have sampled from a wider demographic) and was limited to native or fluent English speakers. The studies also required participants to use their own devices and this may have skewed the sample towards higher income individuals. As such, the work may be bias towards these demographics, and may lack generalisability to wider samples.

7.4 Future Work

The findings of this thesis indicate that the device being used and the associations made with that device, can influence cognitive performance. However as discussed previously, replication of this work is needed to gain a deeper understanding of how consistent these effects may be. As noted in Chapter 7.3, more confounding variables were potentially present in the experiments than desired, due to the need to conduct this research remotely. Future work would also benefit by expanding the number of devices investigated, to gain a more holistic view of the associations and stereotypes held by different devices, leading to consequent work regarding their effect on cognition.

Given the consistent trend in devices holding stereotypes found in Chapter 4, it may also be beneficial to develop new scales designed to investigate device stereotypes specifically. The SCM used in this work was designed to be applied to humans rather than objects, and as such the findings may be confounded by the participant's mental representation of what the user looks like. Furthermore, as discussed in Chapter 4.5.2, the hedonic and utilitarian scale did not use terms that were naturally generated in the free association task. This suggests that the words used to measure these constructs are not ones that people would associate with a device unprompted. Therefore, clearer insights into the stereotypes held about devices may be garnered by developing a device specific stereotype scale.

While the impact of stereotypes held about devices presented in this work are not consistent or conclusive, as technology continues to develop, and the relationship an individual has with their device grows, it is likely that the perceptions and stereotypes an individual holds about their device will grow too. Consequently, the importance of research investigating the impact of stereotypes held about devices will likely grow too. Continuing research in this vein will therefore be necessary to understand how these developing relationships affect people. Furthermore, it is important to develop the understanding of these effects as early as possible. Little is known about the effects of technology on cognitive processes, despite the ubiquitous nature of technology, and as seen in recent years, global events such as the COVID-19 pandemic can lead to technology becoming heavily relied on. Research from 2019 by Carolus *et al.* [24] already indicated that people felt closer to their smartphones than their flatmates. Given that the pandemic led to many people being more reliant on technology, it is possible that closeness to devices may have further grown. This closeness may then have subsequent effects on other areas of life. Thus, as technology develops, it is important to consider all the ways technology may be impacting human abilities, so any negative consequences can be mitigated as early as possible. Therefore, while the work presented in this thesis is not conclusive, it is important that research in this domain is pursued, in order to fully understand the extent to which tangible and intangible device features may effect cognition, and thus prevent negative consequences in the future.

The work in this thesis also lacked ecological validity by design, and so the next step in advancing knowledge in this area should aim to investigate the extent to which these effects persist in real life scenarios. To do this, further work should be conducted to investigate any real world connotations of the findings in Chapters 5 and 6. This could be done by asking participants to review news stories on either a smartphone or desktop, that were either related to one of the devices and the stereotypes held about the device or neutral. Participants would then be asked to retrieve information from these stories to investigate any effects. A comprehension task may also be added similar to that described in Chapter 6.6 (page 146). This would then give insight into the effects of devices and the stereotypes held about them on memory in real life. Furthermore this study could utilise eye tracking to verify the information processing delays suggested in the findings of Chapter 5, where greater time spent looking at elements of the text would be indicative of greater processing time. A further study, investigating the impact of device switching between encoding and retrieval would also give insight into how device switching may be impacting memory performance in everyday life. If this work establishes real world implications, then subsequent work could be conducted in specific application areas, as discussed in the following sections.

7.4.1 Education

The extent to which technology impacts cognitive processes is becoming increasingly relevant in the field of education, as many students in higher education are expressing a preference for using technology during lessons rather than note-taking via analogue methods [58]. In the context of the findings of the present work, it is possible that students will not process the content as deeply if the device they are using is not one they associate with their work, which may have a negative impact on learning outcomes and educational attainment. Furthermore, the use of smartphones in general appears to lead to a trend of poorer memory performance, which may be to the detriment of students.

Moreover, while the present thesis conducted limited investigation into changes in devices in the environment, this is potentially important to investigate in the context of education in future work. For example, most students will attend lectures and classes where the learning materials are projected on a large screen for the class to see. Students will use their individual devices such as laptops to take notes from the projected content. They will then revise from these digital notes for exams, and some may even use their smartphone to review notes prior to entering the exam hall. However, after utilising multiple devices to learn, many exams are still paper based, as it prevents the use of the Internet to look up information. This changing of devices may cause poorer cognitive performance, given the preliminary evidence in this thesis that devices can serve as cues.

Some research has been conducted to investigate the use of screens within educational contexts. For example, Subrahmanyam *et al.* [134] investigated the use of laptops, tablets, and analogue methods on reading comprehension. The results did not find any significant effects, however the study did not consistently distinguish between laptops and tablets. Furthermore, the study gave limited consideration to the potential effects of stimuli on the results. One of the passages related to memories of spending time in the library, an image that may elicit associations with physical books and paper. The second passage pertained to the social theories on the effects of television on people, which may elicit associations with other technological devices. As such it is unclear if these may have had an influence on the results of the study, as the evidence presented in this thesis suggests it might. Other research in this area (e.g. [43, 56]) found negative effects of laptop use on memory, but again did not consider the extent to which the nature of the content being learnt may interact with the device being used. Therefore, ecologically valid research regarding the effect of technology on aspects of the curriculum may be useful in understanding the nuances of device use in educational settings.

7.4.2 Media Consumption

Media is now available to watch across a multitude of devices, rather than just on televisions. In particular, most streaming platforms (e.g. Netflix, Disney+) have apps that allow people to stream shows and films on any smart device. However research in this area tends to focus on the effect of screen size on immersion [112] or the effect of multi-screening (using a second device) while watching television [92], rather than the potential impact the device being used to watch television may have on attention and memory for the content being consumed.

The stereotypes held about everyday computing devices evidenced in this thesis may have the potential to be leveraged by broadcasters. Understanding the ways different devices relate to cognitive abilities such as attention and memory, may aid in the understanding of consumer use trends. This may then contribute to existing research that has sought to understand media consumer engagement [23] by giving an insight into the user's stereotyped attitudes based on their device. For example, if a desktop is being utilised, the findings from Chapter 5 suggest that there will be greater attention paid to desktop-related stimuli, as indicated by the delayed response time for information processing. Consequently, this may indicate to broadcasters that consumer interaction will be improved when they are watching a documentary (which may be considered a more utilitarian genre of media) on a desktop. As such, recommender systems and targeted advertising may be utilised based on the form of device used, in order to best increase audience engagement. As such, future research may wish to test this, and investigate whether genres of media congruent with the stereotypes held about the device being used impact engagement.

7.5 Conclusions

In summary, this thesis explored the effect of different devices and the associations and stereotypes held about those devices, on cognitive processes. The prior literature in this area was explored using a systematic literature search, with a focus on memory processes. This was followed by a survey to investigate the associations and stereotypes held about devices, as well as to investigate the general usage trends of four everyday computing devices. The latter half of this work investigated if information processing and memory were influenced when the stimuli presented were associated with the device being used.

The results of the thesis suggest that individuals hold stereotypes and associations about the device they use, and the effect of technology on cognition can vary between devices. In terms of memory, there is a trend for memory performance to be worse on smartphones than desktops, but this is only significant for the recognition of smartphone-related stimuli. For information processing, terms associated with the device being used can lead to delays in processing speeds.

These results have implications for those who are reliant on one device due to the digital divide, as smartphone users' performance may be worse in cognitive tasks than those who use different devices. Furthermore, the associations held about the device being used may affect individuals, impacting their interaction with devices. However, these associations have the potential to be leveraged by designers to optimise software and interfaces, so they are congruent with the device being used.

Bibliography

[1] ALEX, N. J., AND JOSEPH, A. Hedonic versus utilitarian values: The relative importance of real and ideal self to brand personality and its influence on emotional brand attachment. *Vilakshan: The XIMB Journal of Management 9*, 2 (2012).

[2] ALLÉ, M. C., MANNING, L., POTHEEGADOO, J., COUTELLE, R., DANION, J.-M., AND BERNA, F. Wearable cameras are useful tools to investigate and remediate autobiographical memory impairment: A systematic prisma review. *Neuropsychology review 27*, 1 (2017), 81–99.

[3] ALSOP, T. U.s. households with pc/computer at home 2016, May 2020.

[4] ALSOP, T. Share of households with a computer in africa 2005-2019, Feb 2021.

[5] ANDERSON, M., AND JIANG, J. Teens, social media & technology 2018. *Pew Research Center 31* (2018), 2018.

[6] ANDO, M., AND UENO, M. Analysis of the advantages of using tablet pc in e-learning. In *2010 10th IEEE International Conference on Advanced Learning Technologies* (2010), IEEE, pp. 122–124.

[7] ANTROBUS, V., LARGE, D., BURNETT, G., AND HARE, C. Enhancing environmental engagement with natural language interfaces for in-vehicle navigation systems. *Journal of Navigation 72*, 3 (2019).

[8] APA DICTIONARY OF PSYCHOLOGY. Definition of priming.

[9] ARAYA, T., EKEHAMMAR, B., AND AKRAMI, N. Remembering things that never occured: The effects of to-be-forgotten stereotypical information. *Experimental Psychology 50*, 1 (2003), 27.

[10] ARREOLA, H., FLORES, A. N., LATHAM, A., MACNEW, H., AND VU, K.-P. L. Does the use of tablets lead to more information being recorded and better recall in short-term memory tasks? In *International Conference on Human-Computer Interaction* (2019), Springer, pp. 292–302.

[11] ARTHUR JR, W., KEISER, N. L., AND DOVERSPIKE, D. An information-processing-based conceptual framework of the effects of unproctored internet-based testing devices on scores on employment-related assessments and tests. *Human Performance 31*, 1 (2018), 1–32.

[12] BADDELEY, A. The episodic buffer: a new component of working memory? *Trends in cognitive sciences 4*, 11 (2000), 417–423.

[13] BARKER, V. Older adolescents' motivations for social network site use: The influence of gender, group identity, and collective self-esteem. *Cyberpsychology & behavior 12*, 2 (2009), 209–213.

[14] BARTLETT, F. C., AND BARTLETT, F. C. *Remembering: A study in experimental and social psychology*. Cambridge University Press, 1995.

[15] BATRA, R., AND AHTOLA, O. T. Measuring the hedonic and utilitarian sources of consumer attitudes. *Marketing letters 2*, 2 (1991), 159–170.

[16] BELLIS, C. J. Reaction time and chronological age. *Proceedings of the Society for Experimental Biology and Medicine 30*, 6 (1933), 801–803.

[17] BIRD, S., KLEIN, E., AND LOPER, E. *Natural Language Processing with Python*. O'Reilly Media, 2009.

[18] BLACKER, K. J., CURBY, K. M., KLOBUSICKY, E., AND CHEIN, J. M. Effects of action video game training on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance 40*, 5 (2014), 1992.

[19] BLAIR, K. S., SMITH, B. W., MITCHELL, D. G., MORTON, J., VYTHILINGAM, M., PESSOA, L., FRIDBERG, D., ZAMETKIN, A., NELSON, E. E., DREVETS, W. C., ET AL. Modulation of emotion by cognition and cognition by emotion. *Neuroimage 35*, 1 (2007), 430–440.

[20] BODENHAUSEN, G. V. Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of personality and social psychology 55*, 5 (1988), 726.

[21] BRUYA, B., AND TANG, Y.-Y. Is attention really effort? revisiting daniel kahneman's influential 1973 book attention and effort. *Frontiers in psychology 9* (2018), 1133.

[22] BUHRMESTER, M., KWANG, T., AND GOSLING, S. D. Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science 6*, 1 (2011), 3–5.

[23] CARLTON, J., BROWN, A., JAY, C., AND KEANE, J. Using interaction data to predict engagement with interactive media. In *Proceedings of the 29th ACM International Conference on Multimedia* (2021), pp. 1258–1266.

[24] CAROLUS, A., BINDER, J. F., MUENCH, R., SCHMIDT, C., SCHNEIDER, F., AND BUGLASS, S. L. Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society 21*, 4 (2019), 914–938.

[25] CARTER, L., AND WEERAKKODY, V. E-government adoption: A cultural comparison. *Information systems frontiers 10*, 4 (2008), 473–482.

[26] CHANG, F.-C., CHIU, C.-H., CHEN, P.-H., CHIANG, J.-T., MIAO, N.-F., CHUANG, H.-Y., AND LIU, S. Children's use of mobile devices, smartphone addiction and parental mediation in taiwan. *Computers in Human Behavior 93* (2019), 25–32.

[27] CLINCH, S., MACK, C. C., WARD, G., AND STEEDS, M. Technology-mediated memory impairment. In *Technology-Augmented Perception and Cognition*. Springer, 2021, pp. 71–124.

[28] COLZATO, L. S., VAN DEN WILDENBERG, W. P., ZMIGROD, S., AND HOMMEL, B. Action video gaming and cognitive control: playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological research 77*, 2 (2013), 234–239.

[29] CONSULTANCY.UK. Uk smartphone penetration continues to rise to 85% of adult population.

[30] CONTARELLO, A., AND SARRICA, M. Icts, social thinking and subjective well-being–the internet and its representations in everyday life. *Computers in human behavior 23*, 2 (2007), 1016–1032.

[31] COOPER, A., ET AL. *The inmates are running the asylum:[Why high-tech products drive us crazy and how to restore the sanity].* Sams Indianapolis, 2004.

[32] COVACI, A., GHINEA, G., LIN, C.-H., HUANG, S.-H., AND SHIH, J.-L. Multisensory games-based learning-lessons learnt from olfactory enhancement of a digital board game. *Multimedia Tools and Applications 77*, 16 (2018), 21245–21263.

[33] CRAIK, F. I., GOVONI, R., NAVEH-BENJAMIN, M., AND ANDERSON, N. D. The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General 125*, 2 (1996), 159.

[34] DELIS, D. C. California verbal learning test. *Adult version. Manual. Psychological Corporation* (2000).

[35] DELL, N., VAIDYANATHAN, V., MEDHI, I., CUTRELL, E., AND THIES, W. " yours is better!" participant response bias in hci. In *Proceedings of the sigchi conference on human factors in computing systems* (2012), pp. 1321–1330.

[36] DETENBER, B. H., AND REEVES, B. A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication* (1996).

[37] DEVOS, P., DEBEER, J., OPHALS, J., AND PETROVIC, M. Cognitive impairment screening using m-health: an android implementation of the mini-mental state examination (mmse) using speech recognition. *European Geriatric Medicine 10*, 3 (2019), 501–509.

[38] DODGE, M., AND KITCHIN, R. ‘outlines of a world coming into existence’: pervasive computing and the ethics of forgetting. *Environment and planning B: planning and design 34*, 3 (2007), 431–445.

[39] EICH, E. Context, memory, and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition 11*, 4 (1985), 764.

[40] FISKE, S. T., CUDDY, A. J., GLICK, P., AND XU, J. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology 82*, 6 (2002), 878.

[41] FISKE, S. T., XU, J., CUDDY, A. C., AND GLICK, P. (dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of social issues 55*, 3 (1999), 473–489.

[42] FRANGOU, S.-M., WIKGREN, J., SINTONEN, S., KAIRALUOMA, L., VASARI, P., ET AL. The effect of writing modality on recollection in children and adolescents. *Research in Learning Technology* (2019).

[43] FRIED, C. B. In-class laptop use and its effects on student learning. *Computers & Education 50*, 3 (2008), 906–914.

[44] GARDNER, R. C., AND TAYLOR, D. Ethnic stereotypes: Meaningfulness in ethnic-group labels. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement 1*, 3 (1969), 182.

[45] GARDONY, A. L., BRUNYÉ, T. T., MAHONEY, C. R., AND TAYLOR, H. A. How navigational aids impair spatial memory: Evidence for divided attention. *Spatial Cognition & Computation 13*, 4 (2013), 319–350.

[46] GARDONY, A. L., BRUNYÉ, T. T., AND TAYLOR, H. A. Navigational aids and spatial memory impairment: The role of divided attention. *Spatial Cognition & Computation 15*, 4 (2015), 246–284.

[47] GENTILE, D. Pathological video-game use among youth ages 8 to 18: A national study. *Psychological science 20*, 5 (2009), 594–602.

[48] GIRARDIN, F., AND BLAT, J. The co-evolution of taxi drivers and their in-car navigation systems. *Pervasive and Mobile Computing 6*, 4 (2010), 424–434.

[49] GLASSMAN, N. R., AND SHEN, P. One site fits all: responsive web design. *Journal of Electronic Resources in Medical Libraries 11*, 2 (2014), 78–90.

[50] GODDEN, D., AND BADDELEY, A. When does context influence recognition memory? *British journal of Psychology 71*, 1 (1980), 99–104.

[51] GODDEN, D. R., AND BADDELEY, A. D. Context-dependent memory in two natural environments: On land and underwater. *British Journal of psychology 66*, 3 (1975), 325–331.

[52] GRUDIN, J. Why personas work: The psychological evidence. *The Persona Lifecycle* (2006), 642–663.

[53] HAINES, E. L., DEAUX, K., AND LOFARO, N. The times they are a-changing... or are they not? a comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly 40*, 3 (2016), 353–363.

[54] HART, S. J., GREEN, S. R., CASP, M., AND BELGER, A. Emotional priming effects during stroop task performance. *Neuroimage 49*, 3 (2010), 2662–2670.

[55] HARVEY, M., LANGHEINRICH, M., AND WARD, G. Remembering through lifelogging: A survey of human memory augmentation. *Pervasive and Mobile Computing 27* (2016), 14–26.

[56] HEMBROOKE, H., AND GAY, G. The laptop and the lecture: The effects of multitasking in learning environments. *Journal of computing in higher education 15*, 1 (2003), 46–64.

[57] HEO, N. *The effects of screen size and content type on viewers' attention, arousal, memory, and content evaluations*. PhD thesis, The Pennsylvania State University, 2003.

[58] HERRMANN, K. J., LINDVIG, K., AND AAGAARD, J. Curating the use of digital media in higher education: a case study. *Journal of Further and Higher Education 45*, 3 (2021), 389–400.

[59] HERZ, R. S. The effects of cue distinctiveness on odor-based context-dependent memory. *Memory & Cognition 25*, 3 (1997), 375–380.

[60] HERZ, R. S. Emotion experienced during encoding enhances odor retrieval cue effectiveness. *American Journal of Psychology 110* (1997), 489–506.

[61] HITLIN, P. Internet, social media use and device ownership in us. *Pew Res Cent* (2018).

[62] HODGES, S., WILLIAMS, L., BERRY, E., IZADI, S., SRINIVASAN, J., BUTLER, A., SMYTH, G., KAPUR, N., AND WOOD, K. Sensecam: A retrospective memory aid. In *International Conference on Ubiquitous Computing* (2006), Springer, pp. 177–193.

[63] HOLDENER, A. Does size matter? Master's thesis, University of Geneva, 2008.

[64] HUANG, M.-H. Designing website attributes to induce experiential encounters. *computers in Human Behavior 19*, 4 (2003), 425–442.

[65] HUANG, V., YOUNG, M., AND FIOCCO, A. J. The association between video game play and cognitive function: does gaming platform matter? *Cyberpsychology, Behavior, and Social Networking 20*, 11 (2017), 689–694.

[66] HURTIENNE, J., STÖSSEL, C., AND WEBER, K. Sad is heavy and happy is light: Population stereotypes of tangible object attributes. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (2009), pp. 61–68.

[67] ISHIKAWA, T., FUJIWARA, H., IMAI, O., AND OKABE, A. Wayfinding with a gps-based mobile navigation system: A comparison with maps and direct experience. *Journal of environmental psychology 28*, 1 (2008), 74–82.

[68] JO, J., AND LIM, H. A comparative analysis of the effectiveness of learning english vocabulary between smart device and printed material. *International Journal of Intelligent Information and Database Systems 3 8*, 2 (2014), 150–161.

[69] JOHANSSON, L., LUNDH, L.-G., AND ANDERSSON, G. Attentional bias for negative self-words in young women the role of thin ideal priming and body shape dissatisfaction. *Personality and Individual Differences 38*, 3 (2005), 723–733.

[70] KAHNEMAN, D. *Attention and effort*, vol. 1063. Citeseer, 1973.

[71] KALNIKAITE, V., SELLEN, A., WHITTAKER, S., AND KIRK, D. Now let me see where i was: understanding how lifelogs mediate memory. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2010), pp. 2045–2054.

[72] KANAHARA, S. A review of the definitions of stereotype and a proposal for a progressional model. *Individual Differences Research 4*, 5 (2006).

[73] KAWSAR, F., AND BRUSH, A. B. Home computing unplugged: why, where and when people use different connected devices at home. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (New York, USA, 2013), ACM, pp. 627–636.

[74] KELLEY, M. The effect of screen size and audio delivery system on memory for television news. *Visual Communication Quarterly 14*, 3 (2007), 176–188.

[75] KIM, H.-M. Mobile media technology and popular mobile games in contemporary society. *International Journal of Mobile Marketing 8*, 2 (2013).

[76] KIRCHNER, W. K. Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology 55*, 4 (1958), 352.

[77] KISHIYAMA, M. M., AND YONELINAS, A. P. Novelty effects on recollection and familiarity in recognition memory. *Memory & cognition 31*, 7 (2003), 1045–1051.

[78] KLEIN, G. S. Semantic power measured through the interference of words with color-naming. *The American journal of psychology 77*, 4 (1964), 576–588.

[79] KLEINMAN, Z. Screen use is bad for brain development, scientist claims, May 2013.

[80] KRETZSCHMAR, F., PLEIMLING, D., HOSEMANN, J., FÜSSEL, S., BORNKESSEL-SCHLESEWSKY, I., AND SCHLESEWSKY, M. Subjective impressions do not mirror online reading effort: Concurrent eeg-eyetracking evidence from the reading of books and digital media. *PloS one 8*, 2 (2013), e56178.

[81] KWOK, K., GHREAR, S., LI, V., HADDOCK, T., COLEMAN, P., AND BIRCH, S. A. Children can learn new facts equally well from interactive media versus face to face instruction. *Frontiers in psychology 7* (2016), 1603.

[82] LA HEIJ, W., VAN DER HEIJDEN, A., AND SCHREUDER, R. Semantic priming and stroop-like interference in word-naming tasks. *Journal of Experimental Psychology: Human Perception and Performance 11*, 1 (1985), 62.

[83] LARSEN, R. J., MERCER, K. A., AND BALOTA, D. A. Lexical characteristics of words used in emotional stroop experiments. *Emotion 6*, 1 (2006), 62.

[84] LI, Y., DU, J., SONG, Q., WU, S., AND LIU, L. An erp study of the temporal course of gender–color stroop effect. *Frontiers in Psychology 11* (2020).

[85] LIBERATI, A., ALTMAN, D. G., TETZLAFF, J., MULROW, C., GØTZSCHE, P. C., IOANNIDIS, J. P., CLARKE, M., DEVEREAUX, P. J., KLEIJNEN, J., AND MOHER, D. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine 6*, 7 (2009), e1000100.

[86] LINNMAN, C., CARLBRING, P., ÅHMAN, Å., ANDERSSON, H., AND ANDERSSON, G. The stroop effect on the internet. *Computers in human behavior 22*, 3 (2006), 448–455.

[87] LIU, N., AND YU, R. Identifying design feature factors critical to acceptance and usage behavior of smartphones. *Computers in Human Behavior 70* (2017), 131–142.

[88] LIU, Y., AND WANG, D. How does the device change your choice: A goal-activation perspective. In *International Conference on HCI in Business, Government, and Organizations* (2016), Springer, pp. 446–456.

[89] MACKAY, D. G., SHAFTO, M., TAYLOR, J. K., MARIAN, D. E., ABRAMS, L., AND DYER, J. R. Relations between emotion, memory, and attention: Evidence from taboo stroop, lexical decision, and immediate memory tasks. *Memory & cognition 32*, 3 (2004), 474–488.

[90] MANGEN, A., ANDA, L. G., OXBOROUGH, G. H., AND BRØNNICK, K. Handwriting versus keyboard writing: effect on word recall. *Journal of writing research 7*, 2 (2015).

[91] MANGEN, A., OLIVIER, G., AND VELAY, J.-L. Comparing comprehension of a long text read in print book and on kindle: Where in the text and when in the story? *Frontiers in psychology 10* (2019), 38.

[92] MARINELLI, A., AND ANDÒ, R. Multiscreening and social tv: The changing landscape of tv consumption in italy. *VIEW Journal of European Television History and Culture 3*, 6 (2014), 24–36.

[93] MARSDEN, N., AND HAAG, M. Stereotypes and politics: reflections on personas. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), ACM, pp. 4017–4031.

[94] MAXWELL, S. E., LAU, M. Y., AND HOWARD, G. S. Is psychology suffering from a replication crisis? what does "failure to replicate" really mean? *American Psychologist 70*, 6 (2015), 487.

[95] MAYER, C., WALLNER, S., BUDDE-SPENGLER, N., BRAUNERT, S., ARNDT, P. A., AND KIEFER, M. Literacy training of kindergarten children with pencil, keyboard or tablet stylus: The influence of the writing tool on reading and writing performance at the letter and word level. *Frontiers in psychology 10* (2020), 3054.

[96] MCGREGOR, M., BIDWELL, N. J., SARANGAPANI, V., APPAVOO, J., AND O'NEILL, J. Talking about chat at work in the global south: An ethnographic study of chat use in india and kenya. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), ACM, p. 233.

[97] MEAGHER, B. R. Judging the gender of the inanimate: Benevolent sexism and gender stereotypes guide impressions of physical objects. *British Journal of Social Psychology 56*, 3 (2017), 537–560.

[98] MENKES, S. M. *Children's media comprehension: The relationship between media platform, executive functioning abilities, and age*. PhD thesis, The Claremont Graduate University, 2012.

[99] MIAH, M., AND OMAR, A. Technology advancement in developing countries during digital age. *International Journal 1*, 1 (2012).

[100] MONTEFINESE, M., ZANNINO, G. D., AND AMBROSINI, E. Semantic similarity between old and new items produces false alarms in recognition memory. *Psychological research 79*, 5 (2015), 785–794.

[101] MOST, S. B., SORBER, A. V., AND CUNNINGHAM, J. G. Auditory stroop reveals implicit gender associations in adults and children. *Journal of Experimental Social Psychology 43*, 2 (2007), 287–294.

[102] NAVEH-BENJAMIN, M., CRAIK, F. I., AND BEN-SHAUL, L. Age-related differences in cued recall: Effects of support at encoding and retrieval. *Aging, Neuropsychology, and Cognition 9*, 4 (2002), 276–287.

[103] OFCOM. The communications market 2018: Interactive report, May 2019.

[104] O'HARA, K., MORRIS, R., SHADBOLT, N., HITCH, G. J., HALL, W., AND BEAGRIE, N. Memories for life: a review of the science and technology. *Journal of the Royal Society Interface 3*, 8 (2006), 351–365.

[105] PITICHAT, T. Smartphones in the workplace: Changing organizational behavior, transforming the future. *LUX: A Journal of Transdisciplinary Writing and Research from Claremont Graduate University 3*, 1 (2013), 13.

[106] QU, C., AND SAS, C. Exploring memory interventions in depression through lifelogging lens. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32* (2018), pp. 1–6.

[107] RAGOT, M., MARTIN, N., AND COJEAN, S. Ai-generated vs. human artworks. a perception bias towards artificial intelligence? In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts* (2020), pp. 1–10.

[108] REEVES, B., LANG, A., KIM, E. Y., AND TATAR, D. The effects of screen size and message content on attention and arousal. *Media psychology 1*, 1 (1999), 49–67.

[109] REICH-STIEBERT, N., AND EYSSEL, F. (ir) relevance of gender? on the influence of gender stereotypes on learning with a robot. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI* (2017), IEEE, pp. 166–176.

[110] RICHTER, A., AND COURAGE, M. L. Comparing electronic and paper storybooks for preschoolers: Attention, engagement, and recall. *Journal of Applied Developmental Psychology 48* (2017), 92–102.

[111] RIEGER, C., AND MAJCHRZAK, T. A. Conquering the mobile device jungle: Towards a taxonomy for app-enabled devices. In *WEBIST* (2017), pp. 332–339.

[112] RIGBY, J. M., BRUMBY, D. P., COX, A. L., AND GOULD, S. J. Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services adjunct* (2016), pp. 714–721.

[113] ROEDIGER, H. L., AND MCDERMOTT, K. B. Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition 21*, 4 (1995), 803.

[114] RYAN, C. L., AND LEWIS, J. M. *Computer and internet use in the United States: 2015*. US Department of Commerce, Economics and Statistics Administration, US ..., 2017.

[115] SAUFLEY, W. H., OTAKA, S. R., AND BAVARESCO, J. L. Context effects: Classroom tests and context independence. *Memory & Cognition 13*, 6 (1985), 522–528.

[116] SCHWARTZ, L. A., AND MARKHAM, W. T. Sex stereotyping in children's toy advertisements. *Sex Roles 12*, 1-2 (1985), 157–170.

[117] SCHWIND, V., DEIERLEIN, N., POGUNTKE, R., AND HENZE, N. Understanding the social acceptability of mobile devices using the stereotype content model. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), ACM, p. 361.

[118] SELWYN, N. The information aged: A qualitative study of older adults' use of information and communications technology. *Journal of Aging studies 18*, 4 (2004), 369–384.

[119] SELWYN, N., GORARD, S., FURLONG, J., AND MADDEN, L. Older adults' use of information and communications technology in everyday life. *Ageing & Society 23*, 5 (2003), 561–582.

[120] SHERIF, M., HARVEY, O., WHITE, B. J., HOOD, W., AND SHERIF, C. Intergroup conflict and cooperation: the robbers cave experiment norman. *OK University Book Exchange* (1961).

[121] SHU, Y., HUANG, Y.-Z., CHANG, S.-H., AND CHEN, M.-Y. Do virtual reality head-mounted displays make a difference? a comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Reality 23*, 4 (2019), 437–446.

[122] SIEGRIST, M. Effects of taboo words on color-naming performance on a stroop test. *Perceptual and Motor Skills 81*, 3_suppl (1995), 1119–1122.

[123] SLATER, M., USOH, M., AND STEED, A. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI) 2*, 3 (1995), 201–219.

[124] SMITH, G. Step away from stepwise. *Journal of Big Data 5*, 1 (2018), 1–12.

[125] SMITH, S. M., GLENBERG, A., AND BJORK, R. A. Environmental context and human memory. *Memory & Cognition 6*, 4 (1978), 342–353.

[126] SMITH, S. M., AND VELA, E. Environmental context-dependent memory: A review and meta-analysis. *Psychonomic bulletin & review 8*, 2 (2001), 203–220.

[127] SPANAKIS, P., JONES, A., FIELD, M., AND CHRISTIANSEN, P. A stroop in the hand is worth two on the laptop: Superior reliability of a smartphone based alcohol stroop in the real world. *Substance use & misuse 54*, 4 (2019), 692–698.

[128] SPARROW, B., LIU, J., AND WEGNER, D. M. Google effects on memory: Cognitive consequences of having information at our fingertips. *science 333*, 6043 (2011), 776–778.

[129] STAWARZ, K., COX, A. L., AND BLANDFORD, A. Don't forget your pill! designing effective medication reminder apps that support users' daily routines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), pp. 2269–2278.

[130] STEEDS, M. The impact of device associations on human memory performance. In *Proceedings of the 33rd International BCS Human Computer Interaction Conference* (July 2020).

[131] STEEDS, M., CLINCH, S., AND JAY, C. Uk device use and associations (duaa), 2020.

[132] STEEDS, M., CLINCH, S., AND JAY, C. Device uses and device stereotypes. *Computers in Human Behavior Reports* (2021), 100100.

[133] STROOP, J. R. Studies of interference in serial verbal reactions. *Journal of experimental psychology 18*, 6 (1935), 643.

[134] SUBRAHMANYAM, K., MICHIKYAN, M., CLEMMONS, C., CARRILLO, R., UHLS, Y. T., AND GREENFIELD, P. M. Learning from paper, learning from screens: Impact of screen reading and multitasking conditions on reading and writing among college students. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL) 3*, 4 (2013), 1–27.

[135] TAJFEL, H. Experiments in intergroup discrimination. *Scientific american 223*, 5 (1970), 96–103.

[136] TAY, B., JUNG, Y., AND PARK, T. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior 38* (2014), 75–84.

[137] TAYLOR, S. E. The interface of cognitive and social psychology. *Cognition, social behavior, and the environment* (1981), 189–211.

[138] TOBOLA, C. Left to our own devices: How income, age, and sex affect type and number of electronic devices that individuals use. *Journal of the Communication, Speech & Theatre Association of North Dakota 22* (2009), 31–43.

[139] TSETSI, E., AND RAINS, S. A. Smartphone internet access and use: Extending the digital divide and usage gap. *Mobile Media & Communication 5*, 3 (2017), 239–255.

[140] TULVING, E. Memory and consciousness. *Canadian Psychology/Psychologie canadienne 26*, 1 (1985), 1.

[141] TURNER, P., AND TURNER, S. Is stereotyping inevitable when designing with personas? *Design studies 32*, 1 (2011), 30–44.

[142] UNDERWOOD, B. J. False recognition produced by implicit verbal responses. *Journal of experimental psychology 70*, 1 (1965), 122.

[143] VAN DIJK, J. *The digital divide*. John Wiley & Sons, 2020.

[144] VANNINI, S., NEMER, D., AND REGA, I. Integrating mobile technologies to achieve community development goals: the case of telecenters in brazil. In *Proceedings of the 8th international conference on communities and technologies* (2017), pp. 115–124.

[145] WÄSTLUND, E., NORLANDER, T., AND ARCHER, T. The effect of page layout on mental workload: A dual-task experiment. *Computers in human behavior 24*, 3 (2008), 1229–1245.

[146] WATTS, F. N., MCKENNA, F. P., SHARROCK, R., AND TREZISE, L. Colour naming of phobia-related words. *British journal of Psychology 77*, 1 (1986), 97–108.

[147] WEI, F.-Y. F., WANG, Y. K., AND FASS, W. An experimental study of online chatting and notetaking techniques on college students' cognitive learning from a lecture. *Computers in Human Behavior 34* (2014), 148–156.

[148] WHITE, J. B., AND GARDNER, W. L. Think women, think warm: Stereotype content activation in women with a salient gender identity, using a modified stroop task. *Sex roles 60*, 3-4 (2009), 247–260.

[149] WILLIAMS, J. M. G., MATHEWS, A., AND MACLEOD, C. The emotional stroop task and psychopathology. *Psychological bulletin 120*, 1 (1996), 3.

[150] WILSON, M. Six views of embodied cognition. *Psychonomic Bulletin & Review 9*, 4 (2002), 625–636.

[151] WOLFERS, L. N., KITZMANN, S., SAUER, S., AND SOMMER, N. Phone use while parenting: An observational study to assess the association of maternal sensitivity and smartphone use in a playground setting. *Computers in Human Behavior 102* (2020), 31–38.

[152] WRIGHT, D. L., AND SHEA, C. H. Contextual dependencies in motor skills. *Memory & Cognition 19*, 4 (1991), 361–370.

[153] YARDI, S., AND BRUCKMAN, A. Income, race, and class: exploring socioeconomic differences in family technology use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 3041–3050.

[154] YERKES, R. M., DODSON, J. D., ET AL. The relation of strength of stimulus to rapidity of habit-formation. *Punishment: Issues and experiments* (1908), 27–41.

[155] YOUN, S., SUN, T., WELLS, W. D., AND ZHAO, X. Commercial liking and memory: moderating effects of product categories. *Journal of advertising research 41*, 3 (2001), 7–13.

[156] YUILL, N., AND MARTIN, A. F. Curling up with a good e-book: Mother-child shared story reading on screen or paper affects embodied interaction and warmth. *Frontiers in psychology 7* (2016), 1951.

Appendix A

Ethics for the Device Uses and Stereotypes Study

The letter of ethical approval for the device uses and stereotypes study.

Computer Science School Panel

School of Computer Science
2.32 Kilburn Building
0161-275-0143/5140

The University of Manchester

Manchester

M13 9PL

Email: ethics@cs.manchester.ac.uk

Ref: 2019-6746-11456

24/07/2019

Dear Miss Madeleine Steeds, Dr Sarah Clinch, Dr Caroline Jay

Study Title: Survey investigating device use and stereotypes

Computer Science School Panel

I write to thank you for submitting the final version of your documents for your project to the Committee on 22/07/2019 08:39 . I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation as submitted and approved by the Committee.

Please see below for a table of the titles, version numbers and dates of all the final approved documents for your project:

Document Type	File Name	Date	Version
Advertisement	advert_prolific	09/07/2019	1.1
Additional docs	Debrief v1.2	10/07/2019	1.2
Additional docs	Methodology_justification	10/07/2019	1
Additional docs	Word_list_multiple_choice task	10/07/2019	1
Data Management Plan	Steeds DMP 19-7-19	19/07/2019	2
Default	survey_questions	19/07/2019	2
Participant Information Sheet	Information Sheet v2	19/07/2019	2

This approval is effective for a period of five years and is on delegated authority of the University Research Ethics Committee (UREC) however please note that it is only valid for the specifications of the research project as outlined in the approved documentation set. If the project continues beyond the 5 year period or if you wish to propose any changes to the methodology or any other specifics within the project an application to seek an amendment must be submitted for review. Failure to do so could invalidate the insurance and constitute research misconduct.

You are reminded that, in accordance with University policy, any data carrying personal identifiers must be encrypted when not held on a secure university computer or kept securely as a hard copy in a location which is accessible only to those involved with the research.

For those undertaking research requiring a DBS Certificate: As you have now completed your ethical application if required a colleague at the University of Manchester will be in touch for you to undertake a DBS check. Please note that you do not have DBS approval until you have received a DBS Certificate completed by the University of Manchester, or you are an MA Teach First student who holds a DBS certificate for your current teaching role.

Reporting Requirements:

You are required to report to us the following:

1. Amendments: Guidance on what constitutes an amendment
2. Amendments: How to submit an amendment in the ERM system
3. Ethics Breaches and adverse events
4. Data breaches

We wish you every success with the research.

Yours sincerely,

Dr Markel Vigo

Computer Science School Panel

Appendix B

Survey for Device Uses and Stereotypes

The survey used to collect data in Chapter 4.

MANCHESTER
1824

The University of Manchester

Participant Information

Thank you for volunteering to take part in this study. Please read the following information and, if you are happy, complete the consent information at the bottom of the page.

You are being invited to take part in a research study to explore how adults living in the UK use different technological devices and their opinions of these devices. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. If there is anything that is not clear or if you would like more information, you can contact the primary researcher (details below). Take time to decide whether or not you wish to take part. The full information sheet can be downloaded **here**.

Please keep the information sheet so you can refer back to it at any point.

What would I be asked to do if I took part?

By participating in this study, you are agreeing to take part in an online survey which should take 25 minutes to complete. You will be asked to provide demographic information anonymously. If you do not wish to provide this information but still wish to participate, you may select the option 'prefer not to say'. You will further be asked to indicate which devices you use and how you use them and your opinions of these devices. Your reaction time to answer some questions will also be measured.

Your participation is voluntary and you are free to withdraw at any time by closing the survey. Once the survey has been submitted it will not be possible to withdraw data due to the anonymous nature of the study.

What do I do now?

If you have any questions about the study please contact the primary investigator at madeleine.steeds@postgrad.manchester.ac.uk. If you agree to take part then complete the consent form below to get started on the survey.

If you do not consent to the below you may close the survey and no data will be stored.

☐1) I confirm that I have read and understand the information above. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

☐2) I understand that my participation in the study is voluntary and that I am free to withdraw at any time by closing the survey, without giving any reason and without detriment to myself.

☐3) I understand that any information given by me may be used in future reports, academic articles, publications or presentations by the researcher

☐4) I understand that all data is anonymous and stored securely.

☐5) I understand that the data will be kept according to University guidelines for a minimum of 5 years.

☐6) I agree to take part in this research study.

I consent to take part in this study

MANCHESTER 1824

The University of Manchester

Part 1

Please answer the following questions about yourself

1) What is your age? Please give your answer in years (e.g. 31)

2) When you were born you were described as...

- Male
- Female
- Intersex
- Prefer not to say

3) What is your gender?

- Male
- Female
- Non-binary
- Prefer not to say

4) What is your ethnicity?

- Asian
- Black/African/Caribbean
- Mixed/Multiple Ethnicities
- White
- Other Ethnic Group
- Prefer not to say

5) What is your nationality?

6) Where do you currently live?

Please select

7) What is your yearly household income before tax? If you are a student please select your parent or guardian's household income.

- Less than £10,000
- £10,000 to £19,999
- £20,000 to £29,999
- £30,000 to £39,999
- £40,000 to £49,999
- £50,000 or above
- Prefer not to say

8) What is your native language?

Please select

9) What is your highest level of education?

- No academic qualification
- Left school at 16 (e.g. GCSE)
- Left school at 18 (e.g. A-Levels)
- Vocational qualification (e.g. technical college)
- Degree (e.g. BSc)
- Higher degree (e.g. MSc)
- Prefer not to say

10) What device are you completing this survey on? (e.g. laptop)

Please select

Continue

The University of Manchester

Part 2

Please answer the following questions about technological devices and how confident you are at using them.

1) **How many** of the following devices do you use **3 or more** times per week.

For example if you had 2 phones- one for work and one for home- and you used both more than three times a week, you would answer 2.

A desktop computer (PC)	2
A laptop computer	2
An iPad or tablet	2
A smartphone	2

Please describe the make and model of your **primary desktop computer**.

How would you describe the way that you use this device?.

Please describe the make and model of your **primary laptop**.

How would you describe the way that you use this device?.

Please describe the make and model of your **primary tablet**.

How would you describe the way that you use this device?.

Please describe the make and model of your **primary smartphone**.

How would you describe the way that you use this device?

2a) How confident do you feel using a desktop computer (PC)?

○ Very confident ○ Confident ○ Neither confident or unconfident ○ Unconfident ○ Very unconfident
○ I have never used this device before

2b) How confident do you feel using a smart phone?

○ Very confident ○ Confident ○ Neither confident or unconfident ○ Unconfident ○ Very unconfident
○ I have never used this device before

2c) How confident do you feel using a laptop?

○ Very confident ○ Confident ○ Neither confident or unconfident ○ Unconfident ○ Very unconfident
○ I have never used this device before

2d) How confident do you feel using a tablet/iPad?

- ○ Very confident
- ○ Confident
- ○ Neither confident or unconfident
- ○ Unconfident
- ○ Very unconfident
- ○ I have never used this device before

Continue

MANCHESTER
1824
The University of Manchester

Part 3

For the following device please list as many words as you can that you relate with that device.

For example, words relating to a clock might include time, hand, wall, hour, alarm, slow.

After each word, please click enter or press the enter key on your keyboard.

If the timer runs out you will be moved on to the next task.

Desktop computer

Enter

Continue

The University of Manchester

Part 3

For the following device please list as many words as you can that you relate with that device.

For example, words relating to a clock might include time, hand, wall, hour, alarm, slow.

After each word, please click enter or press the enter key on your keyboard.

If the timer runs out you will be moved on to the next task.

Smartphone

Enter

Continue

MANCHESTER
1824
The University of Manchester

Part 4

You will now be shown some words.

For each word, please select the device you most associate with it as fast as possible. If you do not associate the word with either device, please select the option "Neither".

Start

Presentation

Desktop Computer | Neither | Smartphone | Both

Continue

MANCHESTER 1824

The University of Manchester

Part 5

Please indicate how much you agree with the following statements.

Desktop Computers are ...

Awful Nice

Useful Useless

Valuable Worthless

Pleasant Unpleasant

Disagreeable Agreeable

Harmful Beneficial

Sad Happy

Foolish Wise

Continue

MANCHESTER 1824

The University of Manchester

Part 5

Please indicate how much you agree with the following statements.

Laptops are ...

Awful Nice

Useful Useless

Valuable Worthless

Pleasant Unpleasant

Disagreeable Agreeable

Harmful Beneficial

Sad Happy

Foolish Wise

Continue

The University of Manchester

Part 5

Please indicate how much you agree with the following statements.

Smartphones are ...

Awful Nice

Useful Useless

Valuable Worthless

Pleasant Unpleasant

Disagreeable Agreeable

Harmful Beneficial

Sad Happy

Foolish Wise

Continue

MANCHESTER 1824

The University of Manchester

Part 5

Please indicate how much you agree with the following statements.

Tablets are ...

Awful Nice

Useful Useless

Valuable Worthless

Pleasant Unpleasant

Disagreeable Agreeable

Harmful Beneficial

Sad Happy

Foolish Wise

Continue

The University of Manchester

Part 6

Please indicate the extent to which you agree with the following statements.

As viewed by society, how ... are people with desktop computers:

Sincere

Not at all Extremely

Competent

Not at all Extremely

Tolerant

Not at all Extremely

Confident

Not at all Extremely

Warm

Not at all Extremely

Independent

Not at all Extremely

Good-natured

Not at all Extremely

Competitive

Not at all Extremely

Intelligent

Not at all Extremely

Continue

The University of Manchester

Part 6

Please indicate the extent to which you agree with the following statements.

As viewed by society, how ... are people with Laptops:

Sincere

Not at all Extremely

Competent

Not at all Extremely

Tolerant

Not at all Extremely

Confident

Not at all Extremely

Warm

Not at all Extremely

Independent

Not at all Extremely

Good-natured

Not at all Extremely

Competitive

Not at all Extremely

Intelligent

Not at all Extremely

Continue

The University of Manchester

Part 6

Please indicate the extent to which you agree with the following statements.

As viewed by society, how ... are people with smartphones:

Sincere

Not at all Extremely

Competent

Not at all Extremely

Tolerant

Not at all Extremely

Confident

Not at all Extremely

Warm

Not at all Extremely

Independent

Not at all Extremely

Good-natured

Not at all Extremely

Competitive

Not at all Extremely

Intelligent

Not at all Extremely

Continue

The University of Manchester

Part 6

Please indicate the extent to which you agree with the following statements.

As viewed by society, how ... are people with tablets:

Sincere

Not at all Extremely

Competent

Not at all Extremely

Tolerant

Not at all Extremely

Confident

Not at all Extremely

Warm

Not at all Extremely

Independent

Not at all Extremely

Good-natured

Not at all Extremely

Competitive

Not at all Extremely

Intelligent

Not at all Extremely

Continue

The University of Manchester

Part 7

The following question relate to the amount of hours a day you spend doing specific activities on each device. Please give your answer in hours.

If you do not use the device, please leave the field blank.

On a **desktop computer**, how many **hours a week** do you spend:

On social media (e.g. Facebook, Instagram, Twitter)?

Working?

Gaming/watching television or films?

Emailing?

On a **laptop**, how many **hours a week** do you spend:

On social media (e.g. Facebook, Instagram, Twitter)?

Working?

Gaming/watching television or films?

Emailing

On a **smart phone**, how many **hours a week** do you spend:

On social media (e.g. Facebook, Instagram, Twitter)?

Working?

Gaming/watching television or films?

Emailing

On a **tablet**, how many **hours a week** do you spend:

On social media (e.g. Facebook, Instagram, Twitter)?

Working?

Gaming/watching television or films?

Emailing

Continue

MANCHESTER 1824

The University of Manchester

Part 8

Please indicate which activities you use a **desktop computer** for 3 or more times a week.

- Social Media ☐
- Messaging friends/family ☐
- Video Games ☐
- Online Shopping ☐
- Writing/drawing for pleasure ☐
- Emailing ☐
- Schedule Meetings ☐
- Reading work related documents ☐
- Note taking ☐
- Reading for pleasure (e.g. Novels) ☐
- Watching films, television, YouTube etc. ☐
- Listening to podcasts, music, audiobooks etc. ☐
- General web browsing ☐
- Online banking ☐
- Reading the news ☐
- Arranging travel (e.g. booking a train or taxi) ☐
- Journey planning or navigation (e.g. google maps) ☐
- Fitness or health (e.g. a Fitbit app) ☐
- Online dating ☐
- Taking or viewing photos ☐
- Checking the weather ☐
- Writing reports or essays ☐
- Analysing data or using spreadsheets ☐
- Making presentations ☐
- I do not use this device ☐

Continue

The University of Manchester

Part 8

Please indicate which activities you use a **laptop** for 3 or more times a week.

- Social Media ☐
- Messaging friends/family ☐
- Video Games ☐
- Online Shopping ☐
- Writing/drawing for pleasure ☐
- Emailing ☐
- Schedule Meetings ☐
- Reading work related documents ☐
- Note taking ☐
- Reading for pleasure (e.g. Novels) ☐
- Watching films, television, YouTube etc. ☐
- Listening to podcasts, music, audiobooks etc. ☐
- General web browsing ☐
- Online banking ☐
- Reading the news ☐
- Arranging travel (e.g. booking a train or taxi) ☐
- Journey planning or navigation (e.g. google maps) ☐
- Fitness or health (e.g. a Fitbit app) ☐
- Online dating ☐
- Taking or viewing photos ☐
- Checking the weather ☐
- Writing reports or essays ☐
- Analysing data or using spreadsheets ☐
- Making presentations ☐
- I do not use this device ☐

Continue

The University of Manchester

Part 8

Please indicate which activities you use a **smart phone** for 3 or more times a week.

- Social Media ☐
- Messaging friends/family ☐
- Video Games ☐
- Online Shopping ☐
- Writing/drawing for pleasure ☐
- Emailing ☐
- Schedule Meetings ☐
- Reading work related documents ☐
- Note taking ☐
- Reading for pleasure (e.g. Novels) ☐
- Watching films, television, YouTube etc. ☐
- Listening to podcasts, music, audiobooks etc. ☐
- General web browsing ☐
- Online banking ☐
- Reading the news ☐
- Arranging travel (e.g. booking a train or taxi) ☐
- Journey planning or navigation (e.g. google maps) ☐
- Fitness or health (e.g. a Fitbit app) ☐
- Online dating ☐
- Taking or viewing photos ☐
- Checking the weather ☐
- Writing reports or essays ☐
- Analysing data or using spreadsheets ☐
- Making presentations ☐
- I do not use this device ☐

Continue

The University of Manchester

Part 8

Please indicate which activities you use a **tablet** for 3 or more times a week.

- Social Media ☐
- Messaging friends/family ☐
- Video Games ☐
- Online Shopping ☐
- Writing/drawing for pleasure ☐
- Emailing ☐
- Schedule Meetings ☐
- Reading work related documents ☐
- Note taking ☐
- Reading for pleasure (e.g. Novels) ☐
- Watching films, television, YouTube etc. ☐
- Listening to podcasts, music, audiobooks etc. ☐
- General web browsing ☐
- Online banking ☐
- Reading the news ☐
- Arranging travel (e.g. booking a train or taxi) ☐
- Journey planning or navigation (e.g. google maps) ☐
- Fitness or health (e.g. a Fitbit app) ☐
- Online dating ☐
- Taking or viewing photos ☐
- Checking the weather ☐
- Writing reports or essays ☐
- Analysing data or using spreadsheets ☐
- Making presentations ☐
- I do not use this device ☐

Continue

MANCHESTER 1824

The University of Manchester

Thank you!

Thank you for taking part in this study!

The present study was investigating how individuals use their devices and their opinions of them.

The study also investigated the words people associate with their devices and how strong those associations are.

We hypothesise that if people have a strong perception of stereotypes, this will affect their opinions of them and this will be related to the extent to which they use a device for either work or leisure.

For the full information on the study please download the debriefing sheet **here**

Thank you again for participating in the study!

Continue

Appendix C

Information Sheet for Device Use and Stereotypes

The participant information sheet for the device use and stereotypes study.

MANCHESTER
1824
The University of Manchester

Research Participant Information Sheet

Device uses and opinions

Participant Information Sheet (PIS)

You are being invited to take part in a research study to explore how people use different technological devices and their opinions of them. This is a research project as part of PhD being undertaken at the School of Computer Science at the University of Manchester. Before you decide whether to take part, it is important for you to understand why the research is being conducted and what it will involve. Please take time to read the following information carefully before deciding whether to take part and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information. Thank you for taking the time to read this.

About the research

- **Who will conduct the research?**

Primary researcher
Name: Madeleine Steeds
Address: School of Computer Science, University of Manchester, Oxford Rd, Manchester. M13 9PL.
Email: madeleine.steeds@postgrad.manchester.ac.uk

Supervisors
Name: Dr. Sarah Clinch *and Dr.* Caroline Jay
Address: School of Computer Science, University of Manchester, Oxford Rd, Manchester. M13 9PL.
Email: sarah.clinch@manchester.ac.uk
caroline.jay@manchester.ac.uk

- **What is the purpose of the research?**

Different technological devices are becoming very widespread. This study aims to investigate what devices adults living in the UK use and how they use them. You have been chosen because you responded to an advert placed on prolific.ac. The study aims to recruit a total of 200 participants.

- **Will the outcomes of the research be published?**

The outcomes of the research will be published as part of a student thesis. They may also be presented at academic conferences and published in a peer reviewed journal.

- **Who has reviewed the research project?**

The project has been reviewed by The University of Manchester School of Computer Science Ethics Committee.

What would my involvement be?

- **What would I be asked to do if I took part?**

By participating in this study, you are agreeing to take part in an online survey. You will be asked to provide demographic information anonymously. You will further be asked to indicate which devices you use and how you use them and your opinions of these devices. Associations you make between devices and other words will be investigated. Your reaction time to answer some questions will also be measured. The study will take 25 minutes to complete.

- **Will I be compensated for taking part?**

You will be compensated £2.88 for participating, which will be paid to you through prolific.ac.

- **What happens if I do not want to take part or if I change my mind?**

It is up to you to decide whether or not to take part. If you do decide to take part, please keep this information sheet and be asked to indicate your consent on the online survey. If you decide to take part you are still free to withdraw at any time by closing the survey without giving a reason and without detriment to yourself. However, it will not be possible to remove your data from the project once you submit the survey. This does not affect your data protection rights. If you decide not to take part you do not need to do anything further.

Data Protection and Confidentiality

- **What information will you collect about me?**

In order to participate in this research project we will need to collect information that could identify you, called “personal identifiable information”. Specifically we will need to collect your:

- age
- gender
- sex
- ethnicity
- income
- country of residence
- native language
- education
- nationality

If you do not wish to provide this information but still wish to participate, you may select the option 'prefer not to say' or leave the field blank.

- **Under what legal basis are you collecting this information?**

We are collecting and storing this personal identifiable information in accordance with data protection law which protect your rights. These state that we must have a legal basis (specific reason) for collecting your data. For this study, the specific reason is that it is “a public interest task” and “a process necessary for research purposes”.

- **What are my rights in relation to the information you will collect about me?**

You have a number of rights under data protection law regarding your personal information. For example you can request a copy of the information we hold about you.

If you would like to know more about your different rights or the way we use your personal information to ensure we follow the law, please consult our Privacy Notice for Research.

- **Will my participation in the study be confidential and my personal identifiable information be protected?**

In accordance with data protection law, The University of Manchester is the Data Controller for this project. This means that we are responsible for making sure your personal information is kept secure, confidential and used only in the way you have been told it will be used. All researchers are trained with this in mind, and your data will be looked after in the following way:

- Your data will be anonymous and at no point will your name be attached to the data
- The data will be stored on an encrypted server and will be backed up to a secure store
- The data may be shared for research purposes but you will not be identifiable from the data
- Shared data may be then used in future studies
- The data will be stored for a minimum of five years

Please also note that individuals from The University of Manchester or regulatory authorities may need to look at the data collected for this study to make sure the project is being carried out as planned. This may involve looking at identifiable data. All individuals involved in auditing and monitoring the study will have a strict duty of confidentiality to you as a research participant.

What if I have a complaint?

- **Contact details for complaints**

If you have a complaint that you wish to direct to members of the research team, please contact:

***Primary Researcher*:**
Madeleine Steeds
School of Computer Science
University of Manchester
Oxford Rd
Manchester
M13 9PL.
madeleine.steeds@postgrad.manchester.ac.uk

***Supervisor*:**
Dr Sarah Clinch
School of Computer Science,
University of Manchester
Oxford Rd
Manchester
M13 9PL
sarah.clinch@manchester.ac.uk
0161 275 7190

If you wish to make a formal complaint to someone independent of the research team or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact

The Research Governance and Integrity Officer, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: research.complaints@manchester.ac.uk or by telephoning 0161 275 2674.

If you wish to contact us about your data protection rights, please email dataprotection@manchester.ac.uk or write to The Information Governance Office, Christie Building, The University of Manchester, Oxford Road, M13 9PL at the University and we will guide you through the process of exercising your rights.

You also have a right to complain to the Information Commissioner's Office about complaints relating to your personal identifiable information Tel 0303 123 1113

Contact Details

If you have any queries about the study please contact:

Madeleine Steeds

School of Computer Science, University of Manchester, Oxford Rd, Manchester, M13 9PL.

madeleine.steeds@postgrad.manchester.ac.uk

If you would like to participate, please proceed to the online survey.

Appendix D

Debriefing Sheet for Device Use and Stereotypes

The debriefing sheet for the device use and stereotypes study.

The University of Manchester

Device Uses Questionnaire

Participant Debrief Sheet

Thank you for participating in this study. The present study was investigating how individuals use their devices and their opinions of them. The study also investigated the words people associate with their devices and how strong those associations are. It was thought that strong perceptions of stereotypes, will affect opinions of devices. This would also relate to the extent to which a device is used for work or leisure.

To see if this premise is true, you were asked to complete a questionnaire asking what activities you typically use a laptop or mobile for. The free association task aimed to investigate if there are common trends in words people associate with their devices, and if some devices are perceived differently to others.

If you would like further information about the study, please feel free to contact the primary investigator (below). We hope that you have found it interesting and have not been upset by any of the study. However, if you have found any part of this experience to be distressing and you wish to speak to one of the researchers, please contact the primary investigator or their supervisors:

Primary Investigator:
Madeleine Steeds
Email: madeleine.steeds@postgrad.manchester.ac.uk
Address: School of Computer Science,
University of Manchester,
Oxford Rd,
Manchester. M13 9PL.

Supervisor
Dr. Sarah Clinch
Email: sarah.clinch@manchester.ac.uk

Supervisor
Dr. Caroline Jay
Email: caroline.jay@manchester.ac.uk

This Project Has Been Approved by the University of Manchester's Research Ethics Committee [UREC reference number].

Appendix E

Ethics for the Information Processing Study

The letter of ethical approval for the information processing study.

Computer Science Department Panel

Department of Computer Science
2.32 Kilburn Building
0161-275-0143/5140

The University of Manchester

Manchester

M13 9PL

Email: ethics@cs.manchester.ac.uk

Ref: 2020-10117-16292

14/08/2020

Dear Miss Madeleine Steeds, Dr Caroline Jay, Dr Sarah Clinch

Study Title: Impact of stereotypes on information processing

Computer Science Department Panel

I write to thank you for submitting the final version of your documents for your project to the Committee on 10/08/2020 10:57 . I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation as submitted and approved by the Committee.

Note that there is still some ambiguity about the use of IDs. The DMP says that "The only personal data that will be kept beyond the end of the project is the participant's age, gender and occupation. This is stored under a participant ID code at the point of collection." while the PIS suggests that "it will not be possible to remove your data from the project a week after your participation, as it will no longer be associated with your unique participant ID". I'm we are talking about different IDs here that'd be fine. If not, these statements are contradictory.

COVID-19 Important Note

Please ensure you read the information on the Research Ethics website in relation to data collection in the COVID environment as well as the guidance issued by the University in relation to face-to-face (in person) data collection both on and off campus.

A word document version of this guidance is also available.

Please see below for a table of the titles, version numbers and dates of all the final approved documents for your project:

Document Type	File Name	Date	Version
Additional docs	Debrief	16/07/2020	1.1
Default	DesktopStudy	20/07/2020	1
Default	PhoneStudy	20/07/2020	1
Additional docs	StimuliList	20/07/2020	1.1
Additional docs	StimuliPresentationDesktop	20/07/2020	1
Additional docs	StimuliPresentationPhone	20/07/2020	1
Advertisement	ExampleTweet	21/07/2020	1
Advertisement	advertHorizontal	21/07/2020	1.1
Data Management Plan	DataManagementPlan	10/08/2020	2
Participant Information Sheet	InformationSheet	10/08/2020	2.1

This approval is effective for a period of five years and is on delegated authority of the University Research Ethics Committee (UREC) however please note that it is only valid for the specifications of the research project as outlined in the approved documentation set. If the project continues beyond the 5 year period or if you wish to propose any changes to the methodology or any other specifics within the project an application to seek an amendment must be submitted for review. Failure to do so could invalidate the insurance and constitute research misconduct.

You are reminded that, in accordance with University policy, any data carrying personal identifiers must be encrypted when not held on a secure university computer or kept securely as a hard copy in a location which is accessible only to those involved with the research.

For those undertaking research requiring a DBS Certificate: As you have now completed your ethical application if required a colleague at the University of Manchester will be in touch for you to undertake a DBS check. Please note that you do not have DBS approval until you have received a DBS Certificate completed by the University of Manchester, or you are an MA Teach First student who holds a DBS certificate for your current teaching role.

Reporting Requirements:

You are required to report to us the following:

1. Amendments: Guidance on what constitutes an amendment
2. Amendments: How to submit an amendment in the ERM system
3. Ethics Breaches and adverse events
4. Data breaches

We wish you every success with the research.

Yours sincerely,

Dr Markel Vigo

Computer Science Department Panel

Appendix F

Stimuli

The experimental stimuli used in Chapters 5 and 6 are presented here. Please note that only the desktop and smartphone stimuli were used in Chapter 6.

Desktop	Smartphone	Hedonic/Utilitarian	SCM
big	apps	awful	sincere
work	small	nice	competitive
slow	alarm	useful	competent
bulky	camera	useless	tolerant
heavy	communication	valuable	warm
powerful	text	worthless	independent
desk	touchscreen	pleasant	good-natured
old	calls	unpleasant	intelligent
large	photos	disagreeable	confident
tower	mobile	agreeable	
software	charger	sad	
word	portable	happy	

Appendix G

Information Sheet for the Information Processing Study

The information sheet for the information processing study.

MANCHESTER 1824
The University of Manchester

Research Participant Information Sheet

The effect of device on information processing speed

Participant Information Sheet (PIS)

You are being invited to take part in a research study as part of a PhD project investigating the speed of information processing on different devices. Before you decide whether to take part, it is important for you to understand why the research is being conducted and what it will involve. Please take time to read the following information carefully before deciding whether to take part and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information. Thank you for taking the time to read this.

About the research

- **Who will conduct the research?**

Primary Investigator:

Name: Madeleine Steeds

Address: Department of Computer Science
University of Manchester, Oxford Road
Manchester, M13 9PL

Email: madeleine.steeds@manchester.ac.uk

- **What is the purpose of the research?**

The aim of this research is to investigate the effect of different device on information processing.

- **Will the outcomes of the research be published?**

The primary outcome will be the data collected and analysed to investigate the effect of devices on how quickly information is processed. The anonymised data may be published in academic journals or conferences as well as in the primary investigator's thesis.

- **Who has reviewed the research project?**

This project has been reviewed and given ethical approval by the Department of Computer Science Ethics Committee.

What would my involvement be?

- **What would I be asked to do if I took part?**

To participate in this study you must be aged 18 or over, a fluent English speaker, and not be colour-blind. The study may be completed on a smartphone, laptop or desktop computer.

You will be asked to participate in the research from your own home in a quiet environment. You will be asked to provide demographic information before participating in 10 practice trials of a Stroop test. A stroop test involves looking at words and indicating the colour of the font the words are printed in. After the practice trials you will be asked to complete a longer Stroop test with different words. The study should take no longer than 15 minutes to complete.

- **Will I be compensated for taking part?**

You will not be compensated for your participation.

- **What happens if I do not want to take part or if I change my mind?**

It is up to you to decide whether or not to take part. If you would like to participate please continue to the study. If you do decide to take part you will be given this information sheet to keep and will be asked to complete a checkbox consent form. If you decide to take part you are still free to withdraw up to a week after your participation without giving a reason and without detriment to yourself. However, it will not be possible to remove your data from the project a week after your participation, as it will no longer be associated with your unique participant ID and we will not be able to identify your specific data. This does not affect your data protection rights. If you decide not to take part you do not need to do anything further.

Data Protection and Confidentiality

- **What information will you collect about me?**

In order to participate in this research project we will need to collect information that could identify you, called "personal identifiable information". Specifically we will need to collect:

- Age
- Gender
- Occupation

If you would not like to provide your gender and/or occupation, you may choose "prefer not to say" whilst still participating in the research.

- **Under what legal basis are you collecting this information?**

We are collecting and storing this personal identifiable information in accordance with data protection law which protect your rights. These state that we must have a legal basis (specific reason) for collecting your data. For this study, the specific reason is that it is "a public interest task" and "a process necessary for research purposes".

- **What are my rights in relation to the information you will collect about me?**

You have a number of rights under data protection law regarding your personal information. For example you can request a copy of the information we hold about you.

If you would like to know more about your different rights or the way we use your personal information to ensure we follow the law, please consult our Privacy Notice for Research.

- **Will my participation in the study be confidential and my personal identifiable information be protected?**

In accordance with data protection law, The University of Manchester is the Data Controller for this project. This means that we are responsible for making sure your personal information is kept secure, confidential and used only in the way you have been told it will be used. All researchers are trained with this in mind, and your data will be looked after in the following way:

- Data will be saved throughout the study under a unique ID which will not be able to be linked back to the participant.
- Data will be stored on an encrypted server and will be backed up to a secure store.
- Experimental data will be stored for a minimum of 5 years.

Please also note that individuals from The University of Manchester or regulatory authorities may need to look at the data collected for this study to make sure the project is being carried out as planned. This may involve looking at identifiable data. All individuals involved in auditing and monitoring the study will have a strict duty of confidentiality to you as a research participant.

What if I have a complaint?

If you have a complaint that you wish to direct to members of the research team, please contact:

Dr. Sarah Clinch
Department of Computer Science, University of Manchester
Oxford Road, Manchester, M13 9PL
sarah.clinch@manchester.ac.uk

If you wish to make a formal complaint to someone independent of the research team or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact

The Research Governance and Integrity Officer, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: research.complaints@manchester.ac.uk or by telephoning 0161 275 2674.

If you wish to contact us about your data protection rights, please email dataprotection@manchester.ac.uk or write to The Information Governance Office, Christie Building, The University of Manchester, Oxford Road, M13 9PL at the University and we will guide you through the process of exercising your rights.

You also have a right to complain to the Information Commissioner's Office about complaints relating to your personal identifiable information Tel 0303 123 1113.

Contact Details

If you have any queries about the study or if you are interested in taking part then please contact If you have any queries about the study or if you are interested in taking part then please contact:

Madeleine Steeds

Department of Computer Science, University of Manchester
Oxford Road, Manchester, M13 9PL
madeleine.steeds@manchester.ac.uk

Appendix H

Debriefing Sheet for the Information Processing Study

The debriefing sheet for the information processing study.

MANCHESTER 1824
The University of Manchester

The Effect of Device on Information Processing

Participant Debrief Sheet

Thank you for participating in this study. The present study was investigating whether performance in a Stroop task was impacted when the words presented are related to the device being used. A Stroop task requires people to indicate the colour a word is printed in as fast as possible. Previous research has indicated that when the written word matches expectations. As such it is anticipated that response times will be faster if the word presented matches the device being used.

If you would like further information about the study, please feel free to contact the primary investigator (below). We hope that you have found it interesting and have not been upset by any of the study. However, if you have found any part of this experience to be distressing and you wish to speak to one of the researchers, please contact the primary investigator or their supervisors:

Primary Investigator:
Madeleine Steeds
Email: madeleine.steeds@manchester.ac.uk
Address: Department of Computer Science,
University of Manchester,
Oxford Rd,
Manchester, M13 9PL.

Supervisor
Dr. Sarah Clinch
Email: sarah.clinch@manchester.ac.uk

Supervisor
Dr. Caroline Jay
Email: caroline.jay@manchester.ac.uk

This Project Has Been Approved by the University of Manchester's Research Ethics Committee [2020-10117-1629].

Appendix I

Ethics for the Memory Study

The letter of ethical approval for the memory study.

Computer Science Department Panel

Department of Computer Science
2.32 Kilburn Building
0161-275-0143/5140

The University of Manchester

Manchester

M13 9PL

Email: ethics@cs.manchester.ac.uk

Ref: 2021-11329-18169

15/03/2021

Dear Miss Madeleine Steeds, , Dr Caroline Jay, Dr Sarah Clinch

Study Title: The effect of device on memory- online study

Computer Science Department Panel

I write to thank you for submitting the final version of your documents for your project to the Committee on 11/03/2021 13:58 . I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation as submitted and approved by the Committee.

COVID-19 Important Note

Please ensure you read the information on the Research Ethics website in relation to data collection in the COVID environment as well as the guidance issued by the University in relation to face-to-face (in person) data collection both on and off campus.

A word document version of this guidance is also available.

Please see below for a table of the titles, version numbers and dates of all the final approved documents for your project:

Document Type	File Name	Date	Version
Data Management Plan	DMP_v1	15/02/2021	1
Additional docs	Debrief_1.1	15/02/2021	1.1
Advertisement	exampleTweet	18/02/2021	1
Advertisement	emailTemplate	18/02/2021	1
Additional docs	WordList	23/02/2021	1
Advertisement	RedditAdvertText	25/02/2021	1
Advertisement	memoryAdvert	25/02/2021	1
Default	Full_study	11/03/2021	2
Participant Information Sheet	PIS_2.1	11/03/2021	2.1
Additional docs	PIS_2.1 Highlighted changes	11/03/2021	2.1
Additional docs	LetterReply	11/03/2021	1

This approval is effective for a period of five years and is on delegated authority of the University Research Ethics Committee (UREC) however please note that it is only valid for the specifications of the research project as outlined in the approved documentation set. If the project continues beyond the 5 year period or if you wish to propose any changes to the methodology or any other specifics within the project an application to seek an amendment must be submitted for review. Failure to do so could invalidate the insurance and constitute research misconduct.

You are reminded that, in accordance with University policy, any data carrying personal identifiers must be encrypted when not held on a secure university computer or kept securely as a hard copy in a location which is accessible only to those involved with the research.

For those undertaking research requiring a DBS Certificate: As you have now completed your ethical application if required a colleague at the University of Manchester will be in touch for you to undertake a DBS check. Please note that you do not have DBS approval until you have received a DBS Certificate completed by the University of Manchester, or you are an MA Teach First student who holds a DBS certificate for your current teaching role.

Reporting Requirements:

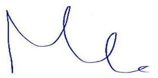
You are required to report to us the following:

1. Amendments: Guidance on what constitutes an amendment
2. Amendments: How to submit an amendment in the ERM system

3. Ethics Breaches and adverse events
4. Data breaches

We wish you every success with the research.

Yours sincerely,



Dr Markel Vigo

Computer Science Department Panel

Appendix J

Information Sheet for the Memory Study

The information sheet for the information processing study.

The Effect of Device on Memory

You are being invited to take part in a research study looking at the effect of device on memory, as part of a PhD project. Before you decide whether to take part, it's important for you to understand why the research is being conducted and what it will involve. Please take time to read the following information carefully before deciding whether to take part and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information.

Who is carrying out this research study?

Madeleine Steeds, Dr Sarah Clinch, and Dr Caroline Jay, Department of Computer Science, The University of Manchester.

What is the purpose of the study?

This research aims to investigate whether using different devices impacts memory performance.

Am I suitable to take part?

To participate you must:

- Be at least 18 years or older
- Have normal or corrected to normal vision
- Be an android phone user

What will I be asked to do if I take part?

You will be asked to come to the Kilburn Building, Oxford Road, to participate in one 20 minute session. You will be shown a series of words and be asked to memorise as many as you can before participating in a recall and recognition test. You will also be asked some mental maths questions.

Will I be compensated for taking part?

You will not be compensated for your participation.

What are the risks if I take part?

There are no risks anticipated with your participation. COVID-19 safety measures will be taken to ensure your safety while participating in the experiment, including sanitizing all surfaces and equipment between participants. The experimenter will also maintain a social distance and wear a mask throughout the procedure.

Who has reviewed this study?

This study has been reviewed by the Department of Computer Science, University of Manchester Research Ethics Committee.

Who is funding this study?

This research is supported by the UK EPSRC under grant number EP/R513131/1 (studentship 2169193).

What happens if I don't want to take part or change my mind?

It is up to you to decide whether or not to take part. If you would like to, then please follow the sign up link or contact Madeleine Steeds (contact details at the end of this sheet). If you do decide to take part, you will be given this information sheet to keep and will be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time without giving a reason and without detriment to yourself. If you would like for us to delete the data that we have collected up until that point, we will do so. This does not affect your data protection rights. If you do decide not to take part, you do not need to

Data Protection and Confidentiality

What information will you collect about me?

In order to participate in this research project we will need to collect information that could identify you, called "personal identifiable information". Specifically we will need to collect:

- Age
- Gender (optional)
- Occupation (optional)

Under what legal basis are you collecting this information?

We are collecting and storing this personal identifiable information in accordance with UK data protection law which protect your rights. These state that we must have a legal basis (specific reason) for collecting your data. For this study, the specific reason is that it is "a public interest task" and "a process necessary for research purposes".

What are my rights in relation to the information you will collect about me?

You have a number of rights under data protection law regarding your personal information. For example you can request a copy of the information we hold about you.

If you would like to know more about your different rights or the way we use your personal information to ensure we follow the law, please consult our Privacy Notice for Research (https://documents.manchester.ac.uk/display.aspx?DocID=37095).

Will my participation in the study be confidential and my personal identifiable information be protected?

In accordance with data protection law, The University of Manchester is the Data Controller for this project. This means that we are responsible for making sure your personal information is kept secure, confidential and used only in the way you have been told it will be used. All researchers are trained with this in mind, and your data will be looked after in the following way:

- Consent forms will be kept in a locked filing cabinet, separate to the data collected in the study
- Data will only be stored under your Participant ID number, and will be fully anonymised.
- Electronic data will be collected on a secure internal server and backed up to the University of Manchester research data storage (RDS). After the data has been moved to the RDS, it will be removed from the server.
- The anonymous dataset will be made available for other researchers on Figshare.
- Data will be stored for 5-10 years from the end of the study.

Please also note that individuals from The University of Manchester or regulatory authorities may need to look at the data collected for this study to make sure the project is being carried out as planned. This may involve looking at identifiable data. All individuals involved in auditing and monitoring the study will have a strict duty of confidentiality to you as a research participant.

What if I want to make a complaint?

If you have a complaint that you wish to direct to members of the research team, please contact Madeleine Steeds (madeleine.steeds@manchester.ac.uk) or their supervisors Sarah Clinch and Caroline Jay (sarah.clinch@manchester.ac.uk; caroline.jay@manchester.ac.uk).

If you wish to make a formal complaint to someone independent of the research team or if you are not satisfied with the response you have gained from the researcher in the first instance then please contact:

The Research Ethics Manager, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: research.complaints@manchester.ac.uk or by telephoning 0161 306 8089.

If you wish to contact us about your data protection rights, please email dataprotection@manchester.ac.uk or write to The Information Governance Office, Christie Building, The University of Manchester, Oxford Road, M13 9PL at the University and we will guide you through the process of exercising your rights.

You also have a right to complain to the Information Commissioner's Office (https://ico.org.uk/make-a-complaint/) about complaints relating to your personal identifiable information Tel 0303 123 1113

What do I do now?

If you have any questions or would like to take part please contact:

Madeleine Steeds

Email: madeleine.steeds@manchester.ac.uk

Appendix K

Debriefing Sheet for the Memory Study

The debriefing sheet for the information processing study.

The University of Manchester

Recognition and Recall Memory

Participant Debrief Sheet

Thank you for participating in this study. The present study was investigating whether using a different device when memorising words compared to when remembering words, would negatively impact memory. We were further investigating whether words associated with a device (for example a “desktop computer” is related to the word “keyboard”) would lead to improved memory performance when using that device.

To see if this premise is true, you were asked to memorise a list of words, some of which had been previously associated with a device, and then recall and recognise those words on either the same device or a different device.

If you would like further information about the study, please feel free to contact the primary investigator (below). We hope that you have found it interesting and have not been upset by any of the study. However, if you have found any part of this experience to be distressing and you wish to speak to one of the researchers, please contact the primary investigator or their supervisors:

Primary Investigator:
Madeleine Steeds
Email: madeleine.steeds@manchester.ac.uk
Address: Department of Computer Science,
University of Manchester,
Oxford Rd,
Manchester, M13 9PL.

Supervisor
Dr. Sarah Clinch
Email: sarah.clinch@manchester.ac.uk

Supervisor
Dr. Caroline Jay
Email: caroline.jay@manchester.ac.uk

This Project Has Been Approved by the University of Manchester’s Research Ethics Committee [2021-12882-20613].