The Use of Personalisation within Manufacturing Environments

Joshua Duvnjak

Faculty of Engineering, Horizon Centre for Doctoral Training

University of Nottingham

Abstract

With increasing emphasis on Industry 4.0 and 5.0 paradigms, there is interest in using data-driven technologies for improvements in the manufacturing sector. Personalisation is one such technology. Using personal data, this technology can change the appearance of a user interface or the action required by a user. This can make computer systems easier to use or better allocation of resources. For manufacturing industries, personalisation could be used to divert complex tasks to advanced users while allowing beginner users to learn on simpler operations. However, integrating data-driven technologies may produce unforeseen negative effects. There is limited current knowledge of how personalisation works in the current manufacturing industry and how this may impact users. This thesis studies the potential effect of personalisation when implemented in manufacturing scenarios.

The overarching goal of the thesis is to provide a set of design insights for personalisation's usage in the manufacturing industry. An early collaborative work into Digital Manufacturing Technologies revealed the concerns around personal data capture and usage. This became a key motivator of the thesis. With personal data being a major component of personalisation, how people feel about data and therefore, the technology could be a potential barrier to its use in manufacturing settings. The current project's aims focused on how people would design their own personalisation systems, understanding stakeholders' acceptance of personalisation systems and how personalisation can affect task completion.

To study a diverse range of topics regarding personalisation a mixed methods approach was utilised. Towards the start of the project, an examination of personalisation systems revealed a range of different types. Thus, a taxonomy was defined, which became the underpinning for how the thesis treats personalisation. The proposed taxonomy categorises personalisation systems into five types whose name implies their function: Suggesters, Rearrangers, Swappers, Generators and Controllers. These categories became a type of card featured within constructed Personalisation Design Cards (PDCs). The PDCs were utilised to understand how users would design their own personalisation systems in a participatory design study. Personalisation's acceptance from stakeholders was examined using a survey, with experimental conditions to represent data types used within the technology and different personalisation types. The last empirical chapter featured multiple types of automated assistance within a quality control inspection task. The impact that the automation had on the participants and their feelings towards the assistance were captured. The types of personalisation system were subject to an

additional systematic review at a later stage to provide further confirmation of their effectiveness in categorising personalisation systems.

The empirical studies each contributed to the understanding of personalisation. The participants used PDCs were used to craft personalisation systems, this experience led to them describing similar worries about data from the motivating work. It appeared that the participants had a preference for systems that are “dynamic” and utilised less intrusive forms of data. The survey indicated that stakeholders preferred Swapper systems when the utilised personal data was heart rate but did not have a preference when performance data was used. Personalisation implementation within a Quality Control Inspection task demonstrated that users viewed the Controllers system (which “completed” the next action) as judgemental; in contrast, the Suggesters system (which suggested the next action) was recognised as having good usability. These insights were collated to form a set of design insights for industry professionals.

The thesis stands as an early look into personalisation’s usage within the manufacturing sector. It provides a resource for systems integrators which describes the theoretical underpinning of how this technology can be implemented. This enables potential integrators to select the type of personalisation that may be useful to them while simultaneously providing a baseline level of user requirements regarding how users may react to this technology. Future work can branch off the included work by specialising in one aspect of personalisation, such as algorithms or design.

Acknowledgements

A first acknowledgement must go to Robert Houghton, Ayse Kucukyilmaz and Sarah Sharples for their supervision over the PhD project.

The Horizon CDT deserves thanks for the support offered throughout the PhD programme. With another shoutout to the 2020 cohort; it was great to be part of such an awesome group.

Another thanks must go to Mixed Reality Lab and to all the support and friendship offered by this group. The HFRG receives a thanks for providing the resources to present my research over the period of the PhD project.

Thank you to the assessors of the PhD project over the years. It aided in keeping the project on track and greatly improved the final outputs.

A word of gratitude for Dr Roger Whitham for providing materials used in the theoretical underpinning of the codesign workshop study. Dr David R. Large for aid with the future paper publications.

Thanks to all my friends and family who have provided support throughout this process.

Publications

Duvnjak, J., Kucukyilmaz, A., & Houghton, R. (2024). *Understanding user needs of personalisation-based automated systems with development and application of novel ideation cards*. 15th International Conference on Applied Human Factors and Ergonomics (AHFE 2024). https://doi.org/10.54941/ahfe1005032

Marinescu, A., Argyle, E. M., Duvnjak, J., Wilson, M. L., Lawson, G., Sharples, S., Hubbard, E., & Justham, L. (2022). The future of manufacturing: Utopia or dystopia? *Human Factors and Ergonomics in Manufacturing & Service Industries*, hfm.20976. https://doi.org/10.1002/hfm.20976

Marinescu, A., Argyle, E. M., Duvnjak, J., Wilson, M., Sharples, S., Lawson, G., Hubbard, E. M., & Justham, L. (2021). *Using the Contravision Approach to Elicit User Reactions to Future Manufacturing Technologies*. https://doi.org/10.6084/m9.figshare.14915277.v1

Notes

This work was supported by the Engineering and Physical Research Council [Grant number EP/S023305/1] and by Digital Toolkit for optimisation of operators and technology in manufacturing partnerships (DigiTOP) [EP/R032718/1] / Made Smarter Innovation Centre for People-led Digitalisation [EP/V062042/1].

The thesis was formatted to match (where possible) the APA 7th Style:

> American Psychological Association (Washington, District of Columbia) (Ed.). (2020). *Publication manual of the American psychological association* (Seventh edition). American Psychological Association.

The motivating work featured in the introduction (Chapter 1: Introduction) has been published prior (Marinescu et al., 2022) and the included work is an expanded form of this work. The Taxonomy of Personalisation systems (Chapter 2: Taxonomy of Personalisation Systems) was attempted to be published prior and thus some sections had significant editing from the co-authors (the two leading PhD supervisors on this project). Since this point, it has received a significant update to its content. It has since be submitted to journal and had input from David R. Large (and supervisor Sarah Sharples). The codesign chapter is the expanded version of the published paper (Chapter 3: Codesign with the Personalisation Design Cards) (Duvnjak et al., 2024), which was contributed to by two of the projects supervisors.

Contents

Figures

Tables

Chapter 1: Introduction

1.1 Towards Industry 4.0/5.0

Technology is constantly being replaced by more powerful and efficient variants (Roser et al., 2023). A set of modern technologies which have the potential to be used in manufacturing have been grouped together as part of the “Industry 4.0” concept (Lasi et al., 2014). The use of Industry 4.0 represents a paradigm shift for manufacturing (Xu et al., 2021) and is quickly becoming mainstream. It has been reported that companies such as DHL, Volkswagen and Bosch have all embraced new technological approaches (AMFG, 2019). Academic research projects are involved with Industry 4.0, such as the external partners for this project: DigiTOP[1] and Made Smarter Centre for People Led Digitisation[2]. These groups aim to provide information and assistance to various groups looking to implement Digital Manufacturing Technologies. As more businesses take interest in Industry 4.0, it is vital that current work continues the path of existing research projects by focusing on how technologies can be utilised.

A sophisticated technological approach can bring benefits to organisations through Industry 4.0. This is best explained by the technologic approaches that are incorporated into Industry 4.0 (e.g. Nahavandi, 2019, Table 1). Nahavandi (2019) states that one of these technologies is Digital Twins. This concept aims to replicate a physical system in a digital form. For example, a physical machine could have each of the inputs and outputs monitored alongside general characteristics of the machine, such as temperature. This type of monitoring can create a “digital twin” which allows for remote maintenance and data analysis to be performed on the machine with the aim to identify potential problems before they occur or improve performance. This is a general theme of Industry 4.0, in which the benefit occurs when existing processes are linked with digital approaches. Further, the building of these technologies can be seen when comparing Industry 4.0 and upcoming Industry 5.0, which shows a

[1] https://digitop.ac.uk/
[2] https://www.madesmarter.uk/adoption/

natural progression of the paradigms over time (Table 1). The need for an Industry 5.0 approach is to promote the user to the front and centre of technology implementation (lacking in industry 4.0 (Xu et al., 2021)).

Table 1 – Comparison of Industry Paradigms

A representation of how elements included within the Industry 5.0 paradigm have been discussed as part of prior "industry" paradigms. Computer numerical control refers to using computers to control physical machining tools. Serus is a modular approach to machining using smaller teams which can be changed when required (Yin et al., 2018). *Artificial Intelligence refers to the use of data to provide insights into a topic using computer analysis. 'Green energy comes from renewable sources'* (Li, 2024). *Systems Collation describes the need for information to be widely available across devices* (Nahavandi, 2019).

Paradigm	Industry 3.0	Industry 4.0	Industry 5.0
Technologies	Computer Numerical Control	Artificial Intelligence, Internet and IoT	Artificial Intelligence
		Systems Collation	'Data transmission, storage, and analysis [...]'
		Digital Twins	Digital Twins
	'Flexible manufacturing systems (FMSs) and serus'	Green Energy and 'Efficient' Manufacturing	Green Manufacturing
			Individualised Human-Machine- Interaction
			'bio-inspired technologies[...]'
Source	Yin et al., 2018; Zakoldaev et al., 2020	Li, 2024; Nahavandi, 2019	European Commission. Directorate General for Research and Innovation., 2021

The interest in Industry 5.0 is a driving force behind the current research project. In addition to the benefits of Industry 4.0, the next generation of paradigm puts an emphasis on being user-conscious (European Commission. Directorate General for Research and Innovation., 2021). This has been described in two ways (Leng et al., 2022): one as being 'to enhance man's activity' and the other relating to 'the well-being of the workers'. Further, Industry 5.0 has been described as a 'forward-looking exercise' (Xu et al., 2021).

Thus, arguably, using an Industry 5.0 approach emphasises integrating a Human Factors approach. This ensures that the upcoming technologies are just as beneficial to the user as the organisation that implements them. Industry 5.0 presents the opportunity for research to utilise the Human Factors and technology-driven approach as a basis for future computing systems.

Industry 5.0 implementation is not straightforward and has its challenges. As will be discussed in the motivating work, there are concerns about what data is captured and how organisations will use this data in Digital Manufacturing Technologies. This is a problem for organisations who wish to implement Industry 5.0, as this data is often a requirement for the beneficial technologies present within the paradigm. Further, as more and more people come into contact with advanced technologies which malfunction in their everyday lives (Clayton, 2024), suspicion will be high for any organisation which wishes to use these types of technology. Thus, the focus should not only be on which type of system will be most effective but in addition, how can this system be implemented in a way users feel comfortable with. For Industry 5.0, the investigation of the potential impact of these challenges is key for the successful implementation of the paradigm.

1.2 Understanding Personalisation

Personalisation is the chosen approach for the thesis to examine Industry 4.0/5.0. Personalisation is defined as functionality that takes personal data and uses it to change a system (Chapter 3: Codesign with the Personalisation Design Cards). For example, a computer-controlled machine could keep track of the time stamps of when it is turned on and over the course of multiple days. The machine tracks one employee's usage schedule (personal data) and uses that data to turn itself on (changing the system state). The machine does this completely autonomously (without user input). If a timer was set by a user to turn it on at a set time, this would be customisation, not personalisation (Sundar and Marathe, 2010; Zhang and Sundar, 2019, Chapter 2: Taxonomy of Personalisation Systems). Further, Personalisation is easily thought of as a whole "system" but is often a separate functionality contributing towards a wider system. In the provided example, the machine (or system) may produce parts as its main role, with personalisation used to improve the usability of the system as additional functionality.

Industry 5.0 is formed of various technologies which are suitable candidates for further exploration (Table 1) (European Commission. Directorate General for Research and Innovation., 2021). Personalisation shines due to its natural basis of being a 'human-centric' technology. The reason personalisation is beneficial is the link between personal data and a person's needs and wants. Sailaja et al. (2019) present the Living Room of the Future: 'The experience begins when the audience sits on the sofa: the pressure sensors in the seat cushions trigger the blinds to close automatically, the lights adjust to their colour, and the movie to begin'. While this is only a demonstration in a non-manufacturing setting, there is the potential for the system to understand when a user performs an action and react in a best way for that specific user, is a useful proposition. Using the UK Governments suggests a Car Manufacturing Worker will have to 'fit interiors' and 'fix engines and frames to vehicle chassis' (National Careers Service (UK Gov), n.d.). This type of assembly work can cause physical fatigue. A personalisation system would be able to understand this situation and put into place measures to alleviate this for the user. The ability to complete this automatically, would allow for fast interventions which could avoid dissatisfaction from users.

The choice of personalisation can be further justified by its benefits. For a Mid-Air Gesture Keyboard, it was found that users' performance rose when using their chosen personalised keyboard (Shen et al., 2022). For organisations which employ many employees, it would be impractical to personalise each interface for each person. Personalisation shows the performance benefit which can be gained by using an almost fully autonomous approach. Personalisation has also been shown to increase engagement with advertisements (Heerwegh & Loosveldt, 2006). This could be used to draw attention to health and safety prompts or important changes to workplace procedures beyond influencing in-the-moment task performance. The flexibility of personalisation to provide benefits to different parts of the manufacturing process is key in its choice as the chosen technology of the thesis.

The current thesis suggests that Personalisation systems have different ways to intervene (Chapter 2: Taxonomy of Personalisation Systems). Intervention refers to when (as in Blom's taxonomy (2000)) and how a personalisation system is activated. This concept forms the basis of a proposed taxonomy, which categorises different personalisation systems into five categories. For example, Swappers will add one piece of content in place of another. In contrast, Generators create new content for users based on their personalisation data. Using this taxonomy, it is possible to differentiate personalisation systems into categories based on the user-facing output of the system (termed intervention). Further, as the intervention amount increases the amount of data required to craft a reliable system increases. Systems implementers need to account for both elements when designing personalisation systems.

1.3 Motivating Work: User Attitudes Towards Digital Manufacturing Technologies

1.3.1 Background

The thesis was motivated by a collaborative work with the DigiTOP project. Before personalisation was chosen as the topic, it was key to understand the current landscape of the Industry paradigms. As Industry 5.0 looks to improve the workplace from the viewpoint of the user (Leng et al., 2022), it is important to understand how users may feel about upcoming technologies (henceforth termed Digital Manufacturing Technologies or DMTs). This will aid in identifying potential consequences before their implementation.

Prior work has examined a range of potential concerns regarding DMTs. One of the main concerns of users has been revealed to be a 'lack of specialized training' (A. C. Shinohara et al., 2017). When technologies are brought in quickly, there can be a gap between the skills currently held by users and those required for the effective operation of the technology. This is a similar concern of 'Reskilling' found by Leesakul et al. (2022), which also identified the topics of 'Trust', 'Data and Privacy' and 'Technology Adoption'. These topics demonstrate the wide set of potential obstacles that need to be resolved in order to successfully implement DMTs.

The ContraVision approach (Mancini et al., 2010) was utilised to understand attitudes towards DMTs. In line with the method (Mancini et al., 2010), two videos presented how DMTs can be utilised in a positive and negative light. This allows participants to react to the content and provide their views about DMTs coming from two different angles: 'Utopic' and 'Distopic' (Marinescu et al., 2022). As the work only serves to provide a motivation for the research questions of the thesis, only a discussion of the results are provided herein (further information on the method is provided in Appendix A).

The current section highlights the key outcomes from the collaborative work which provide motivation for the thesis (Marinescu et al., 2022). As the collaboration was commenced at a late stage in the study, data had already been collected. This section revolves around the thesis authors contribution to the study, a qualitative thematic analysis. This analysis is expanded to provide additional context. The aim of this motivating work is to form an initial look into how people feel about digital manufacturing technologies. This, in turn, can provide a motivation for the building of the research questions of the thesis. As personalisation shares similar characteristics with aspects of DMTs, like using data, the insights generated can be transposed onto personalisation even without its precise inclusion.

1.3.2 Discussion

The thematic analysis provides motivation for the current research project by demonstrating how users currently feel about digital manufacturing technologies. The participants in the current study provided significant amounts of negative viewpoints in both conditions. For example, the discussion around how people may not be viewed as such, with the introduction of digital manufacturing technologies. This is an interesting perspective as the goal of these technologies is often to improve an aspect of the performance of the user or production line as a whole. It is also possible to see elements of distrust, as found in the Negative Outlook on Technology sub-theme. Participants believed that these technologies would bring with them drawbacks that may provide harm to their standing in the workplace. This is a concern for businesses that wish to implement these types of systems to move to an Industry 4.0 paradigm. There has always been a level of pushback against new technologies (e.g. historical luddites), but the large scale of the potential negative aspects discussed here warrant further exploration.

The second major finding was that of discussion into Data. With a theme dedicated to Personal Data, it becomes clear this was a large topic of discussion. Participants were wary of how data is being collected by these technologies and how this data may then be used. It appears that personal data is felt to be private and should be kept to the individual. This is not unexpected (see prior research on value of personal data; Skatova et al., 2013); however, it poses difficulties to businesses looking to engage with technologies that capture personal data. If users do not wish for this data to be captured or used, the implementation of this technology will become restricted.

The research questions included in the PhD thesis are influenced by the findings from this study. While not directly building into each question, the inspiration from these results can be specified. The negative discussion around digital manufacturing technologies regarding Individual Workspaces and People, indicates a need for study into how to make using or the future implementation of these technologies a more positive experience for the user (see Industry 5.0; European Commission. Directorate General for Research and Innovation., 2021). This is reflected in the research questions, which all take a stance of learning towards using technology for the purpose of the users' experience. For example, the first question relates to the "design" of personalisation systems. These concerns also suggest a need for understanding what types of systems users feel concerned with. The second research question links to this by examining the acceptance of personalisation systems. Further, it appeared that there is a potential problem with using Personal Data in future systems. For technologies such as personalisation, the imperative nature of personal data points to a need for investigation to be completed into how data is included within systems while maintaining user approval. This provides further motivation for the inclusion of the user perspective in understanding personalisation within the research questions.

1.4 Contributions and Novelty

As the thesis aimed to provide insight into personalisation systems within manufacturing environments the outcomes will be presented for potential implementation of personalisation technologies. With the external partners of the project being organisations with links to industry (DigiTOP and Made Smarter Centre for People Led Digitalisation), the work incorporates these with a contribution to both academic and industry settings. Thus, the target audience are system implementers, as this could be implementation for future research or implementation for a business use case. The following section identifies novel contributions which are made within the thesis (Table 2).

Table 2 – The contributions of the thesis and how they relate to specific audiences

Contribution	Description	Audience
Taxonomy of Personalisation Systems	A classification of personalisation that will enable easy comparisons between different types of systems.	Academic/ Industry
Personalisation Design Cards	A research method to enable codesign of personalisation systems.	Academic/ Industry
Design Recommendations	A number of design recommendations relating to data in the workplace, workplace procedures and benefits for users.	Industry
Knowledge	Expanded knowledge on the academic side of personalisation systems such as acceptance and personal data.	Academic

The work aimed to contribute to the broad range of work currently published on personalisation systems. As current work produces a new taxonomy of personalisation (Chapter 2: Taxonomy of Personalisation Systems), being able to categorise personalisation effectively will aid in the correct research into personalisation. Further, as prior work may describe personalisation as a universal group (e.g. Kozyreva et al., 2021) or in small categories (e.g. Zhang & Sundar, 2019), often only a few types of personalisation are included. With the provision of the wider categorisation system, authors will be able to better justify the logic as to their implementation and link to other works in a more consistent way.

The contribution will also be the developing of the insights found within existing literature. The Mixed Reality ideation cards (Wetzel et al., 2017) were redesigned with a focus on personalisation systems in manufacturing environments. These new ideation cards contribute by providing a unique way for the academic field to interface with industry. Further, the thesis also contributes to Human Factors literature by linking certain taxonomy categories to the Decision Ladder concept (Jenkins et al., 2010; Frank Gleeson & Vincent Hargaden, 2014) (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). This provides additional validation to the generated taxonomy and provides insights into how systems can be mapped onto the Decision Ladder to provide improvements. As the use of the manufacturing industry for personalisation is a mostly undiscussed topic, the experiments provide a platform for future work to build.

The thesis features a series of design principles for industry developers by providing a basis for the implementation of personalisation systems within manufacturing and potentially spurring on the future development of the technology. As personalisation is not often listed as an Industry 4.0/5.0 technology, organisations may not be aware of its potential benefits to businesses. Further, small to medium organisations may not have the spare capacity to research upcoming technologies. This work provides a “first step” to allow organisations to understand the technology and potential use cases. The empirical chapters include an aspect of the industry use case built into their design. One such chapter (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) included vignettes of theoretical industry use cases for personalisation. Placing the research into the manufacturing sector enables the insights to be easily related back to industry settings and helps a system integrator be confident that personalisation has been tested in the environment where they are aiming to integrate the technology. These factors build a strong case for the contribution to the manufacturing industry.

The work provides a contribution to the “Industry” paradigms. By advancing personalisation, it should be possible to develop more sophisticated automated systems, one of the elements of Industry 4.0 (Lasi et al., 2014). As personalisation can promote additional “training” (Chapter 3: Codesign with the Personalisation Design Cards) or assistance within a task (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). The thesis also heavily suggests that personalisation is one potential technology for user-focused design that is featured in Industry 4.0 (Lasi et al., 2014) and 5.0 (European Commission. Directorate General for Research and Innovation., 2021). The development of personalisation in this thesis will demonstrate how systems can be made with the user in mind.

1.5 Problem Statement

The current work aims to provide a platform for system integrators to implement personalisation systems in a way that is mutually beneficial for users and organisations. As described, certain Industry paradigms may not fully provide strong enough support for users (Xu et al., 2021). The thesis provides a set of design recommendations for the persons responsible for integrating systems, so that they are implemented in a form that is beneficial to the user (e.g. retains users concerns around data; Marinescu et al., 2022).

The choice of personalisation is related to the concerns around Industry 4.0 and the need for a user-first Industry 5.0 paradigm (European Commission. Directorate General for Research and Innovation., 2021; Xu et al., 2021). As personalisation naturally uses personal data with a goal to change a parameter of a system (often to the users benefit). The current thesis argues that personalisation can be used to provide increased usability.

The work provides an exploratory basis for personalisation's usage within manufacturing. As personalisation requires a significant amount of data, it is a technology currently in use in social media platforms like X or Facebook. As the manufacturing industry moves towards increased digitalisation, an increased in data captured may lead to technologies -like personalisation- becoming available for usage by organisations. As there is a limited amount of research into personalisation within manufacturing, by producing the current thesis, the design recommendations are available for future personalisation systems.

The thesis aims to explore the problem of personalisation's usage for an effective Industry 5.0 paradigm. It achieves this through primary research into different aspects of personalisation integration: design, acceptance and functional usage within tasks.

1.6 Research Questions

Personalisation can enable the 'human-centric[ity]' that is found in Industry 5.0 (European Commission. Directorate General for Research and Innovation., 2021). Thus, the inclusion of personalisation speaks to the Industry 4.0/5.0 paradigms that organisations are moving towards. While this technology has the potential to address this idea, personalisation is not widely used in manufacturing. Thus, there is a need to examine how personalisation would function in this specified industry environment. Understanding how personalisation can be used within automated manufacturing environments is the aim of the thesis. This aim can be broken down into smaller research questions. As the research takes a view from Human Factors and usability standpoint, the questions reflect this by focusing on the users as a key aspect for investigation. The three research questions are as follows:

1. How would co-design be used to design personalisation within automated manufacturing environments?
2. How accepting of personalisation are end-users?
3. Can personalisation impact an end user's interaction within an automated system?

For the current project, each question had an associated primary study. In which the main aim of that study was to resolve the research question. The research questions will be described (in a common approach to PhD theses and this is represented in the diagram, Figure 1). The research questions can be linked to the motivation work, which highlighted reservations about 'personal data' which could affect personalisation. The objectives of the thesis can be understood through each empirical chapter (Table 3).

Figure 1 – The relationship between the research questions and the empirical chapters

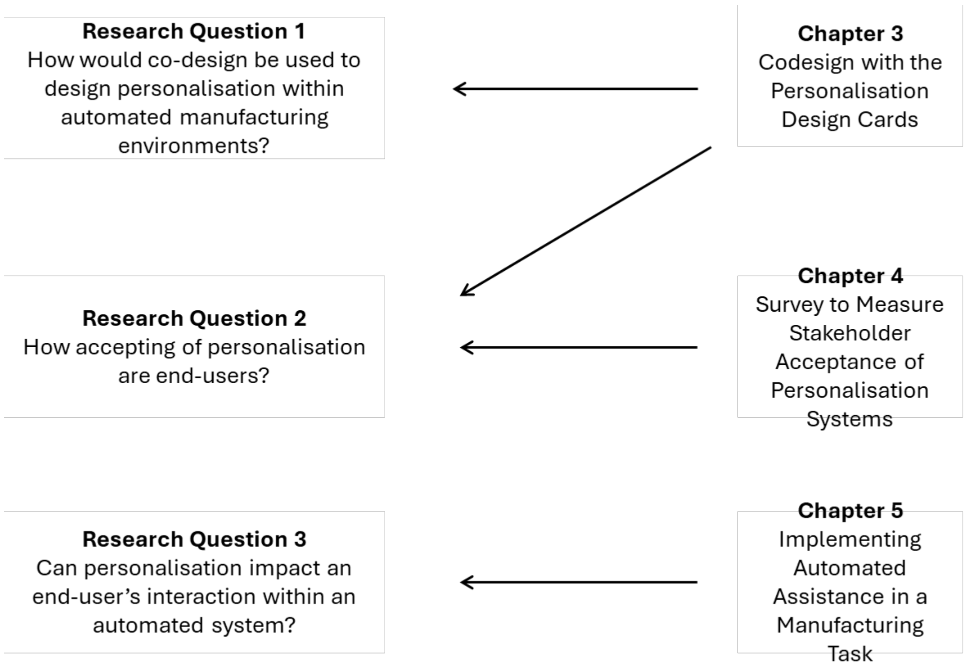


Table 3 – Objectives for the empirical research

As each of the chapters was designed to resolve a specific question (e.g. Chapter 4 to RQ 2) the objectives for each chapter naturally represents the objectives of a research question. The objectives were derived from the sub questions of the empirical chapters.

Study	Objectives
Chapter 3 (RQ1)	To understand how users design personalisation systems within manufacturing environments. To use the created designs to learn what types of personalisation is acceptable. To develop a set of ideation cards (similar to prior work; Lucero and Arrasvuori, 2010).
Chapter 4 (RQ 2)	To explore how acceptance factors from prior literature

	To understand if the type of personalisation system has an effect on acceptance. To identify a link between acceptance and choice to use a system.
Chapter 5 (RQ 3)	To understand if the type of personalisation systems effects user performance. To identify users perception of personalisation systems is varied dependant on the type implemented.

1.6.1 Research Question: Co-design of Personalisation

The first research question encapsulates how co-design could be utilised to design personalisation systems within automated manufacturing environments. The inclusion of co-design relates to the assumption that people may design systems based on their ideas of how a system should be designed. The co-design workshops promoted the participants to discuss and create their own personalisation systems (similar ideas are suggested in participatory design literature (Brandt et al., 2012). Thus, this enables the study of how personalisation should be designed from the perspective of someone who is not the lead designer of a given personalisation system (this being a traditional part of participatory design (like Muller, 1991). By allowing the general members of society to design personalisation, a greater understanding will be gained of how people who may interact with these systems in the future would feel when these systems are actually implemented.

To answer the first research question, a set of ideation cards was implemented, which were used in co-design workshops (Chapter 3: Codesign with the Personalisation Design Cards). These cards included a range of different prompts to turn their thoughts into personalisation systems without the need for aforementioned specialist knowledge (Muller, 1991). The cards, when combined, proved sufficient for participants to generate a strong range of personalisation systems and provide accompanying contextual information about their creations. This enabled the research output of discussion of how people could design personalisation.

1.6.2 Research Question: Acceptance of Personalisation

The second question aimed to understand if end users are accepting of personalisation. Selecting end-users as the population to be studied allows the findings to be related to the potential user base of the technology. As the research topic is geared towards system implementers in industry, ensuring that the outputs represent that environment is instrumental to effective implementation. One metric to understanding how end-users feel about personalisation is acceptance. Prior work has linked acceptance to technology (e.g. Kulviwat *et al.*, 2007). For the current project, using acceptance should allow the discovery as to whether users feel that personalisation in certain settings is appropriate or whether a different approach should be taken.

To understand this problem, an experimental study was conducted, which incorporates a survey-like design (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). As acceptance is a dependent variable, two different methods of gathering acceptance were chosen from existing literature. These were given to participants after viewing different vignettes, which contained short examples of personalisation use cases (and in some instances, the described input data types varied). With two forms of participant characteristics taken using existing scales to account for different backgrounds, this chosen approach was deemed to provide a rigorous answer to the chosen question.

Acceptance was also a secondary topic of examination for the first and third empirical studies. The co-design study (Chapter 3: Codesign with the Personalisation Design Cards) allowed people to design their own personalisation systems, which promoted discussion into personalisation. This contributed to acceptance by the participants naturally talking about the systems they prefer and which they would rather not be utilised. The Implementation study (Chapter 5: Implementing Automated Assistance in a Manufacturing Task), by allowing participants to respond openly after utilising the study, can describe whether they had issues with or found success with using personalisation. By providing this opportunity to participants, insight can be found into notions of acceptance within the utilised systems.

1.6.3 Research Question: Understanding Personalisation Implementation in Automated Systems

The third research question frames the discussion around end-users using personalisation systems. Personalisation would be found within or as part of an existing automated system. Thus, the selected environment should be one with an automated system present. As personalisation in the current work focuses heavily on the user, it follows that the integration of personalisation in the current work will look to affect the user's actions within a task. Thus, thoroughly examining how personalisation impacts the user is key to understanding how personalisation can be used. With the discussion into the range of different personalisation systems (Chapter 2: Taxonomy of Personalisation Systems), there is scope to examine the roles varied personalisation systems can have in a task.

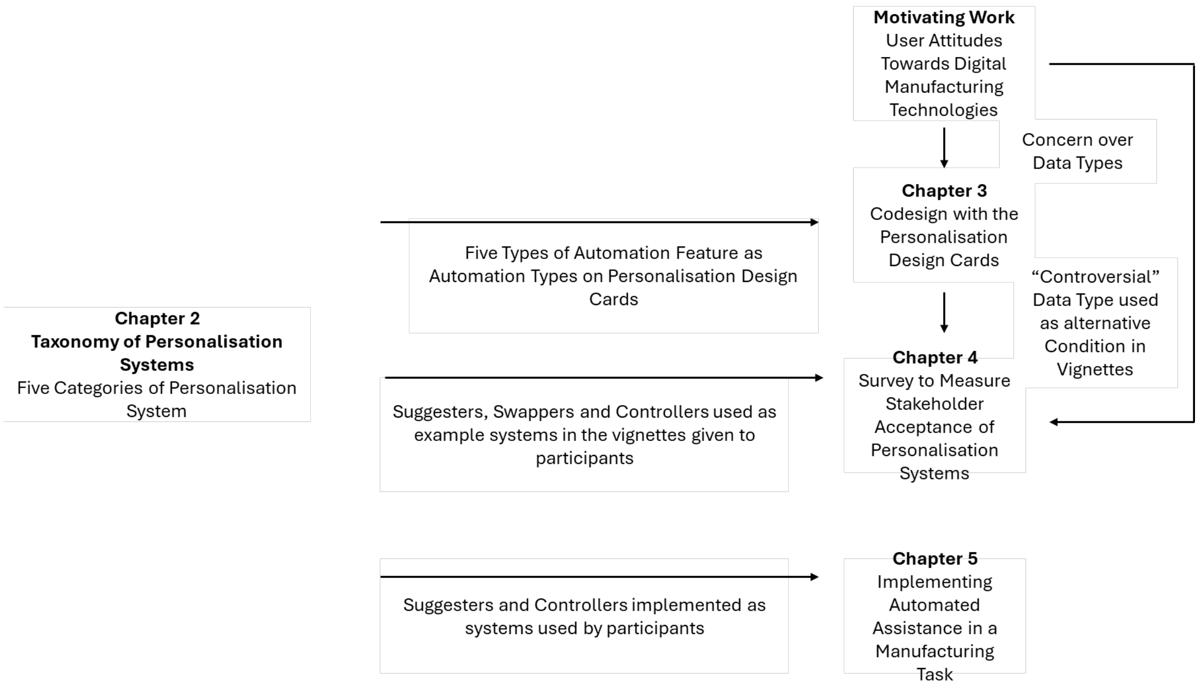
The final empirical chapter used an experimental design to solve the research question. A representative task in the form of Quality Control was chosen, which required participants to identify defects in the cork tiles. An assistive automation was present which aided participants in their choices. For the first experiment, the automation's reliability was altered to aid in understanding how the reliability of automation would affect the task and participants' thoughts towards it. The second experiment had a specific focus on personalisation, incorporating two different types of personalisation system and requesting participants complete the task. The results of the experiments provide an understanding of how participants react to an industry task with different types of automated assistance (the majority being personalised).

1.7 Methodological Approach and Structure

To explore the topic of personalisation within automated manufacturing environments, this current project utilised a set of different methods. The methods used vary due to the nature of personalisation. As Industry 4.0 has multiple associated technologies (Lasi et al., 2014) (and to a lesser extent Industry 5.0) the thesis needs to incorporate potential use cases. This requires examples of personalisation systems to be generated and utilised as part of the research. Thus, the methods chosen to reflect this and use cases are present on the novel ideation cards in the first empirical chapter (Chapter 3: Codesign with the Personalisation Design Cards). In the next chapter, they are in vignette form (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) or were implemented as part of a working system in the final experiment of the thesis (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). The Industry 5.0 focus on the user (European Commission. Directorate General for Research and Innovation., 2021) is echoed in the methodological approach. Each method studies how users could be impacted by the technology, such as the use of an experimental design representing a potential real-world task infused with personalisation (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). The user side was studied through the use of existing scales and a qualitative question given after the utilisation of the personalisation. It is suggested that this forms an interdisciplinary methodological approach by combining both Computer systems development and Human Factors.

The Taxonomy of Personalisation Systems (Chapter 2: Taxonomy of Personalisation Systems) is reflected throughout the thesis and is an important element in each of the empirical chapters' methodological designs (Figure 2). Each of the empirical chapters builds upon the notion that personalisation systems can be categorised by this taxonomy. Defining personalisation clearly allows the thesis to understand how users react to different types. This is key as while personalisation attempts to provide a "one size fits all" approach for users, the environment becomes the area where the system needs to be specifically designed.

Figure 2 - Diagram showing thesis structure and how the Taxonomy of Personalisation Systems (Systematic Literature Review) is built into the empirical chapters



With the challenges presented by the COVID-19 pandemic, the methodological approach was chosen to design the research studies to be resilient to potential disruption. For example, the majority of the studies were conducted remotely using online platforms. Not only does this reduce the effect of disruption on in person studies, it provides greater reach (for potential stakeholders) not possible when recruiting from a local environment. Further, the use of remote experiments provided a level of consistency which is difficult to replicate in offline studies. This is due to the study being almost entirely "hands-off", with the experiment being posted online and participants completing the experiments at their own pace. Thus, it is more certain that participants had similar experiences with the experiment and that the study had to be rigid enough to account for unusual edge cases without requiring experimenter intervention. This usage of remote experiments worked well and allowed for a rigorous scientific approach to be completed while accounting for the practicalities of the COVID-19 pandemic's effect on in-person experiments.

The motivating work uses the ContraVision approach (Mancini et al., 2010) to understand how potential users of Digital Manufacturing Technologies would feel about these systems being implemented (Marinescu et al., 2022). The study utilised two videos showing a positive and negative implementation and then recorded user attitudes. It showed that potential users may have reservations about personal data and how these sorts of technologies can have effects in work and outside of work. This is crucial to understand as DMTs may be controversial (as shown in the current work) and being able to identify potential problem areas is useful in attempting to resolve them. In this instance, the motivating work's findings relating to personal data aided in shaping the current research questions and theme of the PhD thesis.

To refine the topic further, a literature review was conducted in order to describe the current state of the literature and synthesise existing works to validate a novel taxonomy (Chapter 2: Taxonomy of Personalisation Systems). The taxonomy identifies five types of personalisation systems. This crafts a running theme throughout the work in which these different types are explored to understand the effects of different types of personalisation systems on a range of factors, such as user attitudes or task performance:

- Suggesters suggest content to users.
- Rearrangers change locations of content.
- Swappers replace one type of content with another.
- Generators create new content.
- Controllers take control of the wider system.

The work includes a list of systems from the academic literature and their associated personalisation systems (Chapter 2: Taxonomy of Personalisation Systems). The main benefit of this approach is each of the following studies can describe consistently the type of personalisation system they use, and any follow-up studies will be able to understand what system is used and how this system works in practise. For example, in the Acceptance Study (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems), a set - featuring an example of one of each Suggester, Swapper and Controller systems- was utilised to understand how users view these different types of systems. The terminology aimed to create a mental association so that one associates Suggesters with low intrusiveness-low data usage systems and Controllers with highly intrusive-high data usage systems.

The first empirical chapter (Chapter 3: Codesign with the Personalisation Design Cards) utilised a participatory design approach (Brandt et al., 2012) to provide insight into the design of personalisation systems (Duvnjak et al., 2024). Groups of participants utilised a newly designed set of ideation cards to design their own personalisation systems. The cards reflected both the taxonomy of personalisation systems and manufacturing use cases. The rationale was that participants would be able to convey their thoughts and experiences while designing and explaining their own systems (Brandt et al., 2012; Muller, 1991). The data from the workshops were analysed using a qualitative method. The workshops were successful at promoting discussion into personalisation systems, leading to insights into how they should be designed and led to a redesigned set of cards.

The second empirical chapter (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) was an experimental quasi-survey presented to potential stakeholders in the manufacturing industry. The general aim was to understand how participants responded to personalisation systems being used in industry settings. The survey consisted of multiple pre-study scales to capture participant characteristics; this was to be able to examine if a user's background could affect their acceptance of the system. The participant was then presented with a vignette of a personalisation system, which was followed by two acceptance measurement devices. The chapter consists of two experiments that are broadly similar, but the second experiment has participants view multiple example systems and includes an extra differentiator of changing the data type present within the vignette for different groups of participants. The study located a difference between how participants perceive personalisation systems of different data types, only in one of the data types. The study was not able to definitively show any effect of participants' background characteristics on their acceptance of personalisation systems.

The final empirical chapter (Chapter 5: Implementing Automated Assistance in a Manufacturing Task) includes two types of automated assistance for a representative task. The task consisted of identifying defects in cork tiles. The first experiment included a basic form of automated assistance (decision support) and aimed to understand how users would feel about being aided by an automated system of differing reliability levels. This revealed an ability to recognise how reliable the automated system was without being informed that this was a condition in the experiment (Implicit Learning (Reber, 1993)). The second experiment examines working forms of personalisation. The two different types of personalisation were implemented at either end of the taxonomy of personalisation systems. One system suggested the defect to the user, while the other took control away from the user. Both of these assistance types were activated in instances of lowered performance from the user. This experiment was unable to locate any concrete differences in the quantitative results, but the qualitative analysis did demonstrate what users thought about the personalisation systems.

Throughout all of the empirical chapters, a consistent theme appears. This is the inclusion of various data types in the methodological approach. The first empirical chapter places different personal data types as part of the novel ideation cards. The acceptance survey features two different data types as conditions in the second experiment. The final empirical chapter utilises performance data to drive the personalisation system. As 'personal data' was a concern noted in the motivating work (Marinescu et al., 2022), its consistent inclusion demonstrates the timeliness and importance of the current thesis.

1.8 Conclusion

The initial chapter of the thesis provided a foundation for the work. It discusses the motivation being the Industry 4.0/5.0 paradigms and the structure of the thesis. A motivating work section was included to further motivate and provide backing for the research questions. The defined research questions are built from the overall aim of the thesis to understand how best personalisation can be effectively utilised within a manufacturing environment. The specified three questions relate to design, acceptance and implementation and the empirical studies look to address one research question. The following chapter will describe the literature review, featuring a novel taxonomy of personalisation systems, alongside a history of personalisation and a discussion of technical approaches for the technology.

Chapter 2: Taxonomy of Personalisation Systems

2.1 Introduction

Motivated by insights about attitudes towards Digital Manufacturing Technologies (DMTs) (Chapter 1: Introduction) (Marinescu et al., 2022), the thesis' research questions cover design, acceptance and implementation. However, for personalisation to be used in manufacturing, understanding personalisation as a technology is key. The objectives of the chapter are to describe a novel taxonomy of personalisation systems and validate the effectiveness of the taxonomy through a systematic review.

Definitions of personalisation can often relate to non-manufacturing settings, such as 'the objective of personalization is to tailor the promotion and advertisement to match each viewer's interests' (P. S. Yu, 1999). This definition excludes systems like Benyon and Murray's (1988) Monitor System, which aims to restrict user action based on personal data. The output of this system does not fit within the prior definition. Thus, there is room for a new concept of personalisation which can account for the various ways personalisation can interact with users. This work suggests that personalisation uses a user's personal data to provide automated responses to users. A "response" could be content or an action.

With the existing concerns around Personal Data (Marinescu et al., 2022), its role in personalisation should not be excluded from the discussion. Adomavicius and Tuzhilin (2005) state of the recipients of personalisation that 'they may be unwilling to provide some of the personal data needed to improve the system's ability to make recommendations'. Thus, access to personal data may not always be possible for personalisation to take place. This is compounded by the regulatory restrictions that come alongside personal data (e.g. *Regulation (EU) 2016/679*, 2016). Personal data can be suggested to be a limiting factor as to what type of personalisation system can be implemented.

The current chapter attempts to group different personalisation systems into one unified taxonomy. The proposed taxonomy separates personalisation systems by their intervention (or output, a similar approach is taken by Blom, 2000) and the data required. The aim of this chapter is to describe personalisation systems, depict the relationship between personalisation systems and understand how personalisation is currently implemented. This will allow the following chapters to utilise the developed classifications of personalisation system to understand the potential effects on manufacturing settings for different types of system. The taxonomy was initially formulated by separating personalisation into five categories and describing how these are related to one another in an initial review stage of work selected by convenience. In order to validate that the taxonomy is accurate and can be used moving into the future, it was validated by a systematic review method (EBSE, 2007) spanning ten years of existing academic personalisation systems. Before describing the taxonomy, the current historical context of personalisation can be detailed, alongside technical approaches of the technology.

2.2 Personalisation Context

For the current thesis personalisation was chosen as the technology. However, there were other approaches that could have been taken. Augmented Reality can improve usability by displaying additional information to users, often through a user viewing a screen. In an manufacturing context, this could be displaying assistive automated decision-making during an assembly task (for a similar alternative See et al., 2017). The use of Virtual Reality could also provide an alternative to personalisation. Using physical headsets with displays inside, virtual reality allows the users physical movements to be recreated inside a digital world. Personalisation provides a distinct benefit over these technological approaches, the natural integration of Industry 5.0's requirement for user-conscious design (European Commission. Directorate General for Research and Innovation., 2021). As the manufacturing sector will aim to incorporate the new paradigm, focusing the thesis around personalisation will provide a forwarding looking approach which can meet the needs of future industry.

Personalisation is not currently a technology finding use within manufacturing in the form proposed by the thesis (at least not publicly acknowledged). Personalisation is often used to personalise products for end users (as Mass Personalisation as a Service) (Aheleroff et al., 2021). The final personalised products (termed Ultra-Personalised Products and Services) will have 'added value' by better matching the users (Torn & Vaneker, 2019). This is the current trend for personalisation in manufacturing. The current thesis uses personalisation on the employee side rather than the end user. This should provide an alternative use case for the technology which could provide future gains in usability for employees.

At present, Personalisation is mainly found in mainstream information technologies and platforms, such as X or Instagram. However, personalisation significantly predates its use in internet platforms, where it was found in Mail Surveys (Kerin, 1974). This was occasionally similar to modern personalisation, in which user data was used to change the front of the surveys (Kerin, 1974). Personalisation was often making an artefact appear more specific to a user, such as by using 'envelopes with hand written addresses' (Nederhof, 1983). While this is far from the data-driven approach of modern personalisation systems, the concept of making an artefact unique for a user remains the same. As computing was not as developed in this period, it is logical that personalisation was mostly limited to physical artefacts. But with basic text interfaces, there were attempts at personalisation that used a data-driven approach (Ishikawa et al., 1981). Ishikawa et al. (1981) suggest, 'using new instructions made of frequently used patterns of instruction sequences, programs will normally run faster [...]'. This example contains all the elements found in modern personalisation; there is user data being used and the output is beneficial to the experience of the user.

With the increasingly widespread use of internet technologies in the 21st century, personalisation has become more automated and increased its user base to users of online websites, such as MyYahoo! (Manber et al., 2000). The work by Cingil et al. (2000) is an example of the concerns in this time period which focus mainly on technological progress and data privacy. There is a focus on how new internet-based technologies, such as Web Usage Mining (Online data collection) (Mobasher et al., 2000) or Collaborative Filtering (Understanding a user through other users' data) (Mulvenna et al., 2000), can improve the technical side of the personalisation process. But with the greater processing of data, initiatives such as the Platform for Privacy Preferences (Cingil et al., 2000) and The Privacy Act in the United States (Volokh, 2000) were devised. Many of these techniques and concerns are still around over twenty years later, and the turn of the century has marked a shift into data-based personalisation techniques from the rudimentary Mail Survey personalisation. This is best seen from a definition 'personalisation tailors certain offerings (such as content, services, product recommendations, communications, and e-commerce interactions) by providers (such as e-commerce websites) to consumers (such as customers and visitors) based on knowledge about them, with certain goal(s) in mind' (Adomavicius & Tuzhilin, 2005).

The following decade marks a "boom" for web-based personalisation and research is skewed towards this topic. There is a large interest into privacy topics, with personalisation acceptance being discussed as almost transactional in nature (Awad & Krishnan, 2006; van de Garde-Perik et al., 2008). This theme is closely related to the 'tradeoff between information to implement personalization and the potential violations of privacy [...]' (Montgomery & Smith, 2009). Both concepts are formed by users having a resource (personal data) that systems need to benefit the user or the implementer. Awad and Krishnan (2006) also suggest that 'users with previous privacy invasion experience have a lower willingness to be profiled online for personalized advertising'. There appears to be a shift in personalisation being seen as an exciting new technology to a reality that personal data is not an unlimited exploitable resource.

2.3 Technical Implementation of Personalisation

The current thesis places increased emphasis on the outputs of personalisation systems and this is represented in the contributions. What is not discussed in detail is the algorithmic side of personalisation. This was done to focus on the end user's experience, which is the differentiating factor between this and other types of systems. However, the steps of data collection and processing are useful to understand to provide additional context. Further, while these technological approaches to personalisation may have elements that are found in other computational systems, such as Machine Learning (Asaithambi et al., 2021; Yeung, 2018), the novel way in which these are applied to allow for personalisation systems to function needs to be explored to provide additional understanding to how personalisation would be implemented.

A common form of personalisation approach is that of Collaborative Filtering, defined as 'determin[ing] the direction of customer service by grouping customers with similar patterns of preferences/interest based on the basic information of customers and their preferences/expression of interest' (Kim et al., 2020). This type of system is used ubiquitously. Many e-commerce webpages will use systems with taglines in the form of "users also like". It relies heavily on assumptions and thus is found in situations where data on a particular user may be low, but data on groups of users may be large. This is the reason it is so effective in e-commerce websites; a user may be novel without any presence on the site (such as an account or previous visit data), but they can quickly form part of the grouping (or termed Clustering (Cingil et al., 2000)) and receive suggestions. There are also systems such as Interactive Voice Response (Asthana et al., 2013) that rely on previous user data to personalise the interface for new users. This is an example of a system that does not utilise the grouping system of Collaborative Filtering but keeps the rest of the concept intact.

The opposite of Collaborative Filtering is the individual use of the user's data rather than as part of a larger group. This could be gathered from a work task (Bunt et al., 2010) or 'past behaviour of consumers' (Chen et al., 2022). In these cases, the user data is passively gathered; thus, they do not directly specify the data they provide to the personalisation system (Toch et al., 2012). The data gathered from this type of capture is effective for some Personalisation systems. This is because it should be of high validity to the user (in comparison to Collaborative Filtering) and of high detail, allowing for strong insights to be generated. The downside of this approach is the inverse of Collaborative Filtering, as one cannot use the larger group's aggregated data (this may be argued as a benefit). One needs to collect data on each individual user for the personalisation system to work. This can be time-consuming or not realistic in many situations, although Shared User Modelling has been identified to combat this (Montague et al., 2011).

To alleviate these privacy concerns, users could be requested to provide data for the personalisation process to take place (Tam & Ho, 2006; van de Garde-Perik et al., 2008). This provides solutions to the problems with passive data capture, as users can pick and choose what data to provide to the personalisation system. Thus, they could theoretically choose to provide incorrect information or even no data if they feel the system is utilised in a way that the user does not agree with. The downside of this approach is that users who do not provide correct information will often have a reduced experience or be left out of certain functionality. Further, the implementor of these systems may not be able to control the experience for the user; for example, receiving adverts that are not relevant or at worst, could be considered offensive to certain user groups.

There are systems that use parts of the aforementioned concepts together (Mesquita et al., 2002) in what could be described as a mixed approach. Gullà et al. (2015) describe 'novel user interfaces will react with the human behaviour and interact in accordance with the environmental conditions monitored by local sensors'. These types of systems sit along the borderline of Adaptive Interfaces (Haas & Hettinger, 2001) or Context-Aware systems (systems that utilise data collected from the world) (Dey, 2001). In contrast, some systems use non-personal data to predict user data in situations where user data is not usable (Haas & Hettinger, 2001).

Web Scraping links the concept of passive collection and user involvement by using collected user-inputted data from websites (Aguirre et al., 2016). Often, users are not fully aware of the data capture taking place and may not have provided "informed" consent to their data being used in ways outside of common forms of personalisation, for example, online advertising. This data is often gathered in a 'spider[~~s~~]' technique which looks through results of results (Severance, 2016, p148) and thus can gather large amounts of data.

Many personalisation systems function in the digital realm but they can make use of the physical world. The Active Badge (Weiser, 1991) is a concept in which the holder of this device is able to personalise different artefacts (using the term 'customize'). This could be applied to any sort of card-based system, but the idea of using a physical artefact to determine how to personalise is interesting. There are also systems that use the data relating to a person's physical characteristics to generate new physical products (Aheleroff et al., 2021; Zheng et al., 2017). Often, personalisation takes place in the digital realm due to the ease of changing systems online. However, it should be noted that personalisation can also be used for physical artefacts.

There are also ways of using different activities to gather the data required for personalisation. One such way is the ability to play a game to produce insight into the user) (Jimison et al., 2003). The work describes the notion of relating a user's gameplay to their real-life functions by 'compar[ing] user performance to an optimal standard' (Jimison et al., 2003). Game-like qualities are found in other approaches, such as Friendsourcing which relies on associated users providing data about another user (Bernstein et al., 2010). This avoids having to collect data from a user directly while simultaneously providing a form of entertainment to the users participating, avoiding the apathy of completing data collection.

2.4 Methodological Approach

A taxonomy is the proposed approach to best classify personalisation systems. This enables the inclusion of the two selected criteria of data required and intervention. The taxonomy was developed in a multistage process. An initial search of the literature revealed the five categories of personalisation, chosen to reflect the systems found grouped based on the output of the personalisation system. These categories are built to contain the personalisation systems found within the search. However, this set of systems was collected based on their availability and may not translate to the systems currently present within the literature. This formed the part of the final iterative process.

To thus validate the taxonomy, an EBSE (2007) systematic review was performed. The use of the EBSE method, provided a framework in which the review was performed. With this approach, a range of personalisation systems was be taken from a specified time period, which reduced the selection bias and increases the validity of those chosen. The taxonomy was tested against these systems to validate its effectiveness, this found that the taxonomy was able to classify personalisation systems. In the event of a poor taxonomy, it would be expected that systems would fall outside of this taxonomy and further development would need to be performed (but this was not found). The EBSE (2007) review was chosen for its relation to the subject of interest. The development of the initial taxonomy with subsequent systematic review provided an 'iterative' approach found in prior work (Glöggler & Ammenwerth, 2021) to refine the taxonomy where needed. The overall iterative process can be described as follows:

1. Initial search of the literature using a convenience search approach

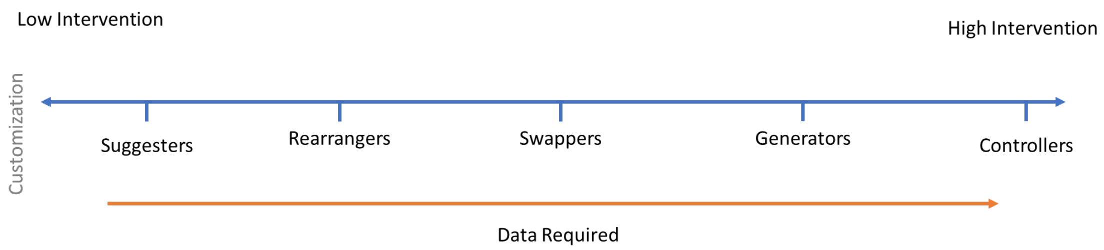
2. Defined categories of personalisation systems
3. Located the personalisation systems present within the literature using a systematic approach
4. Validated the taxonomy using personalisation systems found in the systematic review

2.5 Taxonomy of Personalisation Systems

The novel taxonomy (Figure 3) demonstrates the transition of functionality between categories. Controllers are featured on the right, indicating the high amount of personal data needed and the autonomy needed for this system to operate. The Controllers' systems take over from the user in certain scenarios. The Suggesters fall on the opposite end, being systems that leave most of the task to the user but attempt to have an effect on their experience. It should also be stated that many systems implement many types of personalisation in one application, such as MediaMaps (Wesson et al., 2010).

Figure 3 – The proposed Taxonomy of Personalisation Systems

Intervention is represented by a gradient, with low intervention systems being located on the left and high intervention on the right. The system also links the idea of the data the system has available. Category headers are present to demonstrate the outputs available for the user depending on previous factors.



2.5.1 Suggesters

Suggesters are personalisation systems that increase the visibility of content based on personal data. Increased visibility could be altering the colour of an item, as found with the Visual Popout Interface (Gajos et al., 2006). It could also be the highlighting of content, which is believed would aid in the use of the interface (Bunt et al., 2010). Either way, the intervention is "hands off", allowing the user to have full control over the system, hence the term Suggesters. They are more useful in situations where the personal data is of low quality or there is not the computational means to reliably take control away from the user. In some instances, they could be thought of in the same vein as a warning light often found in interfaces; by design, they make something 'Popout' as named by Gajos et al. (2006). Similar ideas to that of increasing the visibility of prominent content is found in design theory (Brand, 2018, p43-45). This could be thought of along the same lines as the 'Default' concept of Nudges, as users can just assume the automation is increasing the visibility of the Default (Thaler & Sunstein, 2009, p93).

2.5.2 Rearrangers

Rearrangers are a type of personalisation system that alters the location of information based on personal data. The location referred to would be the point in an interface in which the information is displayed. The use of the word location refers to non-visual interfaces such as the IIIT-Delhi system, in which users phone the institution and are directed through an audio interface (Asthana et al., 2013). This work describes a study analysing different types of these systems, one which rearranges the 'node' position based on personal data, which alters the order of the items when they are described to the user. Gajos et al. (2006) describe a Moving Interface in which icons are moved to a more visible location on the interface based on a user's previous actions within the software. In this instance, the location is the interface itself, and the icons being rearranged can make functions that may be more commonly utilised more accessible.

Rearrangers look to alter how the content being shown to the user without adding or removing content. Rearrangers differentiate themselves by moving content around an interface. This is considered a higher level of intervention than Suggesters as users' memory of interface elements is invalidated by the Rearrangers system. There is a discussion that some Rearrangers increase the visibility by the nature of moving an item to the top of a list (for example, the Moving Interface (Gajos et al., 2006)), but the way the term 'visibility' is being used in this case accounts for visual alterations, not locational alterations.

2.5.3 Swappers

Swappers are a type of personalisation system that can add and remove content based on personal data. A clear example can be found in MediaMaps (Wesson et al., 2010). The system is a navigation application similar to common software such as Google Maps, and part of its functionality is the ability to utilise personal data to change the map type. In this case, the maps type the recipient of the swapping functionality. Another example of a Swapper is the concept of a Recommendation Engine (Mobasher et al., 2000). In which a set of items is generated by a user and these items are swapped for different items based on personal data. In both cases, the artefact being swapped already exists, be it all the map variants, and this is changed for the user.

The Swappers type could be considered the most effective of the commonly found personalisation system types. This is due to its ability to balance the need for data and the ability for the automation to have an impact on the task. In many cases, Swappers will provide numerous options for users to choose between, for example, Mobasher et al.'s (2000) Recommendation engine. The

Rearrangers system type is arguably very similar, and this is where the boundaries between personalisation types can become unclear. Generally, both rely on lists and many Rearranger systems will push items off the end of the visible display, which arguably could be swapping content. Thus, the differentiating factor is the amount of items in the list that can be swapped or moved. In the case of Swappers, the items list could be hundreds or even thousands, whereas in the case of Rearrangers this would be much smaller (or only applied to a smaller subset of items), to an extent where the majority of the items which can be rearranged are displayed. It can be suggested that this would mean that there are personalisation systems in-between both categories, which is possible and should be expected.

2.5.4 Generators

Generators are a type of personalisation system that uses personal data to create new content for a user. It is worth making the distinction that this content should be novel and not previously used. An early example is the use of names within documents or communications (Kerin, 1974; Heerwegh and Loosveldt, 2006). By using the user's own personal data, their name, it is possible to place this in various locations within an interface to match it to a user. The fact that the name is not bound by any constraints makes this type of personalisation almost endless. This is the underlying framework of the Generators personalisation system. In general, data has a form of analysis performed before the generations are made. Montgomery and Smith (2009) provide examples of 'personalised price and promotions' in which user data is used to create new sales for customers.

Generators systems are an increase in functionality compared with Swappers systems. In theory, these systems are functionally similar. Swappers replace one item with another item and Generators replace an item with another item. The difference is in the Generators systems with the new item created by the system and a range of potential values is defined. This is what the generation is based on (these ranges generally encompass a larger or potentially near-infinite amount of values). However, in a Swappers system, there are a fixed list of values which are known. The creation of these new items requires a greater level of data as deciding which of the many values is a more difficult task.

2.5.5 Controllers

Controllers are a type of personalisation system that takes control away from a user based on personal data. Due to this, they often require large amounts of data in addition to high amounts of accuracy, as mistakes cause larger problems in comparison to other personalisation types. These systems, thus, are the least common forms and are more theoretical in nature. One such example is the Living Room of the Future (Sailaja et al., 2019). This system represents, as the name implies, a living room experience which uses a user's personal data to affect the user's experience within the room. The idea behind this system is that a user would need less input and would be more free to co-exist with the algorithm which controls the room. In practice, this type of system would require large periods of time for a user to learn and understand how they would like their smart home system to activate around them. A theoretical approach termed Monitor (Benyon & Murray, 1988) takes the opposing form. Rather than controlling systems in anticipation of the user's needs, the system looks to control how users use the system. An example provided in the work is that a 'novice user can be prevented from using certain commands' (Benyon & Murray, 1988). Both of these Controller systems describe functionality that "takes away" control either in an additive light or a restrictive light.

Systems that make use of a Controllers personalisation system would generally use more data and require more in the way of intervention than other types. For example, a Generators system does not take control away from the user and instead relies on influencing a user's behaviour. The user having a role in the decision-making process is useful in situations where the automation is not as accurate or reliable as a human counterpart. The need for reliability comes hand in hand with the need for more data, as in, the more data input into the system, the more confident of the automation outcomes and thus, the ability for the automation to complete the task is increased. In comparison to a Suggesters system, this type of system can be effective with limited data. A new user of a shopping platform could respond positively to personalised advertisements from just three visits to different product pages. The alternative Controllers system that purchases products for you based on three product page visits would seem intrusive (an identified concern (Adomavicius & Tuzhilin, 2005)) and in most instances, not purchase the correct product. As a result, the Controller systems are not commonplace.

2.5.6 Customizers

Customizers are not a form of personalisation, as customisation systems often rely on users making changes to the interface rather than automation (Sundar & Marathe, 2010; Zhang & Sundar, 2019). The inclusion in the current work is to make the distinction between customisation and personalisation. The early format of the My Yahoo page has an example of customisation (Manber et al., 2000). Users are able to edit their page with the content they think they would like (it does feature a basic personalisation functionality, but the main focus is on the customisation aspect). For the most part, the content is not driven by any computation process described in other personalisation systems and is a direct input-to-output system. In traditional UI application settings, customisation can also be used to configure how the information is displayed on-screen (Lallé & Conati, 2019). This is all achieved in real time without any processing required in the background. This means that customisation is very useful in situations such as MQ Transit (Lallé & Conati, 2019), where a user can just walk up to an interface and start using it. The downside is that all the agency is on the side of the user, who may not be inclined to make such changes for themselves.

2.6 Initial Literature Search: Discovering the Taxonomy of Personalisation Systems

Before the systematic review was conducted, an initial taxonomy was developed. This was formed based on an informal literature review including personalisation systems (N=33), as it became clear there were general design traits that each personalisation system incorporates to different extents. The preliminary taxonomy of personalisation systems looks to map the factors of intervention to the outputs of personalisation systems to categorise these variants (Table 4). The taxonomy enables categorisation for different personalisation systems, with the use of a scale format allowing for understanding of how Intervention varies across type of personalisation.

Table 4 - A subset of personalisation systems that formed the initial taxonomy

*Not all systems contain detailed explanations of the practical implementation. Estimations of function were used based on the context. Some systems contain multiple types of customisation and personalisation. * Denoted description included a reference to another work.*

Taxonomy Category	Source	Source's System name/description *and system's function*	Description of Personalisation elements from original/current work
Controllers	Benyon & Murray, 1988	Monitor System *(Personalised Training)*	'the novice user can be prevented from using certain commands, or varying certain parameters [...] the details can be revealed as the users experience grows'.
	Asaithambi et al., 2021	'non-intrusive smart home automation' *(Smart Home Control)*	Uses user data and non-user data to control smart home appliances.
Generators	Montgomery & Smith, 2009	Ding *(Promotional 'applet')*	'This type of system allows Southwest to make unique, personalized offers based upon the customer's history and stated preferences'

	Buganza et al., 2020	Runkeeper *(Fitness Tracker)*	'An algorithm can create personalised goals, combing everything that the app knows about the user, tailoring recommendations and feedback on the user's needs and lifestyle Runkeeper Help Center'
Swappers	Wesson et al., 2010	MediaMaps *(Map software)*	'Visualization adaption is implemented in MediaMaps in terms of adapting the visualization to the previous user's behaviour. [...] In this example, the zoom-level, map style and location have been adapted based on previous user behaviour'
	Gajos et al., 2005 also in Gajos et al., 2006	Split Interface *(custom menu within another application)*	'an adaptive version of the hierarchical interface in which families of the most frequently accessed functions are added to the previously empty "dynamic" area in the middle of the interface [...]'
Rearrangers	Gajos et al., 2006	Moving Interface *(Microsoft Word 'Toolbar')*	'It moves promoted functionality from inside popup panes onto the main toolbar, causing the remaining elements in the popup pane to shift and also causing the existing button on the toolbar to shift and make spaces for the promoted button'
	Wesson et al., 2010	MediaMaps *(Map software)*	'The top section contains the adaptive section of the list, including the most recently used (MRU) option and the two most frequently used (MFU) options' *
Suggesters	Gajos et al., 2006	Visual Popout Interface *(Microsoft Word 'Toolbar')*	'[...] it highlights promoted buttons in magenta'
	Bunt et al., 2010	MICA *(Microsoft Word pop-up screen and 'toolbars')*	'We augment McGrenere and colleagues' direct manipulation personalization facility with system-generated user-tailored recommendations designed to increase an individual's personalization effectiveness

			and decrease his or her personalization effort’
Customizers	Manber et al., 2000	My Yahoo!	‘ [...] is a customized personal copy of Yahoo!. Users can select from hundreds of modules, such as news, stock prices, weather, and sports scores, and place them on one or more pages’
	Lallé & Conati, 2019	MQ Transit *(‘public’ information screen)*	‘a customized mechanism that allows users to hide/display either one of the two visualizations’

Customisations inclusion as part of personalisation is disputed. Seminal sources believe that customisation is not personalisation (Sundar & Marathe, 2010; Zhang & Sundar, 2019); however, it is possible to find work that would include customisation as personalisation (Lallé & Conati, 2019). Lallé and Conati (2019) state that ‘there are two main approaches to support interface personalization: customization done by the user and adaption driven by the system’. In the current work, customisation is differentiated from personalisation by suggesting that personalisation is ‘driven by the system’ and customisation systems are not (Sundar and Marathe, 2010; Zhang and Sundar, 2019). Existing literature has presented similar concepts, such as Reactive and Proactive personalisation (Zhang & Sundar, 2019). However, for the current work, rather than look and when the system is engaged (similar concepts are also found in Blom, 2000; Sundar & Marathe, 2010), the taxonomy presented here identifies and categorises personalisation system’s control on a situation (or rather how much control is taken away from a user).

Personalisation systems rely on personal data to function. The amount of personal data required changes based on the system (Adomavicius & Tuzhilin, 2005). The current work also proposes that the data required for personalisation to work directly affects how much intervention can be made by a system. This is often not a technical limitation; it would be possible to make a highly instructive system with less data. The problem is that the system would suffer greatly in terms of reliability and may not function in a way that would be realistically helpful in a real-world setting. It is doubtful that capturing one single action in an interface is enough data to accurately allow a personalisation system to rearrange one’s layout of said interface. The focus on personal data also differentiates personalisation from the commonly used Adaptive Interface term (Benyon & Murray, 1988; Haas & Hettinger, 2001), of which personalisation systems are a subset. Thus, a novel taxonomy should be able to account for the different amounts of personal data required to allow a personalisation system to function.

2.7 Related Work

In line with the EBSE (2007) review methodology, a 'rationale' for the review was established. To achieve this, a short systematic review was conducted. The EBSE protocol was adapted to also evaluate the existing taxonomies. This would ensure that the novel taxonomy is sufficiently original and contributes to understanding of personalisation systems. The early review included two keyword pairs ("Personalization review", "personalization taxonomy") and only the first ten results on one database (Google Scholar) were featured.

A number of prior taxonomies focus on -what is termed- Web or Internet systems (Malik & Fyfe, 2012; Montgomery & Smith, 2009; Weinmann et al., 2013). Montgomery and Smith (2009) provide a range of personalisation systems: Personalized Search, Personalized Recommenders, Personalized Price and Promotions for Airlines. No specific categorisation is stated to differentiate the chosen groups, but it can be suggested that the effect on the user is the categorisation applied. Malik and Fyfe (2012) take the opposite approach, categorising different approaches to the analytical and data collection parts of personalisation. This provides a key resource to systems developers as to how these systems can be developed but does not categorise the output side of personalisation. Weinmann et al. (2013) attempt a more complete taxonomic approach, split into two halves: User Modelling and System Adaptation (which have further separation). Or, to put it in general terms, these refer to the data and the user-facing output. Further, this is the only work to attempt to specify the search methodology, which is one of the DARE criteria used to evaluate work (EBSE, 2007). In comparison to the other two categorisation systems, Weinmann et al.'s (2013) taxonomy provides a structure that can more precisely group systems based on the whole personalisation process, not just one side.

Blom (2000) follows a similar approach to Weinmann et al. (2013) by providing categorising principles while placing a focus on the outputs of the personalisation system. The taxonomy uses open-ended categories, which could struggle to group systems together and would increase difficulty in following the taxonomy as a formal process. The use of 'degree of initiative' as a categorisation scale is a key concept which separates personalisation systems, as personalisation systems often vary in how they are activated during runtime (Zhang & Sundar, 2019). As the work by Blom (2000) is a poster, it is difficult to ascertain how effectively this taxonomy would group a large number of personalisation systems.

The remaining categorisation systems rely on use cases. Jeevan and Padhi (2006) separate personalisation systems by their specific use cases, such as Personalization in Search Engine Development. Kaaniche et al. (2020) also apply a use case categorisation approach, with Recommendation Services being one such category. Correia and Boavida (2002) include three main categories with one being Content and Presentation Management. Kucirkova (2018) follows this trend by also providing three categories. The benefit of this approach to the categorisation of systems is the specificity of the use case; thus, users in an interest area can easily find personalisation systems that match their needs. The downside of this approach is that systems that could be used in multiple categories either need to be duplicated or systems may not be placeable in any category.

Current taxonomies are often specific to a use case (e.g. Jeevan and Padhi 2006), provide too open-ended categories to definitively categorise systems (Blom, 2000) or focus on specific parts of the process (e.g. Malik & Fyfe, 2012). The taxonomy by Weinmann et al. (2013) is arguably the most effective, allowing distinct categorisation of a wide variety of personalisation systems. However, it lacks distinction between certain types of personalisation outcomes for users, with many systems potentially falling under the 'content' category. The preliminary taxonomy, with its user-facing perspective, finds novelty in this approach. Further, the proposed work represents the categories on a linear scale, relating the categories together by the two defining principles (Intervention and Data Required). These two principles are arguably similar to the main categories suggested by Weinmann et al. (2013) but with a different interpretation of how they should be utilised in a taxonomy. Thus, the preliminary taxonomy is sufficiently original enough in comparison to existing taxonomies to warrant a larger review being applied. This should test its effectiveness as a taxonomy of personalisation systems.

2.8 Systematic Review

The chosen systematic review approach was outlined in the Evidence-Based Software Engineering technical report (EBSE, 2007). The EBSE method was deemed to be relevant and included enough detail to provide the greatest validity to the review in comparison to alternative methods (e.g. Okoli & Schabram, 2010). Practically, it was not always possible to consistently follow the proposed method or defined approach and this was detailed (Staples & Niazi in EBSE, 2007). The current systematic review is naturally adapted in parts of the review where the method provides alternative choices. The next part of this chapter will attempt to retain the terminology from the EBSE method to allow the comparison between the completed review and the theoretical method. The reviewed works will then be used to validate the performance of the proposed taxonomy.

2.8.1 Search Strategy

The search strategy was developed with the guidance of Hart (2018) and the EBSE (2007) methodology. The database was that of SCOPUS, as it should allow for an accurate search of the subject area chosen (Hart, 2018, p266). The terms personalization or personalisation were keywords, alongside the choosing of keywords (or variants) from four papers (Blom, 2000; Gajos et al., 2006; Montgomery & Smith, 2009; Weinmann et al., 2013):

> (TITLE-ABS-KEY (personalization) OR TITLE-ABS-KEY (personalisation) AND TITLE-ABS-KEY (adaptive) OR TITLE-ABS-KEY (interface)) AND PUBYEAR > 1992 AND PUBYEAR < 2024 AND (LIMIT-TO (SUBJAREA , "COMP"))

The earlier systematic review conducted revealed that certain subject areas hold different meanings for words like personalization, so the Computer Science subject area was selected. In addition, only papers published in the last thirty years were included. However, during later stages, this was changed to the last ten years to better match the available resources. The systematic review required the use of a forms to mark the progression of the literature search (Table 5).

Table 5 - Initial Prototype Checklist

The heading search criteria refers to any work which does not pass the study selection criteria which are not covered by the other headings. The systems heading refers to the personalisation systems depicted in the work.

Identifiers			Exclusion Checks				Add. Info.	
Year	Title	URL or location	Title	Abstract	Introduction	Study Selection Criteria	Notes	Systems
2010	Example	url.com		x				

2.8.2 Study Selection Criteria

As the current review is looking to place systems within a taxonomy, only works which adequately describe one or more computational personalisation systems were included. If a system is described repeatedly, only the original description of the system will be taken (unless a significant additional contribution to the description is made), as specified in the method (EBSE, 2007). In addition, work will be manually filtered based on the relevance to the search strategy criteria, such as including personalisation but using a different meaning of the word from an alternative subject area. This includes older forms of personalisation, which can be (but not always) non-computationally based. For example, Nederhof's (1983) “system” is human-operated and the ease of instant algorithmic analysis is not present. Work was excluded based on the focus of the system being on a non-adult population, as it can be assumed that the ethical concern regarding data could mean these systems operate differently compared to other personalisation systems. Papers were filtered based on realistic accessibility criteria, such as whether the paper is available to be accessed or a published English translation of the work is available.

2.8.3 Study Selection Procedure

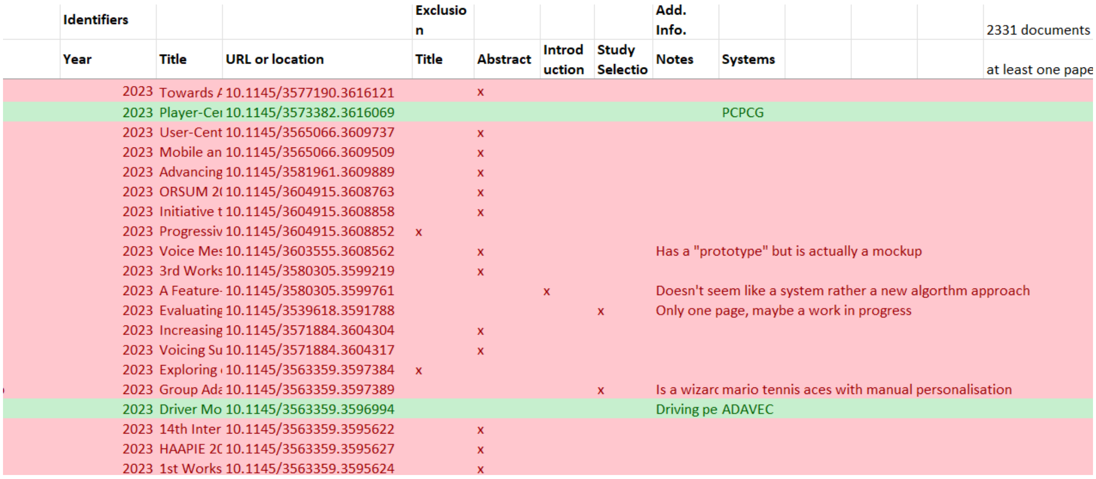
Existing work was excluded based on title and abstract (and other identifying information). Then excluded by introduction, then a full paper scan. When the review was conducted and existing work was related to the research questions, works were removed if they did not fit the above criteria to avoid human error in prior stages being carried into the final review. Following the guidance in the chosen methodology (Brereton et al. 2007 in EBSE, 2007), the review examined a small number of papers at first to examine the approach in practice. The procedures were then adjusted to ensure the review was an accurate portrayal of the literature.

After reviewing a small segment of works, N = 100, it was worth revisiting the search criteria to understand if the current protocol is effective. Fifteen works could not be accessed; it appears some of these were on one platform. Thus, non-accessible works were collected in a list and attempted to be resolved. Otherwise, works that may describe systems may be missed and the validity of the current work would be decreased. Only seven works (which made it past the search criteria) adequately described a personalisation system. One potential issue when reviewing works is that in the constraints of an academic paper, it is difficult to provide all the necessary information and thus, it is not always clear how the system functions. Another point of reflection is that as more papers are reviewed, keywords or phrases are better understood. This quickens or alters the examination of papers. In a general sense, there are no standout problems which would lead to the review protocol needing to be stopped and extensively reviewed. It was found that SCOPUS search results does add papers when viewing the search results, so one extra paper had to be located and excluded. To avoid this problem reoccurring, the search results were saved into two lists of works on the SCOPUS platform. Other than these points, the general plan was effective at finding personalisation systems.

2.8.4 Study Quality Assessment Checklist and Procedure

In the early review phases, work was placed into a table, replicated in a ‘spreadsheet’ (Hart, 2018, p25) (Figure 4), and the checks performed based on the headings provided (**Error! Reference source not found.**). The table includes headings for potential exclusion criteria by title, abstract, introduction and other study selection criteria so work can be indicated if they should be excluded based on these criteria. The different works had additional notes and the systems contained in the works indicated. The final set of works will contain at least one personalisation system to be utilised in the validation of the taxonomy (Figure 5).

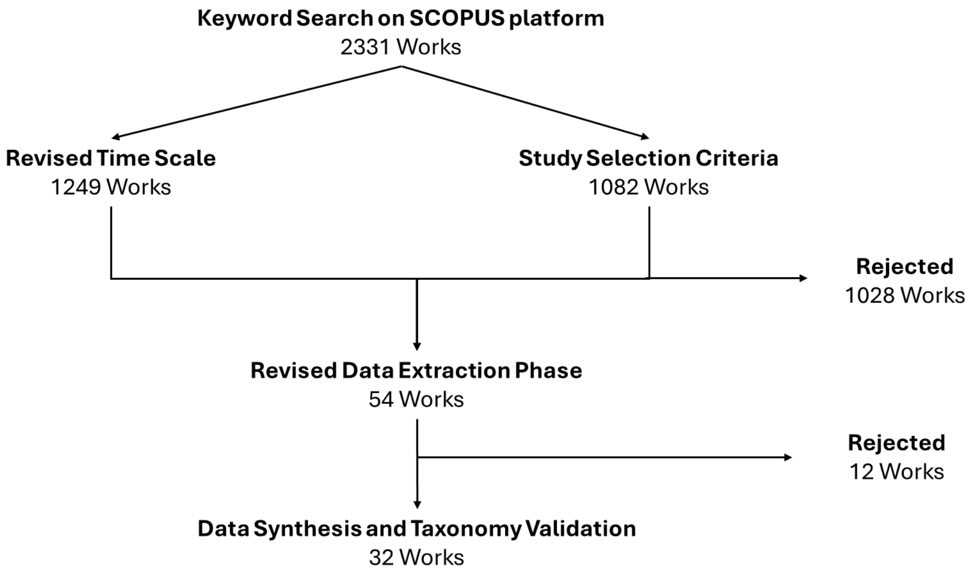
Figure 4 – Screenshot of review spreadsheet



Identifiers			Exclusion				Add. Info.		2331 documents
Year	Title	URL or location	Title	Abstract	Introduction	Study Selectio	Notes	Systems	at least one pape
2023	Towards A	10.1145/3577190.3616121		x					
2023	Player-Cei	10.1145/3573382.3616069						PCPCG	
2023	User-Cent	10.1145/3565066.3609737		x					
2023	Mobile an	10.1145/3565066.3609509		x					
2023	Advancing	10.1145/3581961.3609889		x					
2023	ORSUM 2(	10.1145/3604915.3608763		x					
2023	Initiative t	10.1145/3604915.3608858		x					
2023	Progressiv	10.1145/3604915.3608852	x						
2023	Voice Me:	10.1145/3603555.3608562		x			Has a "prototype" but is actually a mockup		
2023	3rd Works	10.1145/3580305.3599219		x					
2023	A Feature-	10.1145/3580305.3599761			x		Doesn't seem like a system rather a new algorthm approach		
2023	Evaluating	10.1145/3539618.3591788				x	Only one page, maybe a work in progress		
2023	Increasing	10.1145/3571884.3604304		x					
2023	Voicing Su	10.1145/3571884.3604317		x					
2023	Exploring	10.1145/3563359.3597384	x						
2023	Group Ad:	10.1145/3563359.3597389				x	Is a wizarc	mario tennis aces with manual personalisation	
2023	Driver Mo	10.1145/3563359.3596994					Driving pe	ADAVEC	
2023	14th Inter	10.1145/3563359.3595622		x					
2023	HAAPIE 2(	10.1145/3563359.3595627		x					
2023	1st Works	10.1145/3563359.3595624		x					

Figure 5 – Diagram showing the paper selection process

The Revised Time Scale took place during the Study Selection phase, to best optimise available resources. Diagram adapted from Gasteiger et al. (2021).



Upon completion of the search, a number of unexpected factors were identified, which were not accounted for in the original planning. The SCOPUS database has incorrect entries; these appear to be errors with the identifiers of the works. Some works did not follow the standard paper format. To resolve any problems, revaluation was attempted (where appropriate) or excluded, and an additional column was often used to highlight these works.

2.8.5 Data Extraction Strategy

The initial data extraction was to be completed during the document search as works were being read, but this was inconsistent. Works that were unable to be classified by this information and thus, a new data extraction approach was utilised. This included a more in-depth "description" of the system if it fits into the current taxonomy approach and the recommended categorisation for the system (Table 6). In some instances, a work might contain multiple systems and thus, would contain multiple recommendations. Further, the rejection of works which could not be categorised or accessed was required. This revised data extraction was successful and allowed the comparison of the systems present in the review and the initial taxonomy.

Table 6 – Revised Data Extraction Table

Identifiers			Extraction			Old. Info.	
Year	Title	URL or location	New Description of System	Does fit into current taxonomy?	Recommended Categorisation	Notes	Systems
2010	Example	url.com		Yes	Suggester		

2.9 Results

The information about the chosen personalisation systems was extracted and placed alongside other personalisation systems that match the chosen category (via 'document[~~ed~~] and tabular' means (EBSE, 2007)). For example, a theoretical system that matches a category in the taxonomy, such as suggesters, will be placed there. Systems that find no place within the taxonomy would have their own category, but none were located.

Using the information provided within the EBSE documentation, a Qualitative Synthesis was performed (Noblit & Hare, 1998 in EBSE, 2007). The form of this was the Line of Argument Synthesis (LOA) (Noblit and Hare, 1998 in EBSE). Adapting for the taxonomical use case, the LOA was performed as follows: the studies will be examined as 'individual[s]' in a specific taxonomically category and then 'as a whole' by judging the validity of the taxonomy of personalisation systems (Noblit and Hare, 1998 in EBSE). To aid in comprehension, tables are split into the different categories of personalisation systems. At points, works were linked together to aid with description.

2.9.1 Suggesters

The Suggesters category appeared to vary in use cases (Table 7). CiteSee uses data gathered from a user, which can be argued is a normal part of referencing academic work and uses it to format edits to 'inline citations' in a way that conveys information to the user (Chang et al., 2023). One example of this functionality is applied to those works that are 'previously opened' (Chang et al., 2023), being a different colour to those that are considered 'unexplored'. Costa and Duarte (2019) aim to capture patterns in user action data in a TV platform. These patterns identity to users who may be struggling with using the interface. This allows for the 'suggest[ion]' of potential changes to the interface (auditory and visual). WE-nner is a Connect IQ App for the Garmin platform (Martin & Clavel, 2018). As a user uses the application their 'activity' information is stored and analysed to use a bold formatting for one potential activity option in a menu. From these systems, it is clear that the Suggesters functionality is able to be used in a multitude of different environments. This may be due to the lack of intervention, resulting in the ability to have low potential consequences in various situations.

Table 7 – Suggesters systems identified from the systematic review

Source	Name of System *(System Description)*	Description of How System fits within the category
Chang et al., 2023	CiteSee *('prototype scientific paper reading tool')*	Reformats reference text to show different aspects of a 'user's reading history and paper library'.
Costa & Duarte, 2019	'Personalized and Accessible TV Interaction' *(Multi Device TV Interface)*	'user events are logged and analysed and adaptations are suggested to the user'.
Martin & Clavel, 2018	WE-nner *(Fitness Application)*	'consider[s] user's history of previous activities' to highlight one in particular as part of Swapper functionality.

2.9.2 Rearrangers

Often, Rearrangers are used in education situations, where content does not need to be removed (Table 8). PONI is a personalised training tool (Ashtari et al., 2022). It uses an inbuilt questionnaire to capture 'experience in programming and 3D modelling' to produce a 'results page'. The rearrangers functionality allows this page to be reordered based on a user's 'importance', which is a simple three-option selection. Maaliw (2021) present a range of improvements for Virtual Learning Environments (VLEs). By using an existing 'learning style' system linked to each individual user, it is possible for the proposed system to change what is displayed to users based on this data. The Rearrangers functionality is able to change the order of the materials displayed to users.

Hide-n-Seek (or Hide & Seek in the interface) de-personalises traditional 'web search' by utilising fake searches to the main search provider and then re-enables personalisation by keeping its own personal data set and 'reranking' the inaccurate results (R. Ahmed et al., 2019). To the user, their web searches will appear normal. Amazon Stream is a new take on e-commerce search (Teo et al., 2016). By manipulating 'category weights' with 'user click' data, the system can rearrange the options available to users. The use Rearrangers in 'search' situations is arguable one of the most standard examples.

Personalisation in user interfaces was able to be manipulated using a Rearranger approach. Silva *et al.* (2021) define multiple systems that can use personalisation in a similar way. While focusing on the technical element of personalisation which can be used across different systems, there are outputs for users discussed. Which type of user data is not explicitly stated, with general examples given, it appears that this data is part of the process to make different interfaces more efficient for users (by moving parts of the interface). FCT4U is a system developed by Madeira *et al.* (2014) to include multiple types of functionality. It takes usage data to determine which 'widgets' should be shown to a person on a 'large public display'. In this case, all the widgets appear to be present on the page (with some hidden by the length of the display), the arrangement is altered.

Table 8 – Rearrangers systems identified from the systematic review

Source	Name of System *(System Description)*	Description of How System fits within the category
Ashtari et al., 2022	PONI *('AR/VR' Project Search Tool)*	Rearranges content on 'results page' based on user-specified 'importance'.
Maaliw, 2021	'Personalized Virtual Learning Environment'	Using 'learning style' data, change the order of 'learning resources'.
Silva et al., 2021	'Dynamic User Interface Personalization' *(Used on multiple systems)*	Uses a single algorithmic approach to form new 'layouts' for users 'mobile' and 'web application[s]'.
P. Yu et al., 2018	Hide-n-Seek *('Web Search' add on and 'pop-up' menu)*	Takes 'search intents' to rearrange a 'web search' page.
Teo et al., 2016	Amazon Stream *(eCommerce Page)*	Changes the arrangement of eCommerce items based on 'user click' data.
Madeira et al., 2014	FCT4U *('Public Display')*	'widgets' presented in a list format have their arrangement changed based on usage data.

2.9.3 Swappers

Swappers systems have been used in leisure environments (Table 9). The SculptMate system is a personalised 'virtual tour' (Strousopoulos et al., 2023). The author grouped a series of art pieces into historical 'periods' and collected preferences for these as the personal data in this system. The system uses this information to change which piece of art is visually displayed to a user and 'audio guide'. The preferences are refined through a real-time feedback system of a similar style to that presented in Blackburn et al. (2023). Yang *et al.* (2022) developed an 'augmented reality'-based personalisation system. By taking a range of demographics and preferences for topics of interest, the work uses this data to display a range of factors to a user upon detection of famous landmarks (using Augmented Reality). The use of personalisation in these mobile applications shows that the technology is not limited to traditional use cases, such as Web Search (P. Yu et al., 2018).

The PONI system also includes an element of Swappers functionality (Ashtari et al., 2022). As aforementioned, the work uses Experience in Programming and 3D modelling to create results of potential projects that could be completed by a user of that level of experience. From the figures shown in the work, PONI presents a couple of options to participants based on this data, and thus, the projects displayed to participants are replaced by others in a Swapper system approach.

The Alma Assistant is a personalised 'Chatbot' which utilises 'self-reports of knowledge, prior difficulty in understanding and prior exposure to a reference' to change the way it responds to a user's text-based messages (An et al., 2021). This is generally in the form of changing a word or a short sequence of words for another, which matches the personal data of the user. The discussion will thus appear more natural or human-like.

As with Rearrangers, educational scenarios are a clear choice for personalisation and Swappers can play a role by replacing content. The MOOClet enables the inclusion of an 'if-then' paradigm into 'online course[s]' to add, in one 'use case', Swapper functionality (Reza et al., 2021). In practice, this is described as 'sending Version A to students with a low grade in the source, and Version B to students with a higher grade' (Reza et al., 2021). This system is open enough to use any type of information collected from a user. Maaliw's (2021) system, alongside its other functionality, adds the ability to swap content on Virtual Learning Environment pages with others that would aid in the learning of specific users. Pardos *et al.* (2017) provide a traditional 'recommendation interface'. The novelty is provided through the use of one of 'bucketed-time-input and normalized-time-input', which the system uses to attempt to form stronger recommendations in comparison to other algorithmic approaches. These recommendations are given to the user through the use of a 'hyperlink' on a 'MOOC'.

Educational settings can be affected by personalisation in another way. Rudian & Pinkwart, (2021) design a 'Moodle' platform that is custom to a user. By taking a user's 'performance' and comparing this to a target level of performance, a new 'learning path' of materials is presented to the user. This is repeated until the target performance is reached. As there is a fixed number of materials, this system is considered a Swapper, as the material is replaced with others with the intention of improving user performance. AdaptLearn is a Swapper system to create 'personalized learning paths' (Alshammari et al., 2016). This is a common form of personalisation and in this instance, the webpage sidebar changes to meet the personalisation algorithm ('Learning style and Learner knowledge'). Pavlich-Mariscal *et al.* (2015) demonstrate ASHYI-EDU, which 'can generate adaptive plans for each student, based on their specific competencies, skills, and learning styles.'. The plans appear to replace the learning materials with another. The data is collected from 'surveys completed by students'. These plans are updated dependant on current user performances. Each of these systems takes a personalisation Learning Path approach.

Schnabel *et al.* (2020) provide a different perspective, with a work that can be described as providing a significant demonstration of how systems within a single category can vary. However, it is important to note that these themselves are not new categories. The work presents a set of ways to change the 'visibility' of what a system is doing at a given time. For example, one may show the 'differences' between states of the personalisation system. The system described, for the most part, swapped in an out content on a "recommended for you'' container. Martin and Clavel's (2018) WE-nner also uses the aforementioned data to fill the screen of the device with 'activities' which are relevant to the user. This takes a similar approach to that of Schnabel *et al.* (2020) by swapping items into and out of a list.

As Swapper systems do not require large amounts of data, it opens the opportunity to make use of more novel data types to produce effective personalisation. Faria et al. (2019) aim to make use of a Montreal Cognitive Assessment and 'performance' metrics to provide the backing personal data for their Virtual Reality system. The performance data is generated by their use of the system, and this data is used to 'remove' user interface and in-world elements. Stohr et al. (2018) developed a multifaceted system 'for people with disabilities'. While the work does not clearly state how this data is captured, the use of 'NFC [Near Field Communication] for identification' provides an indication of how this personal data will form part of the personalisation. The Swapper functionality takes more than one form; for example, the work provides different formats for content, which are changed to suit the user. Biswas *et al.* (2014) present an underlying 'framework' that is used in multiple systems. This range of systems uses this framework to alter how information is displayed to users. It can achieve this by using metrics like 'grip strength' to understand how best an 'application' should be formatted. The information is the same in each variant of the format; just the visible appearance is changed.

Interfaces are a prime candidate for Swapper systems. Belo *et al.* (2016) describe a system that they believe 'rearrange[s]' the interface. However, there are 'six different segments' that can each contain a different type of 'data visualisation'. These are replaced by a personalised approach using the submitted 'query' information gathered from a user. The FCT4U system also presents Swapper functionality (Madeira et al., 2014). By taking 'transportType' it can change what is shown on one of the aforementioned Widgets. This is either 'road' or 'public transport information'. Usage data and 'whether the user is a teacher or a student' is used to present various options of 'greetings'.

Madeira et al., (2014) include an additional system called LEY, defined as a 'pervasive-based persuasive mobile serious game'. This gathers a set of 'interactions stream resources' to determine whether to display a 'user's contacts' list. The system also features a range of metrics that represent '[...]combats' and these are analysed to provide grouping into 'three cluster profiles'. This then alters the graphic that appears on an 'avatar'.

Table 9 - Swapper systems identified from the systematic review

**The original work references another source which is related to the system. For the review, only the information provided in the cited work was considered.*

Source	Name of System *(System Description)*	Description of How System fits within the category
Strousopoulos et al., 2023	SculptMate *('Sculpture' information 'mobile application')*	The system uses history preference to provide a series of '3D models' to match these preferences. The accompanying 'audio guide' is expanded for the preferred historical 'period'.
Ashtari et al., 2022	PONI *('AR/VR' Project Search Tool)*	Uses 'experience' to present potential project ideas to users.
F. Yang et al., 2022	'Personalised Information Retrieval for Touristic Attractions' *('Mobile AR application')*	Augments physical landmarks with additional information relevant to the user's interests.
An et al., 2021	Alma Assistant *('Chatbot')*	Replaces words in a 'chatbot' interaction such that it better represents the personal data of the user.
Rudian & Pinkwart, 2021	'Course Generating Engine' *('Adaptive online course')*	Forms new 'learning paths' for users based on 'performance'.
Reza et al., 2021	MOOClet *('Online course')*	Takes a 'learner data store' to control how an 'online course' reacts to users.
Maaliw, 2021	'Personalized Virtual Learning Environment'	Replaces one type of content for another to match their 'learning style'.

Schnabel et al., 2020	'News system with interactive recommendations'	Users 'liked' 'articles' and this affected a 'recommended for you' container, which changed which articles were displayed.
Faria et al., 2019	Reh@City* *('VR-based training')*	'remov[al]' of interface and world elements when users display high 'performance'.
Stohr et al., 2018	HRC UI	'Depending on the disability of the user, the HRC UI adapts its user interface by switching between different input and output modalities'
Martin & Clavel, 2018	WE-nner *(Fitness Application)*	Swaps menu interface options with 'activities' based on 'activity history'.
Pardos et al., 2017	'Recommendation interface' *(for 'MOOC', a type of VLE)*	Uses 'student's events' and 'time' to provide a 'recommended page' in the form of a 'hyperlink'.
Belo et al., 2016	'Personalization system for data-visualization'	Uses 'usage profile[s]' data to present relevant 'visualisation[s]' to users.
Alshammari et al., 2016	AdaptLearn* *(VLE)*	New 'learning paths' are displayed to a user using 'learning style and learner knowledge'.
Pavlich-Mariscal et al., 2015	ASHYI-EDU *(for a VLE)*	Lists of learning materials are replaced with others pre-interaction and in real-time based on 'surveys' and performance.
Biswas et al., 2014	'Interface personalisation web service'	Alters visuals of 'applications' based on 'visual acuity, colour blindness, short-term memory capacity, first language and dexterity level'.

Madeira et al., 2014	FCT4U *('Public Display')*	Swaps the 'maps' travel information based on the 'transportType'. Presents a different 'greeting' using a simplistic occupation status.
Madeira et al., 2014	LEY *(Application 'to help people understand domestic energy usage and change their habits')*	Shows and hides a 'contacts' list. Further, displays one of multiple icons with an 'avatar'. Both of these functionalities are based on personal data, which have a level of crossover.

2.9.4 Generators

In the reviewed works, a number of systems utilised elements from video games as the basis for their Generator systems (Table 10). Blackburn et al. (2023) present a system that forms a new 'Pac-Man map' after receiving a "'like" or "dislike"'. The use of personal data in this instance is defined in the work as the 'player's preference' and provides the stage for a wide range of different locations for 'pellets' (Blackburn, Gardone & Brown, 2023) (points-giving items). Thus, users can experience unique level designs, which would be difficult to replicate without this use of personalisation. A similar system has been created by Lyu and Bidarra (2023) which swaps Player Preferences for physical movement data. The work includes multiple 'games' which users complete to aid in increasing their range of movement. The personalisation component is the 'generation' of new 'levels' during these games. The use of computational games is interesting and provides a good platform for personalisation due to the defined start and end points for tasks. The previously described Virtual Reality system includes alternative functionality which is classed as a Generator (Faria et al., 2019). While not every one of the 'Tasks' users complete that are featured in the world created is described in depth by the authors, one utilised the aforementioned data types to create new levels that reflect 'difficulty'. As video games can require movement to new Levels (Lyu & Bidarra, 2023), there is potential -as shown in the works- to apply a Generator approach.

CiteSee (In addition to the Suggester functionality) creates new 'Paper Cards' for references in a paper; this displays additional information about the reference (termed 'context') (Chang et al., 2023). This is suggested to be related to the other work, such as showing 'the citing sentence where the cited paper was saved' (Chang et al., 2023).

Personalisation has also found use in input devices. Mitchell *et al.* (2022) attempt to provide a solution for users who have different requirements for ‘Keyboards’. Their solutions utilised a “Fitt’s Law-based point-select task”; from this, it was possible to form new keyboards that respond to their performance in this task. These keyboards are built uniquely for each individual. AdaptiKeyboard uses ‘performance measures’ to generate new keyboard layouts for users (Shen et al., 2022). This is a ‘mid-air virtual keyboard[s]’, which for personalisation means it is possible to, as is the case here, continually create new keyboards to improve their user performance. The final choice of keyboard is based on the preference of users. To note, the system is not fully functional as would be intended in a final product, but the work is functional enough to be considered a complete system for the purpose of this review.

EmpatheticSDS is a personalisation system that uses personal data to partake in two forms of ‘speech’ manipulation (Zepf et al., 2020). The system would take how a person talks, either in their choice of ‘words’ or ‘emotion’ and replicate this to match the user. These fall in the generators category as there are a large number of potential forms that the system needs to create in order to provide this replication of speech.

TAPrec is a system to ‘compose IF-THEN rules’, which are based on prior rules created by users (Corno et al., 2020). Using related parts of these prior rules, it is possible for the personalisation system to ‘recommend new rules’. The rules are used to control ‘IoT’ devices.

The work by Ni and Kehtamavaz (2023) highlights one of the key differentiators of the Generators type personalisation system. By gathering personal data in the form of ‘hearing preference’ taken from ‘paired audio comparisons’, the personalisation system can select from potentially limitless options (which is a perceived new option). Other types of personalisation systems, such as Swappers, generally have a fixed set of values to choose from. Realmuto et al. (2019) worked to develop a new variant of ‘ankle-foot prostheses’. By capturing how users are ‘walking’ without getting into the mechanics of moving physical artefacts, the system uses a ‘personalisation’ to adjust how the prostheses function. The HRC UI provides personalisation functionality in the form of ‘height’ adjustment (Stohr et al., 2018). Both of these systems use Generators to provide new values as the content rather than a more tangible artefact such as a game level (Blackburn, Gardone & Brown, 2023).

Another key use of Generators is communication. Martin and Clavel (2018) designed a system which takes a ‘user name’ and includes this in ‘message’ given to the user. Madeira *et al.*'s (2014) LEY uses a host of different personal data metrics and classifications to form ‘alerts’. These include a form of usage data, ‘competition level and user level’. As there is the potential -by using time data- to have a wide range of potential times for ‘alerts from the systems’, the system can be suggested to be a Generator. There is a 'persuasive[ness]' element put forward by the authors, but this is not explained in depth.

Table 10 - Generator systems identified from the systematic review

**The original work references another source which is related to the system. For the review, only the information provided in the cited work was considered.*

Source	Name of System *(System Description)*	Description of How System fits within the category
Blackburn et al., 2023	'Player-Centric Procedural Content Generation' *(for Pac-Man)*	Uses 'feedback' provided by the user to make new levels in a 'game'.
Chang et al., 2023	CiteSee *('prototype scientific paper reading tool')*	Includes a 'Paper Card' which displays reference links to other sources from 'reading history and paper library'.
Lyu & Bidarra, 2023	'Procedural Generation of challenges' *(for 'personalized gait rehabilitation' oriented 'game')*	Features a set of activities that users' need to perform physical movements, which vary in 'difficulty' based on this data and the opinion of a medical professional.
Ni & Kehtamavaz, 2023	ADRO *('Smartphone app for personalization of ADRO')*	Adjusts the sound to better suit the 'hearing preferences' of the user.
Mitchell et al., 2022	'Ability-based keyboard generation'	Takes 'movement data' for use in the creation of ergonomic 'keyboards'.
Shen et al., 2022	AdaptiKeyboard *('Mid-air gesture keyboard')*	Uses 'performance measures: speed and accuracy' to create new keyboards for users. At the end of the development period, the user 'select[s] their preferred design'.
Zepf et al., 2020	EmpatheticSDS *(used as a vehicle 'dashboard')*	Uses communication-based personal data to generate new ways of talking to the user to

		replicate their 'speech' types.
Corno et al., 2020	TAPrec *(interface (and 'platform') for setting up smart device automations)*	New ways for automation to control 'IoT' devices based on prior automation 'rules'.
Faria et al., 2019	Reh@City* *('VR-based training')*	Uses 'Cognitive Assessment' and 'performance' data to change the way a task is presented to users by forming new levels.
Realmuto et al., 2019	'Personalised Symmetry Learning Controller' *('Powered ankle-foot prostheses'*	Uses 'walking' 'data' to 'match the achieved ankle torque of the intact limb'.
Stohr et al., 2018	HRC UI	'the worktable was extended to change automatically in height [...]'
Martin & Clavel, 2018	WE-nner *(Fitness Application)*	Integrates personal data into 'personalized messages' for a user.
Madeira et al., 2014	LEY *(Application 'to help people understand domestic energy usage and change their habits')*	'personalize how often a user receives alerts from the system and how persuasive those alerts should be.'

2.9.5 Controllers

Controllers systems can take control of the task by completing them or they can control how users complete tasks (Table 11). The Personalized Virtual Learning Environment takes the latter approach by adjusting how users can 'access' these learning pages (Maaliw, 2021). Araujo *et al.* (2020) have developed a system to create 'study guide[s]'. Premade learning materials are linked to a user's performance score in a 'quiz', which then allows users to access (and not access) materials. The Personalised Adaptive Scheduling System is built into what is arguably a platform of two different systems related to learning (Xiong et al., 2015). Taking all three of these systems as a whole, there is the ability to 'set[~~ting~~] up personalised retention test schedules' for a user using their 'mastery speed of a skill'. In this instance, the system controls when users are able to complete the 'test[s]'. While not as technologically advanced as potentially autonomous systems, all of these three systems focus on blocking access for users when they have not met the prerequisites.

Stohr et al. (2018) also include a Controllers type personalisation in their HRC UI. By using 'cognitive impairments' information, in some instances, the HRC UI controls the flow of the task requiring additional 'user confirmation'. In this example, the system is adding steps for users to complete to ensure task quality based on potential risk.

Table 11 - Controllers systems identified from the systematic review

**The original work references another source which is related to the system. For the review, only the information provided in the cited work was considered.*

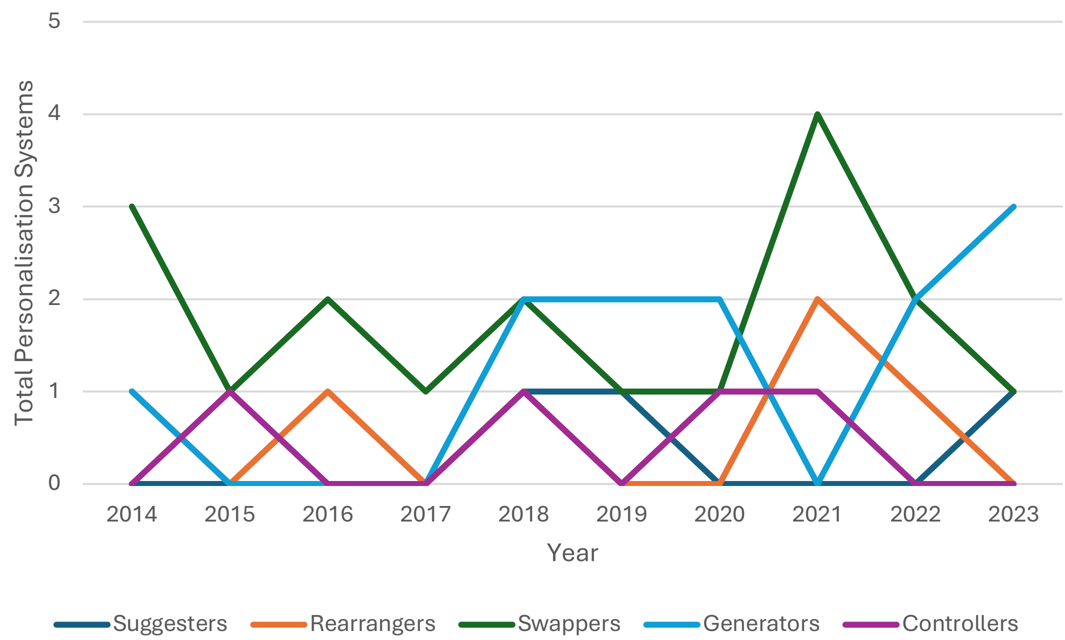
Source	Name of System *(System Description)*	Description of How System fits within the category
Maaliw, 2021	'Personalized Virtual Learning Environment'	'restrict[s] navigational access' to users to reflect their 'learning style'.
Araujo et al., 2020	FCTool* *('personalization of study guides' using 'Google Docs')*	Forms a list of material for learning based on a 'pretest'.
Stohr et al., 2018	HRC UI	Controls the tasks 'learn mode' based on the user's 'cognitive impairments'.
Xiong et al., 2015	Personalised Adaptive Scheduling System	Users are given access to educational tasks using 'personalised retention test schedules' with 'knowledge levels' as the input data type.

2.9.6 Validation

The reviewed literature presented a picture of the current state of personalisation systems within the majority of academic settings. Throughout, there appears to be themes of different types of personalisation systems. For example, in the Controllers category, there were systems which took a use case of online learning. These systems generally take similar approaches to how Personalisation is applied to users. This may signal an element of saturation within the user-facing side of personalisation. There may be a limit to how the technology can be applied, either practically (there are not enough ways to engage with users) or feasibly (the current algorithmic approaches are limited).

With the review taking a segment of almost a decade of academic work, while not the main analysis type (EBSE, 2007), the distribution of work into categories is of interest (Figure 6). It appears for most system types: Suggesters, Rearrangers, Swappers, and Controllers; the systems are scattered throughout without any clear trend. This provides insight that these systems may not be affected by any current trend or novel technological approach. From the current review Swappers seem to enjoy a high level of usage through the past decade. These systems appear popular; the reason could be down to the ease of implementation, or it is arguably down to the user experience. Users may prefer a limit to the amount of options being "swapped" rather than a list being rearranged. Generators take a sharp increase towards the latter half of the decade. As Large Language Models are currently enjoying mainstream coverage (BBC, 2023), the amount of Generator systems is likely to increase.

Figure 6 – Graph showing the distribution of personalisation systems in their taxonomical categories across approximately ten years



It appears that not all systems were categorizable into one type of personalisation system. This is understandable as to use personalisation practically, integrating multiple personalisation functionalities would allow interaction with users in various ways. The FCT4U (Madeira *et al.*, 2014) system was classified as both a Rearranger and a Swapper. This is not a fault of the Taxonomy, as a system can have both categories applied; the taxonomy still effectively describes what types of functionalities are present. The functionality is often separated; for example, in the WE-nner application (Martin & Clavel, 2018), the personalisation is presented to users one at a time. There is potential to apply multiple approaches at once, but the tiered nature of the taxonomy means that it should be possible to distinguish the functionality -even if splitting parts of the same process down is required- into a specific category.

The preliminary taxonomy can be considered valid due to the consideration of a wide range of personalisation systems. It was possible to place the personalisation systems featured in this work into their proposed categories. As aforementioned, due to the multiple functionalities being present, it was not always straightforward, but thus, possible with the inclusion of a decade up to the present. The taxonomy can be thought of as inclusive of systems with the current technological capabilities. Further, during the examination of the literature, systems that fit within the Generators category may not appear at first glance to be in this category. For example, the Personalised Symmetry Learning Controller (Realmuto et al., 2019) may appear not to be "creating" anything for the user. Instead, the argument could be made that this is just the adjustment of a parameter using personal data. However, what is being created is a new value for this parameter. Any system that appears to be adjusting values based on personal data is likely to be a Generator in nature.

2.10 Discussion

With the taxonomy validated, it is possible to compare this to other existing works, as this is not the first type of categorisation of personalisation systems (e.g. Blom, 2000). Nor is personalisation the only way of describing systems of similar functionality, such as the aforementioned Adaptive Interfaces (Benyon & Murray, 1988). While these may be widely used potential alternatives to the proposed taxonomy, the current work takes a different approach to how the classify systems by user interaction and data (Table 12). Thus, these existing approaches can be detailed to provide additional context to the novel taxonomy.

Table 12 – Common alternatives to personalisation which share related features

The different system types are described based on their utilisation of different data types. Internal personal data is captured by the system on specific users. External personal data could also refer to 'Big Data' (Asaithambi et al., 2021). World data refers to information about the physical world (e.g. Context; Dey, 2001). *System data refers to internally collected data about the computer system itself* (Haas & Hettinger, 2001).

System Name	Internal Personal Data	External Personal Data	World Data	System Data	Source
Adaptive Interfaces	Yes	No	Yes	Yes	Haas & Hettinger, 2001
Context Aware Systems	Potentially	No	Yes	No	Dey, 2001
Recommenders (a form of Personalisation)	Yes	Yes	No	No	Kim et al., 2020; Sundar & Marathe, 2010; Tiihonen & Felfernig, 2017
Customisation	No	No	No	No	Sundar & Marathe, 2010

A prior Taxonomy of Personalisation systems (Blom, 2000) provided a concept Degree of Initiation (DoE), with a separate discussion into system outputs. However, they stopped short of having strict categories for each system. The benefit of this approach allows for large levels of precision in terms of categorising systems, at the expense of ease of relating systems. In comparison to the Taxonomy of Web Personalization (Weinmann et al., 2013), the novel taxonomy provides a simplified viewpoint. By reducing the 'Type of Data' categories to feature as a secondary dimension along the taxonomies axis, it allows personalisation systems to be referred to as a single type, rather than as a series of individual factors. Further, the Taxonomy of Personalisation systems takes users' experience of the system as the main differentiator, whereas the prior taxonomy is arguably developer-focused.

One of the key links is between Recommenders and Personalisation. Recommenders are a type of personalisation (Tiihonen & Felfernig, 2017) and arguably, these should find a place within the taxonomy of personalisation systems. The reason for the lack of inclusion is the broad nature of the term, as when described (Tiihonen & Felfernig, 2017), more accurately reflects the technology underpinning found in Swappers and Rearrangers (and to a lesser extent, Suggesters and Generators). Further, existing work has also tried to distinguish between Recommenders and other types of systems (Montgomery & Smith, 2009). The authors provide a structure in which Personalized Search is differentiated from Personalized Recommenders (for media content), but arguably, these abstractly are very similar, the main difference being the usage case. The current taxonomy looks to place outcome as the main factor, in which the other elements (such as usage) can revolve and this should allow for many systems to fit within the five groups without creating multiple categories for many usage cases (as seen in Montgomery & Smith, 2009).

In the same mould as personalisation are Context-Aware Applications (Dey, 2001) and Adaptive Interfaces. Context-Aware Applications use data from the world in order to alter a system (Dey, 2001) (World Data). Dey (2001) refers to this data as Context, in which 'context is any information that can be used to characterize the situation of an entity'. This can be thought of as replacing the personal data from Personalisation systems with Context. It is uncertain whether personal data would be included as part of Context, but it does appear that wider external personal data (found in techniques like Collaborative Filtering (Kim et al., 2020)) are not included. In the current study, 'context' (World Data) was a part of certain systems (Madeira et al., 2014). Where this was not personal data this was not factored into their placement into the taxonomy. As this work is not aiming to classify context, these parts of the functionality were left unexplored.

The use of external data is, however, found within Adaptive systems (Benyon & Murray, 1988). Further, one can reasonably view Adaptive interfaces as building upon personalisation by including additional data types not found in the latter, such as 'the state of the system' (Haas & Hettinger, 2001) (System Data). With both personalisation and adaptive interfaces containing large amounts of crossover, it would not be surprising if personalisation systems include more traditionally adaptive interface data types as they become more complex.

2.11 Limitations

There were a number of limitations that may have affected the end results. Firstly, the choice of the SCOPUS platform. While this is a seminal platform for our subject (Hart, 2018, p266), the database is live and certain papers were added and removed over time. While this was believed not to have affected the results, certain papers may be later removed or updated which could affect how they would have been selected or analysed. For example, if a system contained -what would be classified as an inadequate description- a revision may correct this. A replication study may find papers that have been removed from SCOPUS in the time between events. The review method states a number of considerations for the resource limitations of the project ('Single Researcher' (EBSE, 2007)). Where possible these were followed, such as having a 'review' of the 'protocol'.

There is a further limitation in the search approach. The current thesis places an emphasis on manufacturing. As there are limited available resources on personalisation in manufacturing it would not be possible to form an effective review which would meet the requirement to develop an effective categorisation system. This is reflected in search keywords which do not contain topical manufacturing terms to avoid over-restricting the available results.

2.12 Conclusion

The current chapter set out a novel taxonomy and validated it through a systematic review on a current set of personalisation systems. The taxonomy classifies systems in one of five categories. These were decided upon by informally examining existing literature, taking a user-first viewpoint. Using a prior systematic review methodology, it was possible to examine literature from the SCOPUS platform. This review aimed to identify existing personalisation systems described in the literature. These were narrowed down to a relatively small number which were placed into the preliminary taxonomy. As there were no systems providing a reason to dispute the validity of the taxonomy, it became possible to suggest that the taxonomy is an accurate way to define different types of personalisation systems. The novel taxonomy was compared to other terminology or taxonomical approaches which looked at user interaction or data.

The final taxonomy allows for the categorisation of different types of personalisation systems by intervention and data usage. While prior taxonomies provide points in which to differentiate systems (e.g. Blom, 2000), they lacked distinct overarching categories. The novel taxonomy's categories enable research to distinctly define a system in a way that can be simply understood by others not familiar with the research but who understand the taxonomy. This is of benefit to the manufacturing sector who may want to integrate personalisation but become confused with the various systems. Without being able to see a clear difference between personalisation systems, it will be more difficult for a manufacturer to understand the potential value and integration proposition (e.g. how personalisation would fit within an existing workflow). Thus, the taxonomy has in impact in enabling potential implementers who may not have the time or experience to understand how personalisation systems can be used and how they differ.

The taxonomy contributes to the solving of the research questions. The new categories allows the research to take different types of personalisation system into consideration. The codesign workshops included these categories in the developed card set (Chapter 3: Codesign with the Personalisation Design Cards) which allowed the participants to develop their own personalisation systems. The Acceptance survey (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) included three types of personalisation examples in the second experiment, the taxonomy enabling this differentiation in a consistent way, increasing the validity of the study's approach. With the taxonomies categories lying on a linear plane, it was possible to include systems with different levels of intervention. The final empirical chapter (Chapter 5: Implementing Automated Assistance in a

Manufacturing Task) used the taxonomy as the inspiration behind the two personalisation systems featured.

Chapter 3: Codesign with the Personalisation Design Cards

3.1 Introduction

The first research question aims to answer how people would design their own personalisation systems. The obejectives of this work are to understand how users design personalisation systems within manufacturing environments, to use the created designs to learn what types of personalisation is acceptable and to develop a set of ideation cards (similar to prior work; Lucero and Arrasvuori, 2010).

To achieve the stated objectives, the work leans on a participatory design approach (Brandt et al., 2012; Muller, 1991), the aim of involving people in the design process will involve people discussing personalisation in a great level of depth. The findings from this approach will be used to form design insights for the implementation of personalisation. The insights from this approach will enable the understanding of potential user designs of personalisation and what design features they will be accepting of. Further, the current chapter details a participatory design study using a novel set of ideation cards as the method. These cards have been designed to reflect the varied personalisation system types from the taxonomy (Chapter 2: Taxonomy of Personalisation Systems) and the concerns around data and how this data is then used found within the motiving work (Chapter 1: Introduction) (Marinescu et al., 2022).This chapter is an expanded form of a published paper (Duvnjak et al., 2024).

Personalisation's technical components are often hidden from the perspective of the user, only seeing the final output (a common criticism of many types of automated systems). Personalisation systems rely on users giving away data and these users may be more inclined to want to understand how this data may be being used (Sailaja et al., 2019). This is brought to the forefront with existing work having described users who don't want to engage with these types of systems also being users who want to understand how data is used (Awad & Krishnan, 2006). Thus, there is the question of how and when to display information about an automation's process to a user. It is doubtful that many users should be made to read an instruction manual explaining the complexities of automated systems before using a system (D. A. Norman, 1998, p71). A further problem with the lack of transparency of systems is when problems occur they can be difficult to correct and cause issues for the end user (Norman, 1990; Greenfield, 2006, p147).

The problem around transparency could be related to the problems around the explanation of systems (as described by Sailaja et al., 2019). Potential solutions could be built into the structure of computer systems (e.g. Pawar et al., 2020)) or in the provision of extensive documentation, as suggested by Norman (1998). However, these may not be the correct approachs, as these systems are often created by system experts, they may not understand the transparency issues in the same way as a standard user. With the additional apprehension around DMTs (Marinescu et al., 2022), this chapter proposes to resolve these issues through the involvement of users in the design process. Participatory design was chosen due to its links to system design (Muller, 1991). The theses' first research question of the thesis captures this notion: How would co-design be used to design personalisation within automated manufacturing environments?

The research question was broken down into aims. These aims examined how people design personalisation within automated manufacturing environments, how accepting of personalisation are end-users and the development of a set of cards for participatory design (similar to works such as Lucero and Arrasvuori, 2010), which would allow systems designers to understand the needs of their personalisation system users. To achieve these aims, a set of ideation cards were developed (the Personalisation Design Cards) with insight from the Taxonomy of Personalisation Systems (Chapter 2: Taxonomy of Personalisation Systems) and these were used by end users to understand personalisation systems and allowed insights to be generated into how end-users would design personalisation systems to their needs. The usage in the co-design workshops, in turn, aided in the redesign of the personalisation cards and validated their effectiveness as research tools. This chapter is an expanded form of a published paper (Duvnjak et al., 2024).

3.2 Literature Review

3.2.1 Participatory Design

Participatory Design is a set of methods that 'explore the conditions for user participation in the design and introduction of computer-based systems at work' (Kensing & Blomberg, 1998). Before discussing the chosen participatory design approach, a brief review of current thought into this area will be completed. At a high level, participatory design often leans heavily on being an iterative process defined by Brandt et al.'s (2012) Tell-Make-Enact process. In overview, this process involves the discussion between participants (tell), the creation of artefacts (make) and the use of said artefacts (enact).

Achieving a complete participatory design process can be done in multiple ways, although they follow similar general principles. The foundation has been set by the PICTIVE method (Muller, 1991). The aim of this method is defined as 'the creation of the design of the interface' and involves groups of people working in tandem to produce low-fidelity (lo-fi) prototypes (using the termed Design Items) (Muller, 1991). These basic principles form the blueprint of other participatory design methods. This can be seen in the study by Yao et al. (2019), in which groups of people were asked to design ways to resolve potential problems in a Smart Home using Creation Tools (almost identically to Design Items in the PICTIVE method). While the target and procedure of the participatory design process is different, the key elements of participatory design, being the lo-fi design of systems, remain.

3.2.2 Ideation Card Approach

Ideation Cards are a research method in which people can utilise cards to generate insights into the topic of the cards (or activities). As one of the key parts of participatory design is an Iterative Process (Spinuzzi, 2005), ideation cards excel at being easily made repeatable allowing for strong insights. As with most participatory design methods, ideation cards are often used alongside Muller's Design Objects (1991) to be at their most effective. The cards themselves are often themed on a particular topic, such as Mixed Reality (Wetzel et al., 2017).

The way in which ideation cards are used differs between sets. Halskov and Dalsgård (2006) have a comparatively unstructured approach to their card rules. After a discussion of each card, they suggest an open approach in which participants can use any cards in the design phase. In contrast, Wetzel et al. (2017) utilise their cards in a more traditional game sense. The authors present various sets of 'game rules' which can change the way the cards are used but still provide a structure for participants to follow. This is to say neither approach of utilising their respective cards is invalid, but it would vary depending on one's required research outcomes. A more free-flowing approach could provide a greater outlet for creativity, while a more structured approach could keep participants on track and provide more relevant design outcomes (if one had a niche subject area). This flexibility is one of the key reasons the cards are such useful research tools.

Ideation cards are designed in a number of ways. Lucero and Arrasvuori (2010) present a set of PLEX cards (the authors describe the name PLEX as based on a prior work's 'Pleasurable Experiences Framework'). The authors forgo the usual amount of text found in other card designs (for example, Wetzel et al. 2017) for short, simple descriptors. In contrast, the opposite is found within the card deck by Shinohara et al. (2020). As one of the aspects of the cards is to get people to ‘reflect’ on the findings of prior work (K. Shinohara et al., 2020), the cards provide a comparatively large amount of text on both sides of the cards to achieve this. It is of no surprise to suggest that each card set is developed with a specific focus and that this focus is represented in the design of the ideation cards.

3.3 Method

The study uses participatory design approach as this would that enables users to design their own systems (Brandt et al., 2012; Muller, 1991). This should enable insights into how personalisation systems wish to be designed in terms of features and functionality. To facilitate participatory design, group workshops were held that made use of a novel set of ideation cards, the Personalisation Design Cards (PDCs). As the personalisation systems in manufacturing may reach beyond interfaces, participatory design methods that focus on interface design, like PICTIVE (Muller, 1991) may not generate the aimed range of insights. Ideation cards do not have such limitations. However, as ideation cards are often themed (e.g. Mixed Reality Game (Wetzel et al., 2017)), a new set of cards was developed. These new cards use various data types that could be used in personalisation. The ethics documentation has been made available (Appendix B – Chapter Three Ethics Documentation).

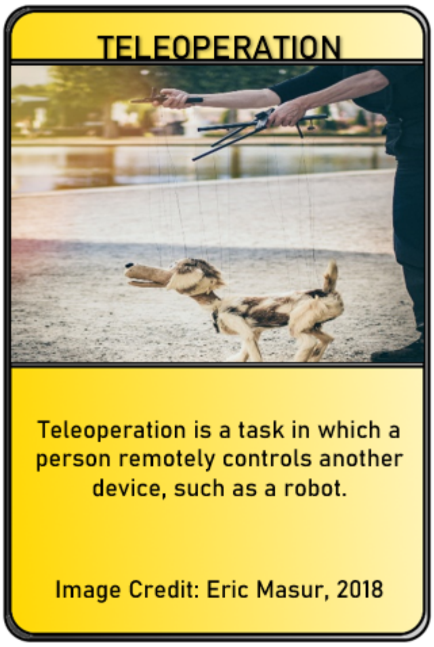
3.3.1 Initial Development of the Personalisation Design Cards

The study’s design rests upon the development of a novel set of Ideation Cards (PDCs) (Appendix C – Original Personalisation Design Cards). The PDCs adapt the Mixed Reality Game’s (MRG) Opportunity, Question and Challenge cards (Wetzel, Rodden and Benford, 2017) into Task, Automation and Data. The Data cards describe personal data types that can be used by participants, similar to the Opportunity MRG cards, which provide additional framing to the game. Some were also chosen to represent more controversial data types (Marinescu et al., 2022). MRG’s question cards were removed, with Task and Automation being MRG-style Challenge cards. This was to create a foundation for the participants in which to build their system without overloading them with too many potential systems designs, in line with common participatory design theory (Sanders and Stappers, 2008). The Automation cards are based on an early taxonomy, which indicates there may be five categories of personalisation systems (Chapter 2: Taxonomy of Personalisation Systems). This was to allow participants to design different kinds of systems and understand how these types of systems differ in practice. The Task cards attempt to represent specific fields in the manufacturing industry while simultaneously are open enough to allow different interpretations of what the user could be doing in each situation.

These cards were based loosely on the MRG Cards designed by Wetzel, Rodden and Benford (2017) to avoid replicating prior card design work. This is notable in the design, in which both sets of cards feature a coloured border to represent the type of card, a title and a description and an image is included on both sets of cards. Wetzel, Rodden and Benford (2017) describe the image in their MRG cards as "an additional source of inspiration". As participants may not understand personalisation terminology, the card's image can clarify what the descriptive words cannot (Figure 7). The cards are designed to be printed on regular paper, as per informal discussion with a developer of MRG Cards, providing the rough appearing 'prototype' cards should visually afford (Norman, 2013, p11) participants the capacity to edit the cards. A similar idea is found in the cards designed by (Halskov and Dalsgård, 2006) that includes an "empty box for comments ".

Figure 7 – Example of an Automation-Type Personalisation Design Card

The 'teleoperation' card's image features a puppet dog, showing a person operating the dog. This will help clarify the card to potential participants who may not be familiar with the concept or term of teleoperation.



3.3.2 Participants

In total, ten participants were utilised during the study, split into three groups (in a 3-3-4 split). These were mostly female (F = 7, M = 2, No Entry = 1), were majority between the ages of 25-34 (34-45 = 2, 25-34 = 6, 18-24 = 1) and mostly had master's degrees (master's degree = 8, bachelor's degree = 1, PhD = 1). They were recruited through email, email lists at the University of Nottingham, word of mouth (including physical and digital variants) or both. Thus, it can be implied that all participants were associated in some format with the University of Nottingham. The sample and group assignments were based on convenience and participant availability. Participants were provided with an honorarium in the form of a £20 Amazon Voucher upon the studies competition.

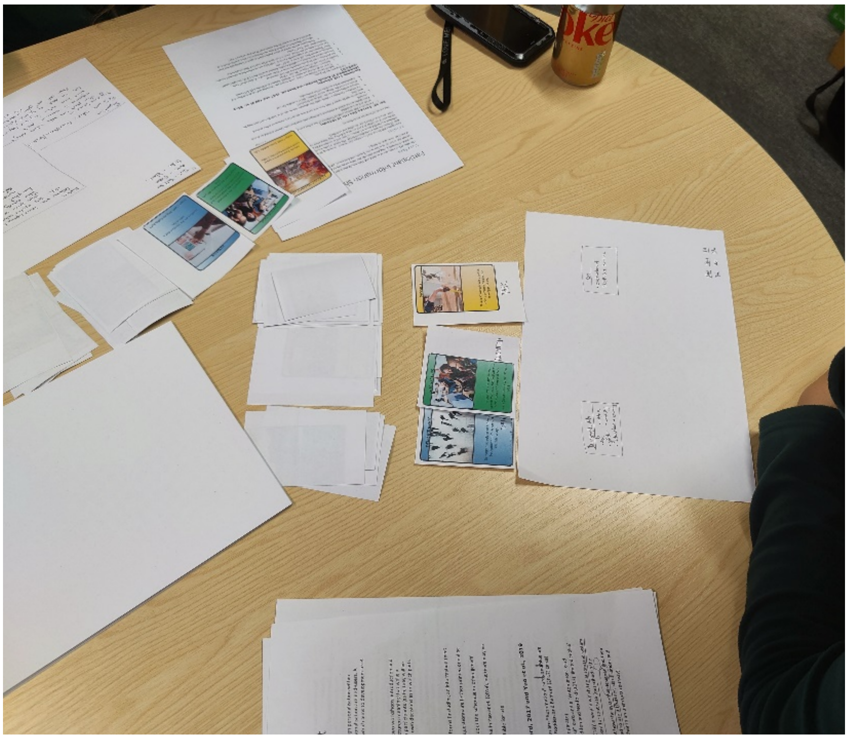
The amount of participants required for qualitative research varies depending on the study type and analysis method (S. K. Ahmed, 2025), thus it is difficult to draw a conclusive “ideal” number of participants. In comparison to other types of reach methodologies, ten participants may be suggested is too low for valid data. However, other studies using a similar approach show that small numbers of participants are not unusual. For example, Demirbilek and Demirkan (2004) had thirteen participants and Ahmed *et al.* (2019) had nine. There is a discussion as to whether more participants would produce greater levels of results, but it did appear that themes were identified across groups. It could be argued that with more participants, themes would be repeated further rather than generating new themes.

3.3.3 Materials

The workshops used printed design ideation cards as the stimuli, one set for each participant. The provided Muller’s (1991) Design Objects were similar to the original work and included basic stationery: pens and paper. These were communal for any participant to use and placed in the centre of the table where the participants were based. The workshops took place in rooms with no unrelated people present. Participants generally sat at a table or a combination of tables to promote discussion (Figure 8). Video equipment was used to capture the workshops visually (a video camera and/or smartphone) and audio equipment auditorily (either the laptop connected to a microphone or a laptop’s built-in microphone). Demographic questionnaires and ethics forms were printed and available to participants during the study. The original and revised cards (created after the workshops, with Game Mat) and rules are available online.

Figure 8 – Cropped Image showing the general physical set-up of an in-progress workshop

The picture shows the participants using the cards in conversational proximity, although in the workshops, participants were not always directly facing others.



3.3.4 Study Design

Participants were arranged into three groups. Each of the participants had the same three sets of cards. The participants were allowed to work individually or alone within their groups. The study was split into five sections and participants took part in all: a short introduction, workshop one, a short break, workshop two and a short finishing section. The researcher was present during the workshops but only existed to guide and answer questions in the event of issues. This was to avoid the researcher having an effect on what the participants designed during the study (as with 'the complete participant' Robson, 2011, p320).

3.3.5 Participant Procedure

In summary, a workshop consisted of two design sessions separated by a break. At the workshop, participants were requested to sit at a table with other participants. They were provided printed ethics documentation and demographic surveys to complete alongside the card sets. Participants were informed about the contents of the workshops and data collection and allowed to ask questions. They were provided the following game rules:

1. Shuffle all three sets of cards (and keep the sets separate)

2. The participant will be given/asked to draw a "Task" card and a "Automation" card
3. The participant will follow a modified Wetzel et al. (2017) 'limited choice' method, in which they will draw a "Data" card
4. Then 'brainstorm' (as suggested by Yao et al. (2019)) a way in which the automated system can use the data (in the "Data" card) in that system for that task (described in the "Automation" and "Task" cards)
    a. Participants will be asked to think about how the system is designed and their attitudes towards this system if they had to use it (Further, clarifications and questions may be asked by the researcher in an unstructured way)
5. They will then draw another "Data" card and repeating step 4, but with two "Data" cards to consider
6. They will then draw another "Data" card and repeating step 4, reaching the 'three [...] cards' described in Wetzel, Rodden and Benford (2017) and consider them all
7. The participant should have three data infused systems, A discussion should be had about their ideas where participants 'tell' (Brandt et al. 2012) their ideas to other participants and the researcher.
    a. Other participants and the researcher are allowed to ask questions about their ideas in an unstructured way

When the workshop formally started (and the participants started to play the game with the PDCs), participants had flexibility in how they wished to participate but an effort was made to make sure every participant completed the game at least once. Once an approximately 30-minute period had been completed, a short break was held (signifying the end of the first session). Images were taken of the participants' paper systems and the cards (except in one case where images were taken shortly into the second session). The participants then restarted the game with new sets of cards for the second session. The workshop was then completed and participants were allowed to leave.

3.4 Results

3.4.1 Thematic Analysis

Participant data was transcribed automatically with the University of Nottingham's Automated Transcription Service which meets the relevant rules regarding data. Manual corrections to parts deemed relevant (Table 13). The analysis method was an adapted form of the Braun and Clarke's (2006) Thematic analysis, which had elements adjusted, rearranged, or removed. The final analysis follows a similar structure to the stated approach (Braun & Clarke, 2006): transcribe data, code data, devise themes, refine into main and sub themes, write report and adjust theme titles. Of the changes, the most notable were the lack of thematic map arrangement (instead opting for a list arrangement). Themes that were replicated in different workshops were considered of greater prevalence. The quotes/transcripts presented herein will have certain sections removed or clarified to retain reading fluidity. There is a discussion as to whether a greater size of population would produce greater levels of results, but it did appear like there was a level of repetition of insights from the participants (and thus, themes were identified across groups). A theme is a common topic referred to by participants, what constitutes a theme is decided by the researcher and has little defined quantitative value. The current work is suggested to comply with the Process Transparency and Data Sharing from SIGCHI communities "Transparency in Qualitative Research" (Talkad Sukumar et al., 2020) by the provision of significant participant quotes and a detailing of the analysis approach and how this was altered from the original method.

Table 13 – A tabular representation of the thematic analysis

The main themes are present in the leftmost column, with participants' quotes available in the rightmost column.

Main Themes	Sub Themes	Participant Quotes
'Dynamic' Systems	Dynamically Different Systems based on Experience	Participant 1: And (...) data is task experience, So I thought we have like low to high task experience and at low task experience the controllers have more control to make decisions. [...] Maybe a lower threshold of risk so the controllers can decide if a robot is about to make a mistake too quickly, override it or stop it.
		Participant 5: Make suggestions on what kind of interface do you want? Do You want like a simple and minimalistic interface for someone who is a beginner. Or do you want a more like advanced user interface for people who are more experienced. So that could be used like this and. To [...] adapt the interface. [...]
	'Show' Shortcuts	Participant 7: Maybe you like show shortcuts to people who are experienced as he uses a lot.

		Participant 3: [...] Shortcuts guess suggesting shortcuts for an (...) would be so thing that came in mind, like Python, Python, so much more easier to use because you just couldn't tap and there's so much more. Suggestions and sometimes you can just like send you the code is not like perfect and it still works based on. Yeah reading on your profile of how how expert you are. [...]
	Dynamically different systems based on Demographics	Participant 6: [...] For example, let's say that you have people of different nationalities and they come to work and they have instructions. So you could give people instructions in their own language. Participant 3: And if you're too old, you like 65 and above, maybe they were just like linked to local handyman or like link you to specialists rather than doing it yourself. So that was demographic data also because thinking about HCI, so maybe education level and linking into more text based or more graphic based, but also implication for that [...]
System Usage Concerns	The Use of Certain Data Types	Participant 1: And I decided to just ignore that because I think it's a big insidious to measure their person's heart rate in terms of this kind of task. I mean, maybe I thought maybe it can, maybe there's a threshold at very high heart rate, it shows that the supervisor is too overwhelmed and maybe the whole the controller overides completely. But I'm not. I don't know enough about this to say that this is an okay thing to do. I think [...] maybe we shouldn't measure people's body data for just this kind of work. Participant 9: its so intrusive
	Accuracy of Conclusions Drawn from Data	Participant 5: I'm sure I feel like maybe fatigue and performance maybe not necessarily linked because I would I would feel like an older person maybe being more tired to do the task but they would probably do it faster because of experience. [...]

		Participant 9: It could be that it should be X amount of difficult. It should take X amount of brain power to do. But if that person is tired, hungry, battling something at home, distracted by someone they fancy sat behind them, like all of these things that are unique to an individual that you can't measure their impact, perhaps it's just too, too broad, like maybe that needs.
	Compensation	Participant 4: sort of bonus or a quota today, 80% of your time was on the difficult task. Therefore your salary this month or for this day's higher or something. that. Then again leads would probably lead to the fact that people were doing everything they can not to seem fatigued because that is the impact on. I mean it would be pretty wild. We come to the office and at the end of the month, every month our boss will tell us, well this month the salary is only 70% of what its normally (...) I thought you were tired a little today I mean. Participant 9: if you are a hgV driver and you are working all hours of the week to get the money that you need and you're working overtime and you're actually really tired and you're really stressed at home. Are you gonna say or any job really? Are you gonna say I'm too tired to do this today for your boss to turn around and say, okay, okay, you don't do it, you don't get paid?

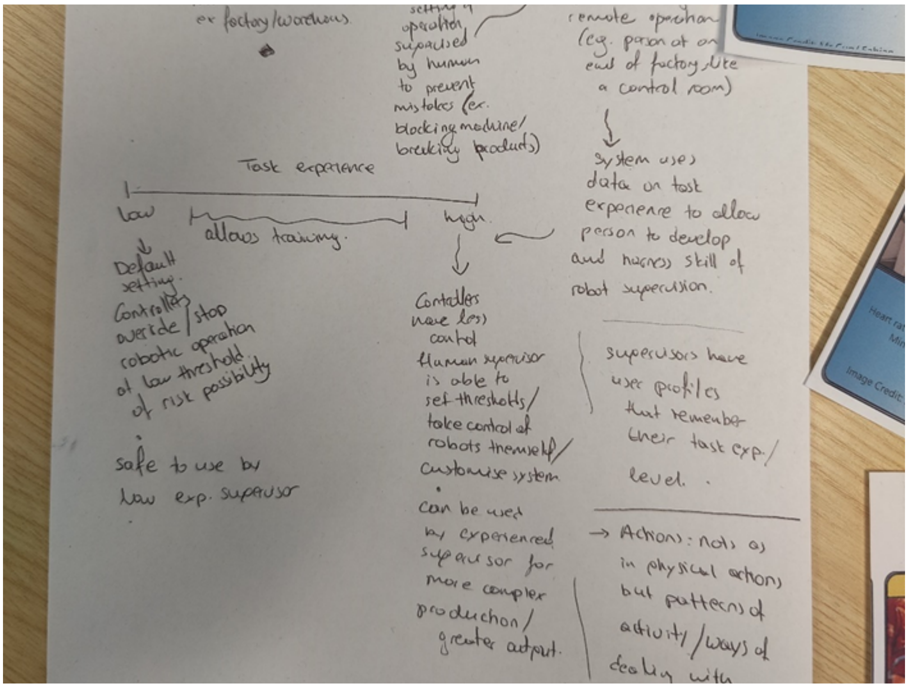
3.4.1.1 Theme: "Dynamic" Systems

The first theme located was that of 'dynamic' systems. The word 'dynamic' comes from a participant (4) (although their system does not partake in this theme). Dynamic in this analysis refers to how systems can create 'shortcuts' or utilise experience/demographic data as part of a changing personalisation process.

The first sub-theme identified was that of Dynamic systems based on user experience. From a participant's (1) drawn material (Figure 9), it is possible to see an example of this approach. In this system, a user is identified to have a level of experience with a system and in turn, alters how much control the automation has over the system. In instances of 'high' experience, users have more control over the system than those with 'low' experience. The participant states how this could make 'low' users safer while also allowing 'high' users to complete more complex work. A quote from a different participant (5): 'Do You want like a simple and minimalistic interface for someone who is a beginner' describes a very similar system. The former uses the 'task experience' card as the data type, where this uses assumably general experience. It appears that experience in various forms must be accepted as a form of usable data by the participants. The systems described could be classed as Swapper systems, which completely change the system for the user.

Figure 9 – A cropped image of a participant's (1) system

The top centre of the image shows a line labelled with low and high task experience. The user has more control in instances of high control and less in instances of low control. Training is allowed for users falling within a middle range of task experience.



The second identified sub-theme was focused on systems that 'show' shortcuts. While only briefly mentioned by participants (3 and 7), there seems to be an idea that you can 'show shortcuts to people who are experienced' (Participant 7). In what format this would take is not directly specified by the participants. There are existing systems which utilise a similar approach. For example, a participant (3) refers to the Python programming language, which may link to the idea of shortcuts but there is no explicit use for their system. The academic literature does talk about these ideas; for example, Billsus et al. (2002) describe an approach for 'mobile technology' which rearranges options to order choices by 'most frequently accessed'. Despite the lack of concrete use case of shortcuts, this sub-theme shows functionality that users may accept in systems.

The final sub-theme identified within the theme of Dynamic Systems was the use of demographic data. One participant said, '[...] people of different nationalities and they come to work and they have instructions. So you could give people instructions in their own language' (Participant 6). This is a clear use of demographic data (the nationality of the worker) and leads to an effect of changing the interface's language. This would be a Swapper-type system, and this is very similar to another participant (3), which had the idea of age data 'linking into more text-based or more graphic based' along with another type of functionality. From this sub-theme, we can assume that demographic data and interface-based swapping type systems are the types of systems users may wish to see.

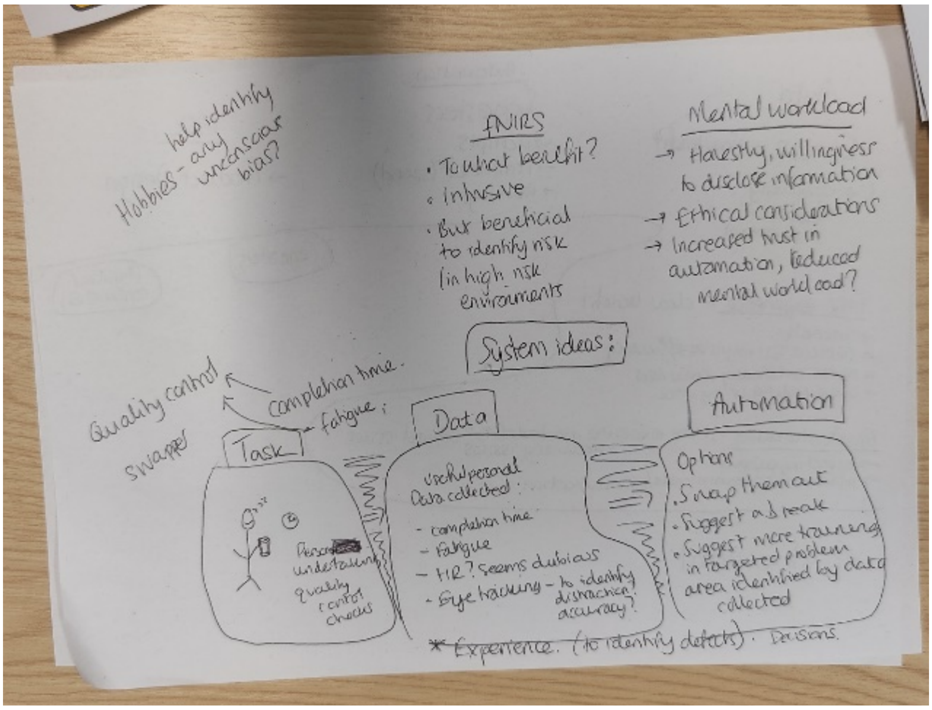
3.4.1.2 Theme: System Usage Concerns

The next identified theme was the concerns around the usage of these systems. The sub-themes brand out this in categories of concerns. These were concerns around the use of certain data types (like Data Usage in Marinescu et al., 2022), the accuracy of conclusions drawn from data and how these systems could affect compensation.

The first identified sub-theme was the concern around the usage of certain data types. Participants (1 and 9, respectively) referred to certain data types as 'insidious' and 'intrusive'. This was in response to the data types: heart rate and fNIRS (the latter is shown in Figure 10). These data types were included to provoke this kind of thought, as previous work indicated that 'personal data' was a theme in which one participant was concerned about certain data (Marinescu et al., 2022). The response from some of these participants clearly show concern, even in a card game situation, which it could be suggested has a bias towards being "light-hearted". In both cases, the participants did try to find 'beneficial' (Figure 10) aspects towards the data types.

Figure 10 – Written content from a participant (9)

The centre heading shows fNIRS and the participant's thoughts relating to this data type.



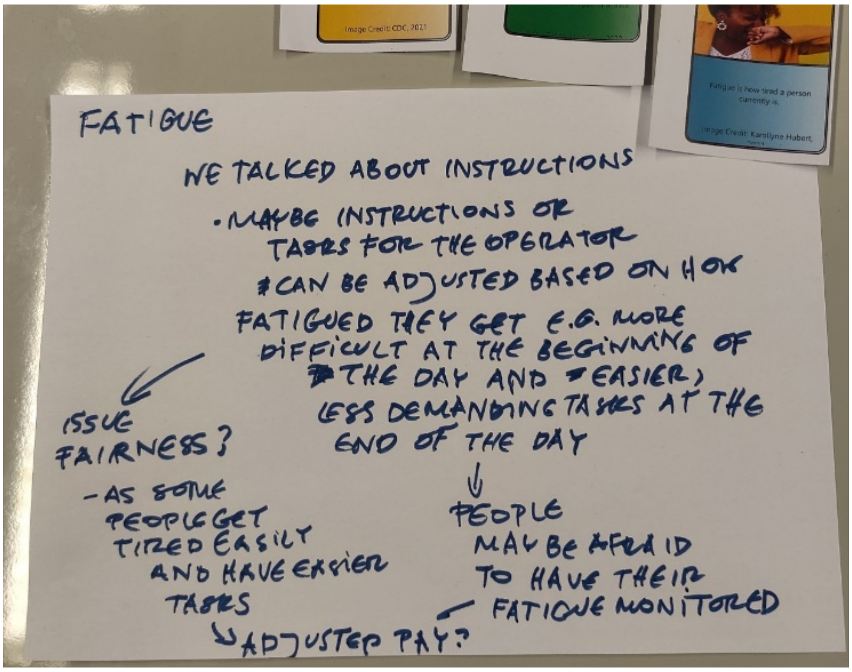
The next identified sub-theme related to the accuracy of conclusions drawn from the data. This can be best seen from a quote from a participant in discussion with other group members about how pay can be connected to fatigue data:

> 'I'm sure I feel like maybe fatigue and performance maybe not necessarily linked because I would I would feel like an older person maybe being more tired to do the task but they would probably do it faster because of experience.' (Participant 5)

From the quote, it is possible to understand the general idea. Participants had thoughts that the data may not connect to the conclusions drawn (from the context, this could be an effect on pay) from the fatigue data. This could indicate a distrust of data analysis on the part of the participant, and this could be a wider factor as the other selected quote suggests a similar idea but refers to how 'unique' (Participant 9) aspects can affect the data and may lead to incorrect conclusions being drawn.

The last sub-theme was that of compensation. As aforementioned, participants raised points about 'worker pay' (Figure 11) and how systems like the ones described in the workshops could affect this. One of the examples provided by a participant (4), which suggested that tasks could be allocated based on difficulty (and difficult tasks pay more) and connected fatigue data, further noting that if you were provided low-difficulty tasks, you would receive less pay for this. A different case can be seen from a different participant (9) who said, '[...] Are you gonna say I'm too tired to do this today for your boss to turn around and say, okay, okay, you don't do it, you don't get paid?'. From this quote, it is possible to see a different side of the argument, perhaps suggesting that rather than employers allocating tasks, users can utilise their data to talk to their employers. They then suggest that an employer may not wish to pay for unworked hours. This perhaps suggests a need for worker rights that could address this imbalance.

Figure 11 – Written content from a participant (6) showing how fatigue can be used and the potential downsides of this approach



3.4.2 Workshop Observations

In the running of the workshops, there are a number of findings -while not directly fitting into the thematic analysis- provides great insight into the use of ideation cards for creating personalisation systems for automated manufacturing environments. The insights from this section relate to observations and material from the workshops.

3.4.2.1 Card Design

How participants used the cards can have an impact on how future cards should be designed. As ideation cards have previously been used in different contexts (such as Mixed Reality by Wetzel et al., 2017), it was interesting to see how they would function in a situation with various factors to consider. With this large amount to mentally process for participants, it is best to keep the cards as clear as possible. The current cards have an identified weakness in that the type of each card was not clearly displayed. There was writing on some of the cards to indicate the type (such as 'data') with another approach being to create a "mat" which outlines the type of cards in a deck (

Figure 12).

Figure 12 – One of the participants produced a mat to show which cards belonged to each type

The mat would avoid users having to remember which colour of the card was associated with each type; this is an idea which can be incorporated into future card designs.

The cards were as intentionally made out of paper to allow participants to edit them in the workshops. The participants were informed that they were allowed to change the cards but only some participants made use of this. A participant (3) added the card type to the card. This may link with the above notion about the need to better clarify the type of card. One participant (1) did change two of the personal data cards slightly, but only small edits to the card itself. These may be indicators that the personal data card designs did not need much adjustment. It was mentioned by two participants (1, 4) that the image for the Controllers cards image may be incorrect. This highlights the importance of making sure the images have the same meaning for the researcher versus what meaning the participant will take from a first impression.

3.4.2.2 Game rules

The game rules are a point of discussion as the three groups took slightly different approaches. From the researcher's perspective, there seemed to be an element of groups wishing to work together rather than independently. This may be due to nerves, especially when given a task which can appear complex such as designing an automated system with a game the participants had not seen before. One group followed the rules the most closely by playing the game, which led to a discussion and then a quick break, which was followed by another round of the game and another discussion round. One group had two game rounds and two discussion ones in one block but had a strong collaborative presence. The final group worked almost entirely as a group and loosely followed the game rules until the researcher requested them to complete the game round independently towards the end of the second session. It would be difficult to argue that any one method of completing the workshop sessions was more or less successful.

3.5 Discussion

3.5.1 Designing Personalisation-based Automated Systems

From the workshops, one of the prominent findings was the using of experience to adjust automated systems. This presents two points of discussion, the first is that potential users may be accepting of using experience as a personal data type in analysis at a workplace. There could be many reasons for this, perhaps users feel as if data collected while at work is less personal. A existing study found that users may be concerned about 'out of work' data capture (Marinescu et al., 2022). The systems designed by our participants would not fall into either category and this may explain why these systems appear frequently. The concept of using experience is not novel. The second discussion point relates to the use of experience as a metric; an existing work has examined 'expert users' vs 'inexpert users', finding that the former would want more involved automation in comparison to the latter (Schiaffino & Amandi, 2004). This is in contrast to the systems designed by our participants, which suggested that high-experience users would receive a less involved system. The concept of new users needing more help with a task is not unreasonable. One potential missing factor is the conflicting nature between trust in automation and experience. The idea that trust in automation would have the effect similar to described by Schiaffino and Amandi (2004) was discussed by participants (3, 7). Thus, systems designers should be aware of this type of interaction when designing automated systems, alongside making sure to use personal data types which you could expect to find at a place of work currently.

An insight from the workshop was the concern around the accuracy of certain personalisation analyses. This is important, as correctly pointed out by participants could affect your perception in your place of work; in the case of the participants, it is related to compensation from your employer. The type of inaccuracy can be identified as being similar to the concept of 'distributive injustice' (Yeung, 2018). Although referring to "customers" rather than employees, the work echoes present the idea that some people will be discriminated against based on 'a commercially rational form of social sorting' (Yeung, 2018). This, coupled with the fact Yeung (2018) suggests systems can be ‘opaque’ it can thus be seen that participants have reasons to be concerned. The identified concerns could also be in a similar vein to the problems with ‘categorisation’ (Monzer et al., 2020). These are potential errors in data analysis leading to negatives for users (Monzer et al., 2020). In the current study, a suggestion from a participant (4) about the potential for ‘unique’ data points to skew the data. If participants cannot see their data or the analysis, they will be unable to know what data point has negatively affected their standing in the place of work.

There is a connection to be made between the identified concern of compensation and the concept of ‘value trade-off’ (Sailaja et al., 2019). Participants suggested that if certain data was used, it could mean employers would be able to reduce or stop paying employees for shifts, they are too tired during or to complete. The potential here is that following existing logic in value trade-off-like situations (Awad & Krishnan, 2006; Sailaja et al., 2019), employees will not want to give away personal data that will bring them lost wages (negative value). Further, there is existing literature that shows many people would put a high value on the type of personal data in a similar category as discussed by participants (Skatova et al., 2013). These factors build into the idea that there may be certain types of data in which the value trade-off can never be fully reached in the current work climate. This is pointed out by workers who discuss potential employer reactions to situations that arise from the data.

Data privacy for certain data types is a topic that participants referred to in the negative during the current study. This is not a new phenomenon, with a survey study finding that a minority are unwilling to provide all their ‘behavioural data for personalisation’ (Yamamoto & Yamamoto, 2020). Further, a related study found concerns with the use of ‘biometrically monitoring’ techniques (Marinescu et al., 2022), which was replicated in the current study in relation to the use of heart rate data. There appears to be a line in which people will feel as if a data type is acceptable to be used and when it is not. Where users draw the line can only be clearly defined at the extremes, for example what is widely accepted and what is not as the case with task experience and heart rate data. The uncertainty lies in the middle ground of when a person feel that the data capture crosses into unacceptable. This is complicated by studies suggesting that people are unaware of the how these systems work in practise (Kobsa, 2007) and can differ in how they value their data (Skatova et al., 2013). Data privacy is a concern for developers of personalisation systems as it could have an impact on how your system is received.

3.5.2 Validation of PDCs

The cards appeared to be promotors of thought into the design of personalisation systems. They formed the basis for the designs created by the participants. For example, the outcome of the thematic analysis was that of using experience presumably from the data type card in their personalisation systems. This can be seen in the quote from a participant (1) ‘And (…) data is task experience, So I thought we have like low to high task experience and at low task experience the controllers have more control to make decisions (…)’. Participants using the cards in their designs is to be expected and is echoed in other ideation card-based work (Halskov & Dalsgård, 2006; Wetzel et al., 2017). If the cards were utilised, one may make an assumption the cards were relevant and a previous ideation card study found that certain cards were not used and they had no evidence as to the cause (Halskov & Dalsgård, 2006). In the case of our personalisation cards, participants noted that certain cards were not used due to concerns over data usage rather than relevance. This adds additional validity to the card sets as the aim of the cards being ideation while simultaneously the production of negative ideas promotes discussion into ideas that one person may feel is unethical but another may not. There were, however, revisions that needed to be made and a redesigned set of cards were developed which address some of the concerns about aforementioned clarity of the card design (Appendix D – Revised Personalisation Design Cards and Game Mat; a list of changes can be found in 3.5.3).

How the participants’ work fits into the taxonomy of automated personalisation systems is an interesting topic. It was previously noted that one of the systems designed (Participant 6) would be a swapper-type system and another system (Participant 9) had suggester-type functionality. It appears that the swapper and suggester cards were the most common, and it appeared that in written material non-suggester type systems would have suggestions as an action (for example participant 3). This may be due to the suggester and swapper systems being common in everyday life, with many systems being one or the other (or elements of both), such as Yahoo! (Manber et al., 2000) and Personalised Search (Montgomery & Smith, 2009). With future automated systems having more personal data, it may become more common for users to be able to more easily visualise more advanced systems, such as Controllers. Even with this limitation, the participants in this study were able to design these types of systems (Participant 1).

As part of the taxonomy, it was argued that customisation was not included, and this is suggested in previous work, as it is users changing the system rather than being personal data-driven (Sundar & Marathe, 2010; Zhang & Sundar, 2019) (Chapter 2: Taxonomy of Personalisation Systems). In the present study it would have been interesting to see whether participants would attempt to include customisation into their personalisation systems. However, the uses of the word customisation are not frequent and appear to refer to changing of the user interface through personalisation, such as the example, ‘So if somebody has […] disability, maybe the video will also be customised to think whatever the person's disability is’ (Participant 3). There may be further scope to examine whether customisation’s definition in terms of personalisation has changed or whether the use of the word "customisable" means for user changes to a system and customised means to have a system changed for a user (as stated prior (Sundar & Marathe, 2010; Zhang & Sundar, 2019, Chapter 2: Taxonomy of Personalisation Systems)).

3.5.3 Redesigned Personalisation Design Cards

The cards were redesigned to address the participants' concerns (Figure 13) (Appendix D – Revised Personalisation Design Cards and Game Mat). The image on the "controllers" card was replaced to better reflect the meaning of the card. The use of mats to identify cards may have added a greater level of complexity to the game and thus, the alternative of writing on the cards was formally included onto the cards. The actions and decisions being similar (which was inferred from a participant's participation during the study) were changed to a new card of Physical Characteristics. The fNIRS card also includes the context of headbands to avoid the confusion which appeared to occur during the study. The Teleoperation card now includes the words robotic to add further context to the card. The errors which occurred due to printing required some changing of the text format to reduce the likelihood of these errors occurring in the future.

Figure 13 – The Redesigned PDCs

As the cards were shown to be valid as tools to design personalisation systems, only small changes were made to increase the usability and reduce the potential for confusion.

3.6 Limitations

One limitation is the effect of accidentally becoming 'the complete participant' (termed by Robson, 2011, p320) due to one answering participant questions. This was minimized by providing abstract answers. Another limitation was that certain participants may have had less input than others and this could affect the data. To reduce this, the researcher attempted to make each participant complete one round of the game individually. Another limitation was the lack of participant experience in manufacturing. Many of the participants discussions, while leaning slightly into manufacturing, do not contain vivid potential use cases or description of industry specific issues. This was too be expected with the chosen population, but for an initial look into this type of method, the cards demonstrate an effectiveness in enabling the design of personalisation systems.

3.7 Conclusion

In conclusion, this chapter examined how people would personalisation design systems to resolve the thesis' first research question through the development of the Personalisation Design Cards and the running of multiple workshop studies. These workshops showed a design preference for systems to be dynamic. Some participants achieved this through creating systems that generated new shortcuts or changed based on user experience and demographics. There is thus room for personalisation systems to innovate in these areas. Further, in regards to acceptance, the study found that certain data types appeared to be “off the table” for some people and there may need to be further insight into the theme of compensation. The study echoed part of the findings from the motiving work (Marinescu et al., 2022). The novel cards were developed with insight from prior work (e.g. Wetzel et al., 2017; Marinescu et al., 2022) and proved successful at allowing users to design personalisation systems. The cards were revised based on the outcomes of the workshops.

Future work in this area could look at expanding the card set, either by including more cards or changing the task cards to represent different industries (perhaps as an Expansion Pack). This would allow the design of personalisation systems in the workshops to better match fields outside of manufacturing. Further, while the work found certain insights, the group sizes were small and these findings may not represent the broader scope of people within the manufacturing sector. It could be possible to relay these findings to other people and see whether they are consistent or whether, when they are not the people designing the systems, their views change.

Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems

4.1 Introduction

With different categories of personalisation defined (Chapter 2: Taxonomy of Personalisation Systems), there is potential to use these categories to understand how users feel about different types of personalisation. This, combined with the findings relating to how users design their own personalisation systems (Chapter 3: Codesign with the Personalisation Design Cards), forms a basis into how acceptance into personalisation can be studied. A survey forms the basis for this chapter; by including a range of existing acceptance metrics, it is possible to detail current stakeholder acceptance. Thus, for this chapter, two experimental studies featuring an experimental quasi-survey provided to stakeholders about personalisation in the workplace was the chosen method.

For automated systems, user acceptance is crucial for their use (Baldwin & Rouleau, 1981) and it would not be amiss to suggest that personalisation systems (which can be considered as automated) have their own acceptance challenges. Users have been known to balance '[...] their disclosure decisions by the cost and benefits [...]' of providing personal data for personalisation (van de Garde-Perik et al., 2008). These user calculations are not without some merit, as user data may be incorrect (Norman, 1998, p256) and data may be misused (Weiser, 1991) (Chapter 3: Codesign with the Personalisation Design Cards). However, Kobsa (2007) theorised that 'internet users often lack sufficient information to be able to make educated privacy-related decisions'. Users may bring previously bad experiences with them when using a new personalisation system (Awad & Krishnan, 2006), even if that is not applicable to the current system. To enable acceptance of personalisation, the case to be made for personalisation needs to be carefully balanced. It is possible to argue that the difficulty in achieving this balance is a factor in the lack of personalisation systems currently in workplace use.

To realise the potential of personalisation in the workplace, the acceptance problem needs to be resolved. While acceptance could be thought of as user-centric, the use of automation in the workplace requires involvement from a wide variety of stakeholders (Baldwin & Rouleau, 1981). Personalisation systems were presented to users in the form of vignettes using an experimental style to alter the systems experienced (like Abraham et al., 2019). The acceptance of the systems was measured using the Advanced Transport Telematics Acceptance Assessment (ATTAA) (Van Der Laan et al., 1997) and a single-question acceptance statement (Abraham et al., 2019) (henceforth termed AQ). Using two approaches for acceptance capture provides additional validity to our findings. Further, it allows the exploration of how these acceptance measuring methods work in comparison to each other. As another measure of validity, our work tracked the effect of user characteristics that may alter acceptance. These were the Need for Cognition (NFC) (Cacioppo & Petty, 1982) and Affinity for Technology (AFT) (Edison & Geissler, 2003). The capture of these two metrics provides a greater understanding of acceptance of personalisation systems. Thus, the contribution of this research is the examination of personalisation acceptance in various scenarios and how the characteristics of users affect this acceptance. This will enable potential integrators of personalisation systems to understand which systems their users would be more accepting of and under what data provision circumstances. To achieve this, the first step was to understand the existing work around acceptance of automated systems and personalisation.

4.2 Literature Review

4.2.1 System Acceptance

The Consumer Acceptance of Technology model (Kulviwat et al., 2007) proposed that there are two main categories of factors (Cognition and Affect) that determine a user's 'attitude towards adoption', which in turn would affect a user's Adoption Intention. These categories contain metrics that could be considered personal, such as 'Dominance'. The difference between users' personal factors could alter their acceptance of personalisation (as seen in Kobsa, 2007). Another model describes users Affinity for Technology (AFT) (Edison & Geissler, 2003). The model combined elements similar to Cognition and Affect together. Thus, there appears to be a crossover between a person's AFT and their attitude towards adoption in terms of the concept of taking non-technology characteristics of the user into account. It is possible to suggest a potential link between AFT and a user's acceptance.

Kulviwat et al.'s (2007) model is a variant of the seminal Technology Acceptance Model (the earliest being found in (Davis, 1985); this model suggests when a user wishes to use a system, they work out whether it is worthwhile for them to use. While this model only discusses the benefits, it is not hard to argue that the potential cost of a system could affect their acceptance of the system. The original technology acceptance model revolves around the 'attitude toward using, in turn, is a function of two major beliefs: perceived usefulness and perceived ease of use' (Davis, 1985). This is similar to the notion present in the Consumer Acceptance of Technology model (Kulviwat et al., 2007), and for the current work, the acceptance of personalisation could be heavily reliant on Davis' (1985) 'usefulness'. Furthermore, if users can bring with them 'positive past experience [...] whose impact on the disclose of personal information is well supported' (Kobsa, 2007), there may be other factors which alter how users perceive personalisation systems. This is important to explore as if personalisation is used by a variety of users, knowing how they would respond could be vital for effective installation.

Kulviwat et al. (2007) include the self-explanatory concept of Adoption Intention with the Consumer Acceptance of Technology model. However, Taviss (1972) has suggested that a person's usage of technology does not match up with their own held opinions. The concept that people may not be wholly consistent is not novel, but it does further add complexity to the idea that the previous model can be used to accurately understand users' intent to use a new technology. In manufacturing settings, there may not be a choice to use personalisation, this being a potential problem with models such as the Theory of Planned Behaviour (Ajzen, 1991). After adoption, users' feelings are important for continued use. In one study, buyers of 'personalized product[s]' were generally happy with them (in this instance, the term personalized is manually-driven 'customization') (Goldsmith & Freiden, 2004; Sundar & Marathe, 2010; Zhang & Sundar, 2019). There is the argument that acceptance lasts beyond initial adoption and feelings about the system are also important to consider.

4.2.2 Attitudes Towards Personalisation

As attitudes and acceptance may be linked (Kulviwat et al., 2007), the current attitudes towards personalisation need to be explored to best understand acceptance. Describing user attitudes towards personalisation is complex as the technology is available in different formats reflecting different Cost/benefit scenarios. Thus, the approach was taken to discuss existing personalisation systems and abstract formats, which may prove useful in presenting current use attitudes. As personalisation is used predominantly in non-workplace settings, the research chosen within this discussion reflects this availability.

Existing research has examined systems that resemble those available in the real world. Personalisation within invitations has been shown to lead to greater take-up by participants (Heerwegh & Loosveldt, 2006). Compared with other personalisation systems, this form of output is simplistic and would require the lowest amount of personal data but provide the least benefit (Chapter 2: Taxonomy of Personalisation Systems). For music personalisation, some participants were concerned about data privacy but the participants in the personalisation condition scored their music more highly (van de Garde-Perik et al., 2008). This shows that while personalisation can make useful outcomes for end users, the impact on data privacy needs to be addressed or mitigated. Whether these findings apply directly to workplace personalisation systems rather than the informal context studied remains to be seen. If the concerns around privacy (such as Marinescu *et al.*, 2022) are substantial one would expect an impact on the acceptance of these systems.

Personalisation systems can be advanced to the point of being autonomous (Chapter 2: Taxonomy of Personalisation Systems). Although rarer, autonomous personalisation systems would harbour similar attitudes to that of other autonomous systems. One study noted that most participants were willing to utilise an automated transport system (Nordhoff et al., 2018). This is a type of system that one would expect the least amount of user resistance to, or the most accepting of, due to the material difference between a human driver and an autonomous driver not (in theory) changing the way the system is used (Hancock, 2019). Hancock (2019) links this to the idea of 'control', which can be linked to the technological approach of the Controllers category from the Taxonomy of Personalisation Systems (Chapter 2: Taxonomy of Personalisation Systems). This link can be made due to the difference between an autonomous driving system and an autonomous system being low when taken in the abstract, as by the nature of automation, they both complete part of a task (Bainbridge, 1983).

User attitudes and understanding of personalisation have changed over time. One such work from earlier in the century (Lavie et al., 2010) suggested that many people may not be used to the idea of personalisation. In contrast, a study closer to the present suggested the opposite (Kozyreva et al., 2021). In just over ten years, there appears to be a dramatic shift in just the basic understanding of personalisation. To explain this change, one needs to view the Office for National Statistics Bulletins. These show that while overall household internet access remains high for the years 2010 and 2019, they show an increase in the number of people online shopping, especially in the eldest age category (ONS, 2013; 2019). Online shopping is most likely one of the main ways people will interact with personalisation systems in the real world, so a demonstrated increase in the number of people using online shopping systems would arguably increase the knowledge of such systems. Further, it has been suggested that the knowledge of technical terminology is not based on the required knowledge for the usage of systems but by the advancement of technology (Norman, 1998, p70). This may have caused a situation where people may be familiar with the systems but not familiar with the terminology, which is evidenced by differences between users' knowledge of 'personalized advertisements' but less so of 'recommender systems' (Kozyreva et al., 2021). If users are more knowledgeable about personalisation, they may be more able to tell these systems apart.

4.3 Method and Materials

To understand acceptance a survey-style experiment is the chosen approach. This is due to prior work use of surveys (e.g. Kozyreva et al., 2021) and ability to include developed acceptance measure approaches (e.g. Van Der Laan et al., 1997). The survey provides an alternative to the codesign workshops from the previous chapter. The choice to use a different method was to cover off the limitations of the prior study design. The small population size of the workshops (which were not manufacturing stakeholders) was addressed by having a lower contact time survey which could be scaled to many more participants (who could be filtered for manufacturing experience). This increases the validity of the thesis and aids in the understanding of personalisation in manufacturing. Ethics approval was granted (

Appendix F – Ethics Approval for Survey).

4.3.1 Scale Choice

The experiment utilised existing scales to form the basis for the survey (apart from the demographic questions (Appendix G – Survey Demographic Questions). The scales can be separated into two categories: pre-treatment and post-treatment. The pre-treatment scales were chosen to best report what factors affect how users accept automation. For this, the AFT scale (Edison & Geissler, 2003) was chosen to better understand how a potential user's affinity (or preference) is related to how and if they would choose to use an automated system. The NFC scale (Cacioppo & Petty, 1982) was chosen to provide a more baseline check of a participant's characteristics relating to what could be described as 'intellectual' qualities and its relationship with the AFT (Edison & Geissler, 2003). These scales allowed the nature of participants to be understood before they were presented with the experimental condition and the potential effect of these participant characters on the conditions.

The post-treatment scales were given after the vignette experimental condition. This was to measure acceptance of the personalisation-based automated system by the participant. The first questionnaire selected for this was the ATTAA (Van Der Laan et al., 1997). While the name suggests an affiliation with transport, the questions are generalised and do not specifically refer to transport. As aforementioned, automated driving is not too dissimilar to the abstracted idea of automation, the scale is a valid way to examine a participant's acceptance of an automated system. The ATTAA was answered as one scale, but the results are split into 'Satisfying' (ATTAAS) and 'Usefulness' (ATTAAU) as specified an option in the original work (Van Der Laan et al., 1997). The NFC was altered to a 7-point Likert scale and the ATTAA was altered to be a Likert scale. For the second experiment, the AQ (like Abraham et al., 2019) was changed from a yes/no choice to a five-point Likert scale due to the results of the first experiment.

4.3.2 Scales and Single Question Acceptance Statement

4.3.2.1 Acceptance Question

The acceptance question is a single question based on a similar question in an existing study, which looks at whether or not a user will use a system after reading one (of many potential types) vignette (Abraham et al., 2019). The scoring was a binary choice in the pilot and expanded to a five-point Likert for the main study. The question was as follows:

> If you were in the situation presented in the Vignette, would you utilise this system?

4.3.2.2 Advanced Transport Telematics Acceptance Assessment

The Advanced Transport Telematics Assessment was an existing scale (Van Der Laan et al., 1997) given to participants after receiving a vignette. Participants answered each question on the scale one by one due to the limitations of the Microsoft Forms platform. These were prefaced by the phrase 'My judgements of the (...) system are...' as in the original work. The original scale has been converted into a standard Likert format with additional headings over each point in the scale (Table 14). This was done for clarity for the participants when representing the form digitally.

Table 14 – Items from the ATTAA scale

Items 1, 3, 5, 7 and 9 formed the "Usefulness" questions and items 2, 4, 6 and 8 formed the "Satisfying" items.

Question	Response				
1.	Useful	Slightly Useful	Neither Useful or Useless	Slightly Useless	Useless
2.	Pleasant	Slightly Pleasant	Neither Pleasant or Unpleasant	Slightly Unpleasant	Unpleasant
3.	Bad	Slightly Bad	Neither Bad or Good	Slightly Good	Good
4.	Nice	Slightly Nice	Neither Nice or Annoying	Slightly Annoying	Annoying

5.	Effective	Slightly Effective	Neither Effective or Superfluous	Slightly Superfluous	Superfluous
6.	Irritating	Slightly Irritating	Neither Irritating or Likeable	Slightly Likeable	Likeable
7.	Assisting	Slightly Assisting	Neither Assisting or Worthless	Slightly Worthless	Worthless
8.	Undesirable	Slightly Undesirable	Neither Undesirable or Desirable	Slightly Desirable	Desirable
9.	Raising Alertness	Slightly Raising Alertness	Neither Raising Alertness or Sleep-inducing	Slightly Sleep-inducing	Sleep-inducing

4.3.2.3 Need for Cognition Scale

The Need for Cognition scale is a previously developed scale by Cacioppo and Petty (1982) (Table 15). The original scale was a Likert of nine; this has been changed to a Likert of seven due to technical limitations. Each question shared the same Likert, ranging from strong disagreement to strong agreement.

Table 15 – Items from the NFC Scale

Question Number	Question Text
1.	I really enjoy a task that involves coming up with new solutions to problems.
2.	I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
3.	I tend to set goals that can be accomplished only by expending considerable mental effort.
4.	I am usually tempted to put more thought into a task than the job minimally requires.

5.	Learning new ways to think doesn't excite me very much.
6.	I am hesitant about making important decisions after thinking about them.
7.	I usually end up deliberating about issues even when they do not affect me personally.
8.	I prefer just to let things happen rather than try to understand why they turned out that way.
9.	I have difficulty thinking in a new and unfamiliar situations.
10.	The idea of relying on thought to make my way to the top does not appeal to me.
11.	The notion of thinking abstractly is not appealing to me.
12.	I am an intellectual.
13.	I only think as hard as I have to.
14.	I don't reason well under pressure.
15.	I like tasks that require little thought once I've learned them.
16.	I prefer to think about small, daily projects to long-term ones.
17.	I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.
18.	I find little satisfaction in deliberating hard and long hours.
19.	I more often talk with other people about the reasons for and possible solutions to international problems than gossip or tidbits of what famous people are doing.
20.	These days, I see little chance for performing well, even in "intellectual" jobs, unless one knows the right people.
21.	More often than not, more thinking just leads to more errors.
22.	I don't like to have the responsibility of handling a situation that requires a lot of thinking.
23.	I appreciate opportunities to discover the strengths and weaknesses of my own reasoning.
24.	I feel relief rather than satisfaction after completing a task that requires a lot of mental effort.
25.	Thinking is not my idea of fun.

26.	I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.
27.	I prefer watching educational to entertainment programs.
28.	I think best when those around me are very intelligent.
29.	I prefer my life to be filled with puzzles that I must solve.
30.	I would prefer complex to simple problems.
31.	Simply knowing the answer rather than understanding the reasons for the answer to a problem is fine with me.
32.	It's enough for me that something gets the job done, I don't care how or why it works.
33.	Ignorance is bliss.
34.	I enjoy thinking about an issue even when the results of my thought will have no effect on the outcome of the issue.

4.3.2.4 Affinity for Technology Scale

The affinity for technology scale by Edison and Geisser (2003) was utilised during the survey (Table 16). The scale was kept in the same format as described in the original work.

Table 16 – Items from the Affinity for Technology scale

Question Number	Question Text
1.	Technology is my Friend
2.	I enjoy learning new computer programs and hearing about new technologies
3.	People expect me to know about technology and I don't want to let them down
4.	If I am given an assignment that requires that I learn to use a new program or how to use a machine, I usually succeed
5.	I relate well to technology and machines
6.	I am comfortable learning new technology
7.	I know how to deal with technological malfunctions or problems
8.	Solving a technological problem seems like a fun challenge

9.	I find most technology easy to learn
10.	I feel as up-to-date on technology as my peers

4.3.2 Vignette Experimental Condition

Vignettes were utilised to form the experimental condition, as they have been used in an existing acceptance-style study (Abraham et al., 2019). Using an early variant of the Taxonomy of Personalisation Systems as a guide (Chapter 2: Taxonomy of Personalisation Systems), three different personalisation systems were developed matching the categories present in the work: Suggesters, Swappers and Controllers. Using this approach avoided the perceived differences for the end user being undetectable, as real-world systems can share similar characteristics. The Personal data type described in the vignette was chosen based on the results found in previous studies (Chapter 3: Codesign with the Personalisation Design Cards) (Marinescu et al., 2022) that found that users were concerned about certain data types more than others. To evaluate the personalisation system, the single-item scale used by Abraham et al. (2019) was changed to a commonly used range of Likert for the second experiment (but a binary choice in the first experiment). The AQ should attempt to link the participants' responses to their actual likelihood of using the system, avoiding mismatches as discussed in prior work (Taviss, 1972).

Experiment One Vignettes

The first experiment was simpler in design than the second and featured fewer vignettes. A preceding statement was provided with an experimental conditional statement (Table 17). Preceding statement:

> You work as a quality control assistant. Your task is to identify potential defects in manufactured parts and select potentially faulty parts for further inspection by a colleague. Each manufactured part is shown to you one by one on a computer screen.

Table 17 – Experiment One Conditional statements

Condition	Vignette
Suggesters	There is an automated system which uses your previous performance data to place a warning symbol on screen when it believes you miss a type of defect found in the current part. The automated system cannot do the job by itself as it does not have perfect reliability.
Controllers	There is an automated system which uses your previous performance data to automatically complete inspections of parts with defects you are likely to miss. The automated

	system cannot do the job by itself as it does not have perfect reliability.

Experiment Two Vignettes

The vignettes are separated into two conditions of data type. They are further split into three conditions of personalisation system type. The Vignettes that were given to participants are made up of two parts: a preceding statement about the context and a following statement about the personalisation system and data types (Table 18). The preceding statement:

> You work as driver of a robotic arm to assemble manufactured parts. Your task is to control a robotic arm on an assembly line with a joystick controller. You are then able to attach two parts together on the assembly line using the robotic arm to form a part. The robot has some capability to complete tasks but cannot complete the whole task by itself.

Table 18 – Experiment Two Conditional statements

The vignettes link of performance and fatigue justified by Barker & Nussbaum, (2011) who state 'physical fatigue had a significant effect on physical performance' following their experiment.

Condition	Vignette
Performance Data	
Suggesters	The automation system uses your **Performance data** to light up the LEDs on the robotic arm when your fatigue reaches a high level.
Swappers	The automation system uses your **Performance data** to Swap the assemble button with a 'Take a break' button when your fatigue is high.
Controllers	The automation system uses your **Performance data** to Automatically complete assembly tasks when your fatigue is high.
Heart Rate Data	
Suggesters	The automation system uses your **Heart Rate data** to light up the LEDs on the robotic arm when your fatigue reaches a high level.
Swappers	The automation system uses your **Heart Rate data** to Swap the assemble button with a 'Take a break' button when your fatigue is high.

Controllers	The automation system uses your **Heart Rate data** to Automatically complete assembly tasks when your fatigue is high.

4.3.3 Hypotheses

From the literature review, there is a clear indication for the potential outcome of studying personalisation in an acceptance setting and this has been used to form a series of hypotheses. The first set relates to the participant characteristics: the Need for Cognition (Cacioppo & Petty, 1982) and Affinity for Technology (Edison & Geissler, 2003). The second and third sets of hypotheses refer to the experimental conditions of personalisation types and data types (the latter only present in experiment two). The two acceptance metrics to be studied were the Acceptance Question (like Abraham et al., 2019) and the Advanced Transport Telematics Acceptance Assessment (Van Der Laan et al., 1997):

H_{1A} – NFC and AFT would have an influence on AQ

H_{1B} - NFC and AFT would have an influence on ATTAA

H_{2A} – Personalisation Condition would have an influence on AQ

H_{2B} – Personalisation Condition would have an influence on ATTAA

H_3 – The data type used by the personalisation system will have an effect on the AQ and ATTAA

4.4 Experiment One

4.4.1 Participants

The participants were recruited on the Prolific platform (an online participant recruitment platform) with the filters/sub-filters: Work, Industry and Manufacturing. This would ensure that the participants have a connection to manufacturing and thus can be considered potential stakeholders. To ensure participants were correctly recruited, they were queried about their experience in the questionnaire. The platform automatically replaced participants who did not complete the study (or those who were "rejected" for not correctly participating) with new participants. Due to the limitations of the Microsoft Forms and Prolific platforms, participants were recruited in sequential waves for each of the conditions of the vignette.

In the first experiment study, the participants' self-identify was majority male (Male = 31, Female = 16, Non-binary = 1). The participants indicated that their age ranges were skewed towards the 25-34 bracket (18-24 = 7, 25-34 = 19, 35-44 = 16, 45-54 = 3, 55-64 = 2, 65+ = 1). Participants indicated that their education attainment was majority grouped between the A-level and

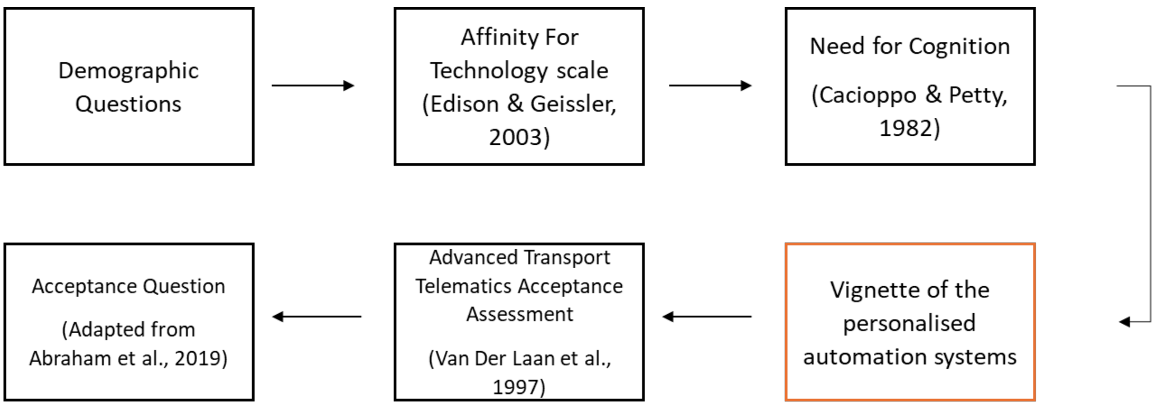
Undergraduate categories (or equivalents, GCSE = 3, A-Level =13, Foundation Degree = 7, Undergraduate Degree = 13, Master's Degree = 9, PhD = 1, Prefer not to Say = 2). Participants were asked to confirm that they had experience of manufacturing environments (Yes = 43, No = 5). From this, it is possible to ascertain that the majority were working-aged stakeholders in the manufacturing industry from a wide variety of different educational backgrounds. Two participants were removed from the analysis for selecting the same value for each question for two out of the three surveys. There were no formal exclusion or inclusion criteria (unless failing the attention checks or providing clearly erroneous data).

4.4.2 Study Design

The first experiment was a between-subjects type. Participants were split into two conditions of personalisation system type (Suggesters N = 23, Controllers N = 25) based on the Taxonomy of Personalisation systems (Conditions = Suggester or Controller). Upon joining the study, the participants were assigned to one condition until the required amount of participants was met. The stimuli were the provision of a vignette which described a personalisation system of quality control context and of condition type. This was provided to participants after they had completed the demographic, AFT and NFC questions. After seeing the vignette, the participants completed the ATTAA and AQ.

Figure 14 - Participants will complete the scales during the study in the order shown in the diagram

The vignette is the experimental condition and will change to represent the Suggesters and Controllers conditions. Diagram similar to Yao et al. (2019).



4.4.3 Procedure

The participants started on the Prolific platform. The participants were then directed to the MS Forms in which the survey was contained. The ethics documentation was included in the survey, which the participants read and consented to the study. Participants who did not consent were able to withdraw by closing their browser window. Participants worked through the scales, questions and vignettes, all held in the same MS Form (separated where required into separate "pages"). The first questions were demographic-focused (Appendix G – Survey Demographic Questions). The next section was to read vignette stimuli (this approach was adapted from Abraham et al., 2019). After this, there was the requirement to complete the ATTAA (Van Der Laan et al., 1997) on the subject of the vignette. The final section was an adapted version of the single-question acceptance statement (AQ) (Abraham et al. (2019)). After this, the participant was redirected to the Prolific platform, where they received payment if they completed the study. During the study, participants had to correctly complete two attention checks; failure of these would make a participant ineligible to receive payment and their data could not be used.

4.4.4 Results

4.4.4.1 Data Analysis Approach

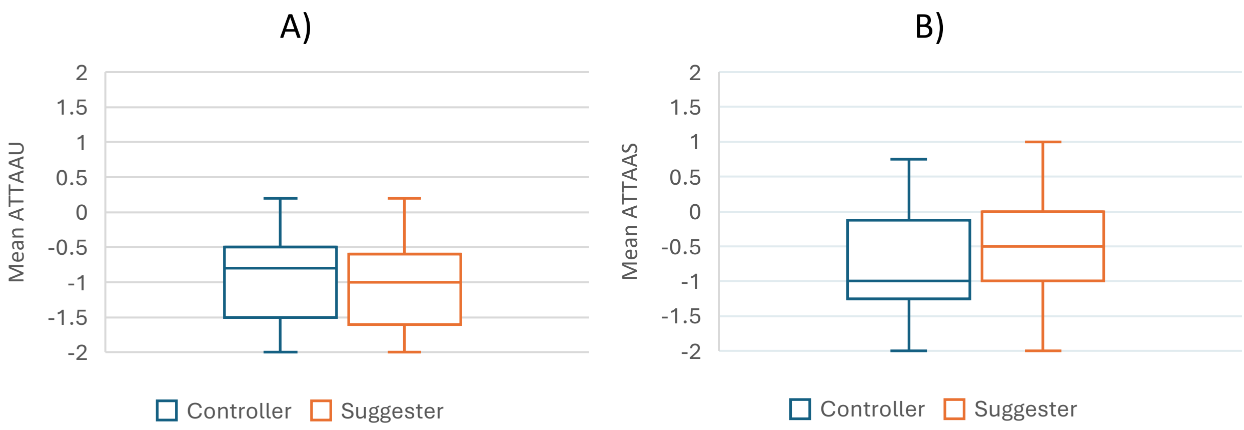
To understand the effect of the different personalisation systems, a statistical test was used to compare differences between the conditions. To analyse how the different participant characteristics affected the acceptance and how the ATTAA related to the AQ a regression approach was taken. This allows for the understanding of the relationship between two variables (Beers, 2024) by trying to use a line to track the link (e.g. if one variable increases does the other). The B metric mathematically describes the relationship or 'slope' (Beers, 2024). The Wald X^2 is the test statistic, testing whether a 'variable add some incremental value to the model' or line (Analyttica Datalab, 2021; Statistics How To, n.d.). For the current study, the regression is used to see whether a variable can predict another (e.g. can the participant characteristics predict the AQ results). Probit (and later Logit) refer to the approach taken to create the line and the choice of which to use does not materially affect the results (Grace-Martin, n.d.-a).

4.4.4.2 Effect of Personalisation Condition on ATTAA Survey

Non-parametric analyses were performed to understand if the personalisation condition had an effect on the ATTAA survey. A Mann-Whitney U for the effect of Personalisation Vignette condition on the ATTAAU to be not significant, p = .487, Controller Mdn = -.8, Suggester Mdn = -1. (H_{2B}). Further Mann-Whitney U analysis for the effect of Personalisation Vignette condition on the ATTAAS to not be significant, p= .220, Controller Mdn = -1, Suggester Mdn = -.5 (H_{2B}).

Figure 15 - Graphs A, B show the mean of the ATTAA scale scores (separated by Usefulness and Satisfying)

For the ATTAA, the lower score represents a greater level of usefulness or satisfaction (Van Der Laan et al., 1997).



4.4.4.3 Relationship Between Existing Surveys and Binary Acceptance of Automated Personalisation System

With a similar experiment design to Abraham et al. (2019), utilising a similar analysis approach of a multilevel regression model would be theoretically valid. However, using a probit multilevel model (due to the binary data type of the Acceptance Question) produces a result but is not able to produce a positive hessian matrix, suggesting the result would be invalid. This can occur when there is not enough differentiation between groups (Karen Grace-Martin, n.d.-b). In the current study, there was an equal amount of not accepts in both groups and 18 and 20 accepts for Suggesters and Controllers (respectively), which could be the reason for the lack of difference. Thus, for the analysis on the Acceptance Question, the conditions were combined.

To investigate how the three different Surveys (AFT, NFC, ATTAA U/S) correlated to the AQ (H1A, H1B), a probit analysis was performed (Table 19). The total model was significant, LR X^2 = 17.729, p = .001 and a significance was found for AFT, B = -1.371, p = .025, suggesting a correlation between this scale and the AQ. No significance was found between either of the ATTAA conditions: ATTAAU B = .427, p = .330, ATTAAS b = .218, p = .641 or NFC, B = -.024, p = .108. Thus, the ATTAA scale and NFC were not to be related to the AQ. The AFT was shown to have a relationship with the AQ.

Table 19 - Probit Binary Regression Scores and Significance Values per condition

Scale Scoring	Wald X^2	B	p
NFC	2.579	-.024	.108

AFT	4.995	-1.371	.025
ATTAAU	.950	.427	.330
ATTAAS	.218	.191	.641

4.4.4.4 Relationship Between the Participant Characteristics and the ATTAA

The participant characteristics can be analysed with an ordinal probit linear regression to understand the relationship between the ATTAA, NFC and AFT (Table 20). This was chosen due to the hessian matrix's variance being acceptable for one half of the survey but not the other. It was decided that retaining coherence and analysing each separately was more valid than mixing statistical analysis methods. Within the ATTAAU with both conditions, the total model was significant: Suggesters, LR X^2 = 8.178, p = .017; Controllers, LR X^2 = 13.018, p = .001. Further, the AFT was significant in both suggesters, B = -1.257, p = .005) and controllers, B = -1.498, p = <.001. The NFC was not significant in either condition. For the ATTAAS part of the scale, both conditions had a significant total model, Suggesters, LR X^2 = 10.483, p =.005; Controllers, LR X^2 = 13.161, p =.001. As with the ATTAAU, the ATTAAS had a significant relationship with the AFT, Suggesters, B = -1.289, p =.004, Controllers b = -1.558, p = <.001. There appeared to be a consistent link between the AFT and the ATTAA in both conditions. It is worth noting that a higher ATTAA score is a lower level of "acceptance". The NFC was able to show a significant relationship, B = .028, p = .048 in the Suggesters condition. This same relationship was not found for the Controllers condition.

Table 20 - Probit Linear Regression scores for AFT and NFC scales on the ATTAA scale, separated by usefulness and satisfying

Condition	Scale Scoring	Wald X^2	B	p
ATTAAU				
Suggesters	AFT	7.845	-1.257	.005
	NFC	.685	.011	.408
Controllers	AFT	11.140	-1.498	<.001
	NFC	.014	-.001	.904
ATTAAS				
Suggesters	AFT	8.174	-1.289	.004
	NFC	3.897	.028	.048
Controllers	AFT	11.216	-1.558	<.001
	NFC	0.001	.000	.973

4.4.5 Discussion

The hypothesis that relates to the current experiment are:

H_{1A} – NFC and AFT would have an influence on AQ

H_{1B} - NFC and AFT would have an influence on ATTAA

H_{2A} – Personalisation Condition would have an influence on AQ

H_{2B} – Personalisation Condition would have an influence on ATTAA

The single question acceptance statement (AQ) and acceptance scale (ATTAA) have not been shown to be related in the first experiment. In each case, the general scores showed that participants were mainly scoring the systems as acceptable. The result could demonstrate that the yes/no answer to the question was not precise enough to capture the differences between the acceptance scale scores. This is further evidenced by the fact that both sets of participants were generally accepting of the systems in their responses to both metrics and thus, there is little room to make strong inferences about the effect. To rule out the effect of imprecision on the result, a repetition with a large range scale for this single question should be conducted (the original range in Abraham et al. (2019) was 11).

The effect of a participant's characteristics on their acceptance of the systems can be explored. A relationship was found between the Affinity for Technology scale (AFT) and single question acceptance statement (AQ). This effect was not found for any of the other scales. The results may also be related to the aforementioned imprecision effect of this question. In any case, it is not possible to reject the hypothesis (H_{1A}).

For the effect of participant characteristics on the acceptance scale (ATTAA), the results showed that the Affinity for Technology (AFT) had a clear relationship with the acceptance scale responses. It can be expected that one's preference for technology would affect one's acceptance of technology, and the inverse of this link was previously suggested, with the author discussing the TAM (Edison & Geissler, 2003). It is an interesting finding as many of the ATTAA questions relate to different aspects of the system's functionality, and thus the participants who liked technology more scored the system more highly. This is important to note as in manufacturing settings, users who are less naturally inclined to like technology may need greater support in terms of how getting used to the technology. The result could also be a 'consistency' effect (Cialdini, 2007, p57-58) where users who rated that they have high AFT feel they need to be consistent and score in a similar way on the ATTAA. In contrast to the AFT, the participants' Need for Cognition (NFC) scores were only shown to be related in one of four cases. This would indicate that the NFC has less of an effect on the participants' acceptance, which could be expected due to the NFC being one of the factors that were utilised in the creation of the AFT scale (Edison & Geissler, 2003). Taking both sets of results, the null hypothesis should be provisionally rejected (H_{1B}) but this examination should be revisited in order to produce more conclusive results.

The effect of the personalisation condition on the two-acceptance metrics (ATTAA and AQ) were also examined. The validation issues within the hessian matrix for the multilevel regression demonstrated a lack of differentiation between the two personalisation conditions for the AQ (Karen Grace-Martin, n.d.-b). The similarity in many of the results between the ATTAA surveys for the conditions further leads us to the conclusion that there may not be a large amount of difference for participants relating to their views on personalisation systems. With prior evidence that users may not understand the terminology in terms of personalisation (Kozyreva et al., 2021), responses found in the current study could be due to a lack of knowledge in which to differentiate personalisation types and may view them in a similar light. It is understandable that users who are not designers or developers of personalisation systems would not understand the intricate differences between the vignettes. In each case (H_{2A}, H_{2B}), the null hypothesis cannot be ruled out.

4.5 Experiment Two

The second experiment aims to solve second thesis research question. This was to understand how accepting of personalisation are stakeholders. An increase in participant population will increase the validity of the results, compared with the first experiment. The design of the second experiment was refined based on the results from the first. The first experiment used a two-condition between-subjects study design for two personalisation systems. The results of this first experiment indicated a need to change this personalisation conditional element of the study to a within-subjects design, as this would allow more examination of the vignettes by participants to increase validity and reduce the effect of participant numbers on the statistical methods utilised. Additionally, the AQ was changed from a yes/no choice to a five-point scale. It was theorised that the question style might have been forcing the choice of a definite opinion when a user may feel unsure about whether they would use a system. Thus, the second experiment builds upon the results and shortcomings of the first to more definitive provide research to the wider aims of the thesis.

4.5.1 Participants

The study utilised 204 participants from an adult population recruited on Prolific using the same filters as the experiment one. From the demographic data, it is possible to determine that the majority of our participants were male (Male = 150, Female 52, Non-binary = 2), and many were around 25-34 years old (18-24 = 39, 25-34 = 64, 35-44 = 50, 45-54 = 37, 55-64 = 12, 65+ = 2), had undergraduate degrees or equivalent (GCSE or lower = 17, A-level = 34, Foundation Degree = 19, Undergraduate Degree = 88, Master's Degree = 42, PhD = 0, Higher than PhD = 0, Prefer not to say = 4), and the vast majority had some form of manufacturing experience (Yes = 189, No = 15). The fact that 92.6% of our participants had some form of manufacturing experience satisfies our aim to examine manufacturing stakeholders.

4.5.2 Study Design

The experiment utilised a mixed-subjects type for the vignette conditions (Figure 16). Further, an additional midway personalisation condition (Swappers) was present, with these conditions being presented to all participants. This was alongside the addition of a data between-subjects condition. The data condition related to suggesting the provision of more “personal” personal data in the hypothetical scenario. This change turned the experiment into a 2x3 design (

Table 21). The experiment made use of two Latin Squares (one for each level of context condition) for participant allocation to cover all cases (R. A. Bailey & Peter J. Cameron, 2003). A Latin square was used to reorder the vignettes order given to participant groups. For example, one participant group may receive the vignettes in an order of A B C, another in B C A and another C A B (where A B and C refer to the personalisation conditions of the vignette).

Figure 16 – Study Design Diagram

The vignette is the experimental condition and will change to represent the two data type conditions. Further, the presentation of the Suggesters, Swapper and Controllers conditions will change to match the Latin Square design (R. A. Bailey & Peter J. Cameron, 2003). Diagram similar to Yao et al. (2019).

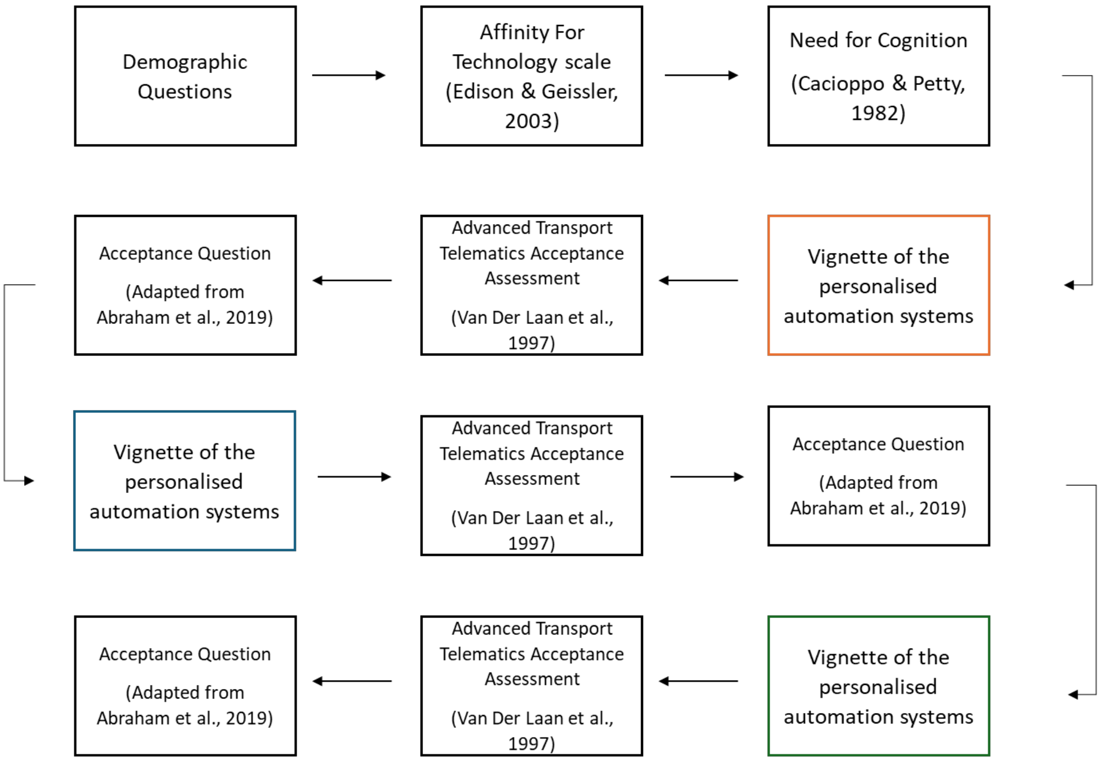


Table 21 - The Mixed subjects design presents the context as a between-subjects condition

With each participant then being presented with each condition of the personalisation-based automated system. Six groups of participants were utilised. These are mapped onto the table rows.

Data/Personalisation	1st Vignette	2nd Vignette	3rd Vignette
Performance	Suggesters	Swappers	Controllers
	Controllers	Suggesters	Swappers
	Swappers	Controllers	Suggesters
Heart Rate	Suggesters	Swappers	Controllers
	Controllers	Suggesters	Swappers
	Swappers	Controllers	Suggesters

4.5.3 Procedure

The second experiment started in a similar way to the first. After the participants completed the AFT scale (Edison & Geissler, 2003) and the NFC scale (Cacioppo & Petty, 1982). The participant then proceeded to the experimental section containing the vignettes. After the participant read the first vignette and answered the ATTAA scale, the participant was asked to rate their willingness to use the system on a 5-point Likert scale. The participant then read a different vignette and re-complete the ATTAA and AQ questions. This was repeated a further time for the final vignette condition. Thus, the participant would have completed a Swappers, Suggesters and Controllers vignette in the same data condition (the order of the vignettes was altered for the different groups).

4.5.4 Results

4.5.4.1 Data Analysis Approach

In comparison to the first experiment, the mixed-subjects design requires a more complex data analysis approach. Rather than a Mann-Whitney U, a multivariate ANOVA would allow the accounting for all the potential difference effects of the conditions (personalisation system as within subjects and data type as between subjects) on the acceptance scale data. A regression was still the appropriate choice for the remaining analyses as it can still show how the collected data may relate to each other. The binary regression was replaced with an ordinal regression to match how the binary choice was replaced by a five-point Likert for the AQ. The descriptive statistics were for the NFC, Performance Mdn = 26.5, Heart Rate Mdn = 24.5, and the AFT, Performance Mdn = 4.2, Heart Rate Mdn = 4.2. The descriptive statistics for the acceptance tables can be found in accompanying table (Table 22) (Figure 17,

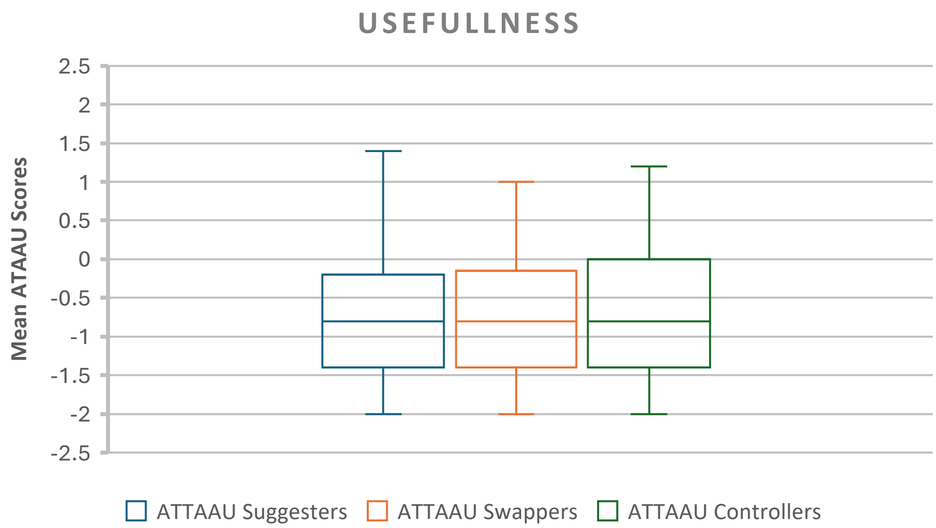
Figure 18).

Table 22 – Medians of the Acceptance Scales separated by data condition and personalisation condition

Condition	ATTAAU Mdn	ATTAAS Mdn	AQ Mdn
Performance Condition			
Suggesters	-0.8	-0.25	4
Swappers	-0.8	-0.625	4
Controllers	-0.8	-0.75	4
Heart Rate Condition			
Suggesters	-0.5	0	4
Swappers	-0.8	-0.75	4
Controllers	-0.8	-0.375	4

Figure 17 - Graphs A, B showing the mean scores of the ATTAA "usefulness" and ATTAA "satisfying" scale halves for the Performance Condition

The ATTAA survey leftmost (in this case lower scores) trending scores as more useful. The graphs show limited difference of the personalisation conditions upon the ATTAA survey.


USEFULLNESS
Mean ATAAU Scores
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
ATTAAU Suggesters
ATTAAU Swappers
ATTAAU Controllers

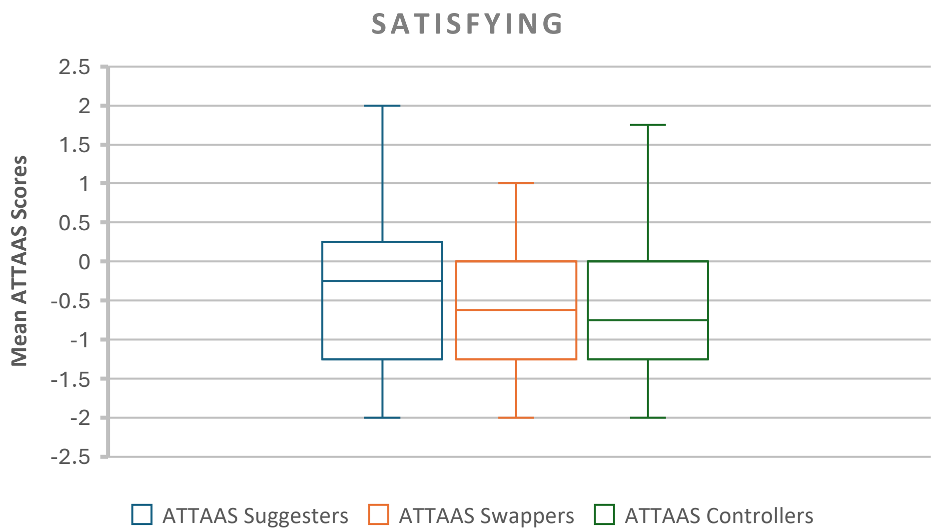
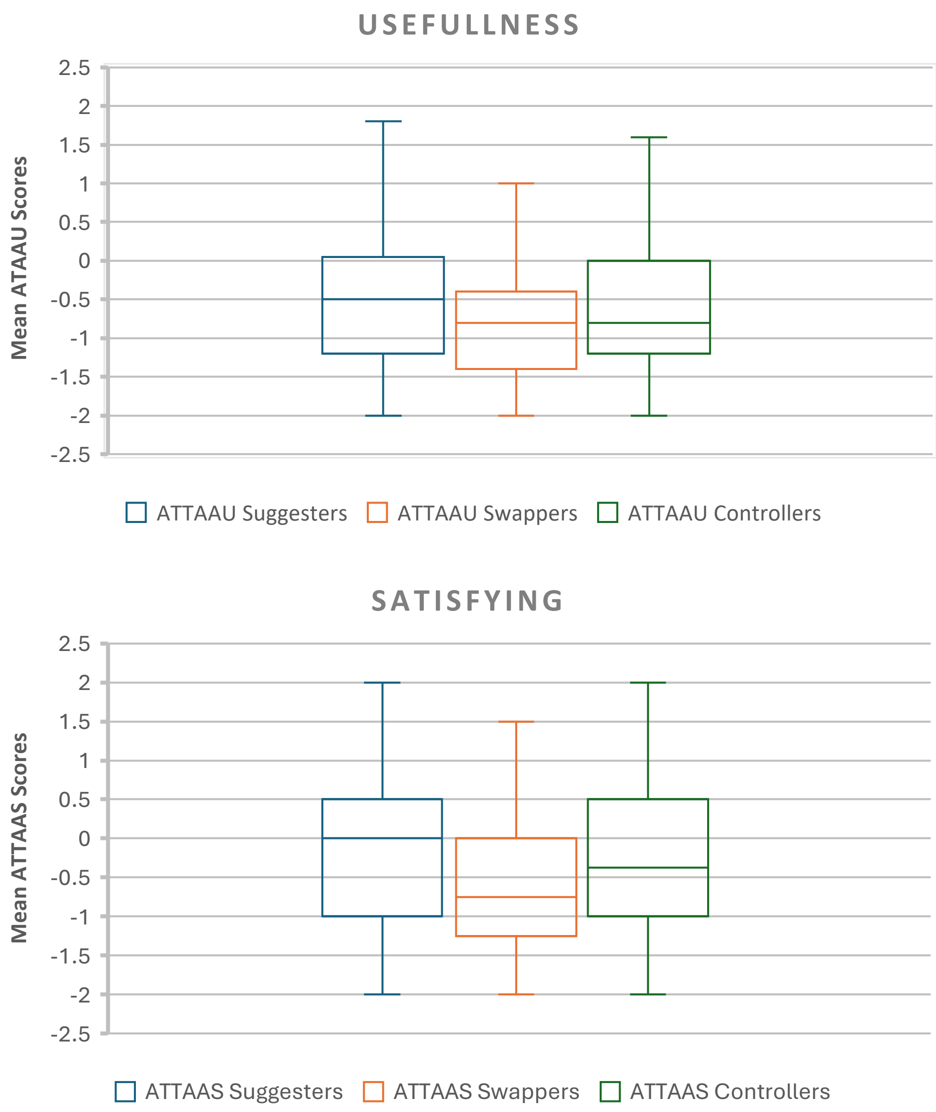

SATISFYING
Mean ATTAAS Scores
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
ATTAAS Suggesters
ATTAAS Swappers
ATTAAS Controllers

Figure 18 – Graphs A, B showing the mean scores of the ATTAA "usefulness" and ATTAA "satisfying" scale halves for the Heart Rate Condition

The ATTAA survey leftmost (in this case lower scores) trending scores as more satisfying. The graphs show that the Swappers personalisation condition is more Useful and Satisfying than the alternative personalisation conditions.



4.5.4.2 Examining the difference between how users perceive different types of personalisation systems

A multivariate ANOVA was conducted to understand the difference between the personalisation conditions and the data conditions. The analysis was conducted in line with and existing approach described in online audiovisual material by Mike Crowson (2020). Using this approach, the multivariate results were utilised as no significant violation of Box's was found. The analysis was conducted for the personalisation conditions with each of the acceptance reporting methods, AQ F = 3.460, p = .003; ATTAAU F = 5.514, p = .005; ATTAAS F

= 8.895, p = <.001. The data conditions were more varied, with only the ATTAAU reporting a significant value in results, F = 3.945, p = .021. The pairwise results were mostly non-significant, apart from a consistent difference found across all the acceptance metrics between the Suggesters and Swappers conditions within the heart rate group (Table 23) and the ATTAAS for the performance group. Also, in the heart rate group, a difference between the Suggesters and Controllers was also found for the ATTAAU/S. A between-subjects difference was only found within the ATTAAS survey, F = 4.457, p = .036, this was not replicated for the ATTAAU or AQ scores.

Table 23 - Pairwise comparisons from the Multivariate analysis of different acceptance measures and the associated within and between experimental conditions

Data Type	Personalisation Condition	Personalisation Condition	p
ATTAAU			
Heart Rate	Suggesters	Swappers	<.001
Heart Rate	Suggesters	Controllers	.514
Heart Rate	Swappers	Controllers	.008
Performance	Suggesters	Swappers	1
Performance	Suggesters	Controllers	1
Performance	Swappers	Controllers	1
ATTAAS			
Heart Rate	Suggesters	Swappers	.002
Heart Rate	Suggesters	Controllers	1
Heart Rate	Swappers	Controllers	.006
Performance	Suggesters	Swappers	.042
Performance	Suggesters	Controllers	.197
Performance	Swappers	Controllers	1
AQ			
Heart Rate	Suggesters	Swappers	.045
Heart Rate	Suggesters	Controllers	.945
Heart Rate	Swappers	Controllers	.455
Performance	Suggesters	Swappers	.607
Performance	Suggesters	Controllers	1
Performance	Swappers	Controllers	1

As significant differences were only found between the Suggesters and Swappers condition (for all scores) and controllers and swappers (for ATTAA scores) for the heart rate group, the null hypothesis cannot be completely ruled out (H_{2A}/H_{2B}). The results show that only in terms of the ATTAA Usefulness there was a significant difference between the conditions. However, this was not replicated across all the other acceptance measures and thus, the null hypothesis cannot be discredited (H_3).

4.5.4.3 Examining the relationship between existing acceptance reporting methods and a user's acceptance of a system

The analysis of the ATTAA and its relationship with the AQ was conducted using an ordinal logit linear regression model. The models were generated for both halves of the ATTAA and for each condition of the personalisation system and data condition (this repetition is required due to limitations in data structure (Aggarwal & Ranganathan, 2017)). As further explanation, a model was created for mapping the ATTAA U/S (together) with the AQ for each unique experimental combination. The results for a model, for example, Performance Suggesters is then split into ATTAAU and ATTAAS (see discussion in Analyttica Datalab, 2021) and shows that one the former is significant, and the latter is not. This provides evidence that the ATTAAU and AQ have a correlation.

For the performance data condition, the suggesters model fit was significant, LR X^2 = 26.747, p = <.001, the Swappers model fit was significant LR X^2 = 48.620, p = <.001 and this was also found in Controllers model fit, LR X^2 = 32.794, p = <.001. Significances were also found in the data condition models: Suggesters model fit: LR X^2 = 34.327, p = <.001; Swappers model fit: LR X^2 = 55.617, p = <.001; Controllers model fit: LR X^2 = 46.705, p = <.001. The analysis revealed a significant negative correlation for 5 of the 6 ATTAAU conditions with the acceptance condition (Table 24), thus showing a relationship between strong "usefulness" and the acceptance question. This was not found for the ATTAAS, with none of the analyses providing a significant result.

Table 24 - Ordinal logit linear regression scores for the relationship between AQ and ATTAA separated by experimental condition

Condition	ATTAA type	Wald X^2	B	p
Performance Condition				
Suggesters	U	8.131	-.905	.004
	S	1.910	-.372	.167
Swappers	U	17.558	-1.757	<.001
	S	.992	-.337	.337
Controllers	U	9.467	-1.226	.002
	S	.960	-.292	.327
Heart Rate Condition				
Suggesters	U	3.788	-.716	.052
	S	2.440	-.567	.118
Swappers	U	18.398	-1.974	<.001
	S	.050	-.084	.823
Controllers	U	8.199	-.996	.004
	S	2.516	-.501	.113

4.5.4.4 Examining existing acceptance factors and how these are related to acceptance

As with the examination of acceptance scores, an ordinal logit linear regression was used to examine how the AQ is related to the NFC and AFT surveys (

Table 25). For the performance conditions, the Suggesters model fit was not significant, LR X^2 = 3.849, p = .146, the Swappers model fit was significant: LR X^2 = 18.092, p =<.001 and significance was found for the Controllers model fit: LR X^2 = 8.475, p =.014. For the data conditions, the Suggesters model fit was not significant, LR X^2 = 3.942, p = .139, neither was the Swappers model fit: LR X^2 = 5.179, p = .075 or the Controllers model fit: LR X^2 = 1.366, p = .505. The results show a limited link between the AQ and the NFC at an expanded p value >0.1 within the performance condition. None of the other combinations of participant characteristics and personalisation/data conditions provided a consistent relationship.

Table 25 - Ordinal logit linear regression for AQ scores

Condition	User Characteristic	Wald X^2	B	p
Performance Condition				
Suggesters	AFT	.210	-.147	.647
	NFC	3.705	.014	.054
Swappers	AFT	.032	-.056	.858
	NFC	14.873	.031	<.001
Controllers	AFT	1.570	.429	.210
	NFC	3.594	.014	.058
Heart Rate Condition				
Suggesters	AFT	3.365	.724	.067
	NFC	.048	-.002	.826
Swappers	AFT	.234	-.178	.628
	NFC	4.846	.017	.028
Controllers	AFT	.173	.148	.677
	NFC	.587	.006	.444

To analyse the ATTAA data in relation with the AFT and NFC scores, the same ordinal logit linear regression was utilised (Table 26). The model fits for the Performance data conditions were all significant, Suggesters: LR X^2 = 8.033, p =. 018), Swappers: LR X^2 = 14.502, p = <.001, Controllers: LR X^2 = 11.377, p = 0.003. The model fits for the Heart rate data conditions were less consistent: Suggesters: LR X^2 = 1.026, P = .599, Swappers: LR X^2 = 8.396, p = .015, Controllers: LR X^2 = 1.536, p = .464. ATTAAS scores were less conclusive (Table 27). For the performance condition, the NFC scores were related to ATTAAU scores. This was unable to be replicated with the AFT for the performance condition. The heart rate condition was unable to determine any conclusive relationship across all personalisation conditions.

Table 26 - Ordinal logit linear regression for ATTAAU scores for each of the data and personalisation type conditions

Condition	User Characteristic	Wald X^2	B	p
Performance Condition				
Suggesters	AFT	1.960	.425	.162
	NFC	7.885	-.019	.005
Swappers	AFT	.037	-.064	.847
	NFC	11.492	-.025	<.001
Controllers	AFT	1.803	-.475	.179
	NFC	4.428	-.016	.035
Heart Rate Condition				
Suggesters	AFT	.094	-.108	.759
	NFC	.526	-.005	.468
Swappers	AFT	.211	.166	.646
	NFC	7.404	-.020	.007
Controllers	AFT	.047	.076	.829
	NFC	1.397	-.009	.237

Table 27 - Ordinal logit linear regression for ATTAAS scores for each of the data and personalisation conditions

**Refers to significant model fit.*

Condition	User Characteristic	Wald X^2		p
Performance Condition				
Suggesters*	AFT	1.803	.398	.179
	NFC	6.590	-.018	.010
Swappers*	AFT	1.757	-.444	.185
	NFC	7.501	-.020	.006
Controllers*	AFT	3.230	-.638	.072
	NFC	1.990	-.010	.158
Heart Rate Condition				
Suggesters	AFT	2.213	-.525	.137
	NFC	.002	.000	.968
Swappers	AFT	.521	-.251	.470
	NFC	2.518	-.011	.113
Controllers	AFT	.368	-.210	.544
	NFC	.281	-.004	.596

When the analysis was performed on the AQ, significant differences were only found in one condition for one NFC (of twelve). These results were similar in the first experiment study, and it is not possible to rule out the null hypothesis (H_{1A}). In terms of the ATTAA scale, the NFC did appear to have some effect on the identified Usefulness of our participants. However, it was not possible to establish any link between the AFT and how participants rated the systems in terms of the ATTAA, showing almost the opposite of the first experiment. Thus, the null hypothesis (H_{1B}) cannot be ruled out due to the results found in both the first and second experiments.

4.5.5 Discussion

User Characteristics Effect on Acceptance

As an early study into understanding acceptance of personalisation (and their accompanying data types), the first set of hypotheses relates to how user characteristics would influence how they accept the personalisation system. Although, an effect was found most strongly in the performance data condition. As the heart rate condition was chosen to map onto the previous study's ideas of controversial data types (Chapter 3: Codesign with the Personalisation Design Cards), this could be attributed to the lower Usefulness found in this condition. To put it another way, participants were more concerned about the heart rate data type itself and focused less on the Usefulness of the system.

This would make sense as it is often suggested that 'data privacy' is key for users (Canhoto et al., 2023; Chen et al., 2022).

Personalisation Systems Effect on Acceptance

As the work attempted to discover how users would respond to the different types of personalisation systems. There appears to be a curve of Usefulness and Satisfying, which peaks at the Swappers and decreases on the edges. This result helps validate the differences between the systems present within the Taxonomy of Personalisation Systems (Chapter 2: Taxonomy of Personalisation Systems). For the performance condition and the ATTAA scale, the participants rated the systems similarly enough not to produce a large difference. This may be due to the participants' belief that performance data is less controversial.

For the heart rate data type, the results are more definitive. The Swapper type system was indicated to have more Usefulness and be more Satisfying to use (looking at the IQRs), and this can be attributed to the nature of swapping elements of the system, which was thought to be more useful than the Suggester system. Further, the controller system is seen as less useful than the swapper system. The reason for this is unclear and it could be experience-related (Schiaffino & Amandi, 2004). The Satisfying scores portray the participants preferring the Swapper system. The Suggesters being low in satisfaction could be directly related to Usefulness, such that they believe the system would not be as Satisfying as it is not useful. The Controllers system would also have a low satisfying score as the system is taking control of the task, rendering the user as a bystander (or a supervisor who shares and adds its own set of problems (Bainbridge, 1983)) and the original ATTAA paper suggests an Autonomous Intelligent Cruise Control system would be rated poorly (compared to the 'control') on the scale (Van Der Laan et al., 1997). From this, a picture can be built of the Suggester and Controller systems generally falling behind the Swapper in terms of Usefulness and Satisfying scores. Systems designers should be cautious about implementing these types of systems into the workplace on these grounds and thus should focus on implementing Swapper systems.

Data Type's Effect on Acceptance

The final hypothesis relates to the difference in data type in the vignettes (H_3). From the data, it can be suggested that the usefulness is directly affected by the data type chosen to be utilised by the system. As seen from the graphs, the means and IQRs show the data type was causing greater differentiation between the conditions. The result may be due to participants considering the systems more forensically due to the potentially negative data type (Marinescu et al., 2022). This effect on how users rate these systems has an impact on the implementation; systems designers should be careful when using different data types.

Link between Acceptance Metrics

The current study was able to explore the problem of mis-calibration between user feelings and the choice to use a system (Taviss, 1972). The analysis showed a relationship between the single-question acceptance statement (AQ) and the ATTAAU scores. This suggests that a user considers the Usefulness of a personalisation system when considering whether to use it. A user may be concerned that the addition of a personalisation system -if it is not useful- could be more cumbersome and inhibit work.

4.6 Limitations

It is possible to highlight a set of limitations. The first is in the study's design. The use of vignettes, while a replication of an existing study's method (Abraham et al., 2019), cannot provide the same results as a design in which the participants use the systems. In the case of the current work, it was decided that the requiring of a large number of participants to use a system and report on it would generate data too complex to reasonably be able to interpret, combined with the practical problems of setting up a task in which many participants could partake. Another potential limitation is in the vignette's design; for example, the suggestion of a 'take a break button' may make more sense in a performance data setting than a heart rate data setting. An attempt was made to mitigate this potential issue. However, it was impossible to ensure perfect parity as the data types were chosen in such a way that one was more controversial than the other (Chapter 3: Codesign with the Personalisation Design Cards) (Marinescu et al., 2022).

Another potential limitation is in the number of factors chosen to be examined. The Affinity for Technology scale was chosen due to its inbuilt consideration of multiple factors (Abraham et al., 2019). With one of these factors, the NFC is being utilised in the current study (Cacioppo & Petty, 1982). It could be argued that these two scales do not represent a large set of potential factors that relate to acceptance. For example, the Contributions of Technology and Dangers of Technology surveys could have been used (Taviss, 1972). This was not done in order to avoid a survey of intensive length, which could cause potential issues of participants rushing through (Nodder, 2013, p64-65).

4.7 Conclusion

In conclusion, this chapter examined stakeholder acceptance of personalisation systems in manufacturing settings in two experiments to contribute to the second thesis research question. This was achieved by using a set of existing scales that represent potential acceptance factors (Cacioppo & Petty, 1982; Edison & Geissler, 2003) and an experimental vignette condition (Abraham et al., 2019) and the acceptance question or scale (Abraham et al., 2019; Van Der Laan et al., 1997) in the form of a questionnaire presented to stakeholders.

In terms of this objectives, the current study was unable to consistently verify whether user characteristics could affect the acceptance of personalisation systems by users. There did appear to be a link between the ATTAA Usefulness scores and the participants' choice to use a system. It was, however, able to show a link to the type of personalisation system and the acceptance of the stakeholders. The current chapter described how potential stakeholders felt about personalisation systems. Building on the findings relating to data types of the previous chapters (Chapter 1: Introduction, Chapter 3: Codesign with the Personalisation Design Cards) (Marinescu et al., 2022), it is clear that there are restrictions on what is considered acceptable. The final objective of understanding acceptance and choice to use a system was explored. However, only a link between ATTAA Usefulness was linked to the Acceptance Question results in the latter experiment.

As these studies have been theoretical and may not apply in reality (Taviss, 1972), to best understand personalisation systems, an implementation needs to be developed. The following chapter explores automated assistance. Initially, developing a basic task assistance and then incorporating a personalisation system to control a refined automation.

Chapter 5: Implementing Automated Assistance in a Manufacturing Task

5.1 Introduction

In prior chapters, the focus has been on depictions of personalisation systems rather than an implementation. However, personalisation needs to be explored in a potential real-world use case. For this chapter, the use of personalisation is examined through the lens of a quality control inspection task. Currently in manufacturing environments, human working with automation allows increasingly complex functions to be possible. This collaborative approach maximizes the 'capabilities and limitations' of each type of worker (The Role of Automation and Humans in Nuclear Power Plants, 1992). To achieve optimal performance, the computational systems that form the environment for collaboration must be designed correctly. Failure to implement automation properly has been suggested as part of one high-profile failure (Leggett, 2019). One technique for potentially increasing ease of use is personalisation. By adding user data to automated systems, the personalisation approach will allow automated systems to closely align with the users with whom they work. This will enable higher levels of collaboration between humans and automation.

Existing work has linked personalisation-like systems to automated environments (Rieth & Hagemann, 2022) but is unable to define exactly how this should be achieved. Work has also explored different types of personalisation on user acceptance (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) and noted differences between different types of personalisation system. However, these studies are suggesting personalisation as a way forward rather than directly testing their assumptions that the technology is worth implementing. Without this knowledge of manufacturing tasks, it is impossible for organisations to confidently invest in personalisation implementation.

One of the potential reasons why personalisation has not been investigated in an implementation state is the high requirements for entry into this area. For a system implementer, a viable task has to be identified and a personalisation system that can understand user data needs to be developed. To form this understanding, a small-scale experiment was conducted to understand how automated assistance can benefit users in a manufacturing task. To understand assistive automation in greater detail before the personalisation element was added, the reliability of this automation was altered. With this understood, it will be possible to understand how poor-quality automated affects users and how this may alter how personalisation is perceived. The second experiment uses a performance-based algorithm to serve as the personalisation element and a Quality Control inspection task to represent the manufacturing element. Thus, the study aimed to understand the effect of personalisation on users of automated systems, in terms of their experience towards the system. A further aim is to understand whether different types of personalisation also contribute to the user's experience. From this work, it will be possible to understand how personalisation can be implemented in decision-support tasks. The effect of these systems on a user's qualitative, performance and subjective data should be understood. The chapter builds upon the existing studies by the inclusion of the Taxonomy of personalisation systems (Chapter 2: Taxonomy of Personalisation Systems) as a theoretical backing for the type of personalisation system included in the study. These results provide a foundation for discussion into the integration of personalisation in manufacturing tasks. First, the prior literature on personalisation and automation shall be detailed.

5.2 Literature Review

The main task was that of Quality Control Inspection. This is a task which has seen implementations of decision support automation (e.g. Dixon & Wickens, 2006). These systems and their associated theoretical underpinning can be described to locate what types of systems have been successful in this area. As the aim is to link personalisation with a decision support system. How personalisation has found use within tasks can present valuable insight into the potential effects of this approach.

5.2.1 Decision Support Automation

A seminal work in this area has been conducted on a participant's detection of elements within a UAV task (Dixon & Wickens, 2006). Its utility for examining the current area is the similarity between this task and a Quality Control Inspection task. Alongside this, assistive automation would indicate potential 'system failure' with the UAV system, which had to be reported by participants. The automation's reliability was tweaked across the work to form the conditions. In the first experiment, fully reliable automation had lower '[System Failure] SF Detection time' than no automation. Although there is no direct comparison between low and high reliability, this effect was not found in the lower reliability conditions. It is possible to suggest that this could appear as a difference between higher and lower reliability automation as part of the current experiment.

In a mobile robotic setting, a demonstration of a robot's reliability was shown to not have an effect on the actions of participants (Salem et al., 2015). In the work, a robot was shown to be 'faulty' to participants and one that behaved as expected. Participants were then required to follow instructions from said robot. The work noted that between the two conditions, the robot's reliability had no effect on how the participants responded to instructions. This forms the hypothesis for the current experiment. The concept of using participants' actions as a way of capturing a participant's trust in an automated system is not limited to robotics, with a 'trust fall' method detailed for automation car-based automation (Miller et al., 2016).

Existing theory has identified a link between 'automation capability' and 'trust' (Lee & See, 2004). However, experimental studies have shown Trust not to be affected by differences in Capability (Lochner et al., 2019; Miller et al., 2016). In these experimental works, the automation capability conditions differ generally by feature set or adjustment of characteristics. While capability is not reliability, there is the argument that participants may have perceived reliability as a type of capability, more so in cases like the current study where the participants are unaware of the potential for different automation reliabilities.

A previous iteration of the experiment (Lammert, 2021) identified a link between the Quality Control Identification task and literature describing a 'Forced choice: four possible outcomes' matrix within Signal Detection Theory (Heeger, 2006). While existing work has identified 'false alarms' (Dixon & Wickens, 2006), the current work adds additional automation actions, incorporating them into the prior matrix (Table 28). The now eight outcomes provide insight into when the user trusts or does not trust the automation (by following its advice or not). Further, the automation alters the factor of Noise and Signal (Pashler et al., 2004, p 43, 44). By providing additional information about whether the tile contains a defect or not, the automation adjusts when a user makes a decision. As the reliability is being varied, the effect on the Noise and Signal thresholds between errors (Pashler et al., 2004, p 43, 44) is different for each reliability level.

Table 28 – Automation and User Trust Table. A matrix of all possible response types after adding a decision support automation, built from Heeger's non-automation matrix (2006).

			Automation Action	
			Correct Alarm	Missed Alarm
User Action	Defect	Hit	Trusting Hit	Distrusting Hit
		Miss	Distrusting Miss	Trusting Miss
	No Defect	Correct Reject	Trusting Correct Reject	Distrusting False Alarm
		Incorrect Reject	Distrusting Incorrect Reject	Trusting Incorrect Reject

Trust between humans and automation is not straightforward, with a highlighted difference in automation usage between pilots and students (Riley, 1994b in Parasuraman and Riley, 1997) in which pilots were not disabling the automation in situations where the students did. The ability of the working demographic of a user to alter the usage of automation could be a potential issue in industry environments. Employees may vary by demographic when integrating an automated system into their work processes. Further, a user's prior experience with a non-automated version of a system has been shown to alter their trust in an automated system (Niu et al., 2018). The work also contained a discussion on the use of a trust scale to capture participants' trust in the system; the scale itself was modified to fit the context. Both works show how users of different experience levels can affect trust in automated systems.

5.2.2 Personalisation Systems Use Within Tasks

Although personalisation as a technology is well established, personalisation's usage in manufacturing tasks is less so. Thus, when examining the current literature, focusing on only manufacturing-based personalised systems would cause relevant literature to be missed. However, even with the limited existing content on the subject area, it is worth discussing how personalisation systems can affect use performance in tasks (such as Dede et al., 2022), as there is limited argument to implement systems that are not effective. Further, understanding how users feel about the inclusion of these types of systems (Gajos et al., 2006) will shape how they are realised in real-world systems.

It is worth examining the non-task-based literature on personalisation. This is due to their having been existing literature discussing trust and personalisation (Monzer et al., 2020). It has been identified that personalisation could increase trust by association with new sources that users already have trust in (Monzer et al., 2020). It is not a leap to suggest that users may be more trusting of personalisation systems that provide information they trust. In a marketing setting they found that personalisation started lower but finished higher in a metric of marketing performance (Postma & Brokke, 2002). Further in a study of this length, it may be possible to see the baseline non-personalisation system have higher levels of trust.

Personalisation systems can also have an effect on how users complete tasks (Dede et al., 2022). In a visual supervision task, participants were found to perform better in situations where an Adaptive Interface system was present (Letsu-Dake & Ntuen, 2009). While this system is not personalisation (due to a non-personal data type utilised (Chapter 2: Taxonomy of Personalisation Systems), this system is close to a Suggester in nature and it can be argued that the effects could transfer. Further work has presented a positive effect of personalisation on performance in a Human Robot interaction setting (Dede et al., 2022). While this work's results are not as statistically validated, in combination with the prior examined work, it still presents a picture of the expected effect.

There may also be an effect of different types of personalisation system on user attitudes (Gajos et al., 2006). Gajos et al. (2006) found that the Visual Popout (a Suggester equivalent) interface was generally the worst received out of the variants they tested (one of these close to a Swapper interface). The previous study made use of a wide range of self-report methods and the current work should expect a similar dislike towards the Suggester system.

5.3 Experiment One: Assistive Automation

The first experiment examined the effect of assistive automation on a quality control task manipulated by its reliability. The automation itself is a Suggester type, which places a border around tiles that it "believes" has defects (implemented using a Wizard-of-Oz approach). The study looks at performance, automation reliability (Dixon & Wickens, 2006) and participant demographics (Niu et al., 2018; Riley, 1994b in Parasuraman & Riley, 1997), as understanding these elements is key to the successful implementation of assistive automation system into real-world environments.

5.3.1 Method

Participants

The experiment involved 36 participants (N = 36). Participants were recruited using either the Prolific online platform or were conveniently available (termed 'local' participants). However, only 18 of the participants had a total data set available and thus were used in the data analysis. The assumption for the majority of the missing data is a unclear user interface element, which led to participants perhaps missing questions unintentionally. Each condition of automation reliability (100%, 75%, 50%) had six participants as an even split. The participants' self-identified gender was split into two categories (Male = 9, Female = 9). Participants' age ranges fell between the 18-24 to 35-44 categories. In addition, participants had to self-report their usage of different types of computer systems and automation types (Figure 19, Figure 20, Figure 21, Figure 22). Prolific participants were compensated for their time and this differed per group due to Prolific's required compensation adjustments. Local participants were not compensated for their time. There is the potential for the immediate results to be affected by this difference. Arguably, the participants compensated will have a different incentive structure to favour completing the task quickly to maximise the time spent over study completion reward. The locally recruited participants have no such incentive. As the study is meant to be demonstrate the validity of the quality control inspection approach, the lack of compensation should have a minimal effect on the tangible outputs. Future work will need to address these concerns by using a unified population (as featured in the second experiment).

Figure 19 - Participants self-reported timescales of smartphone usage

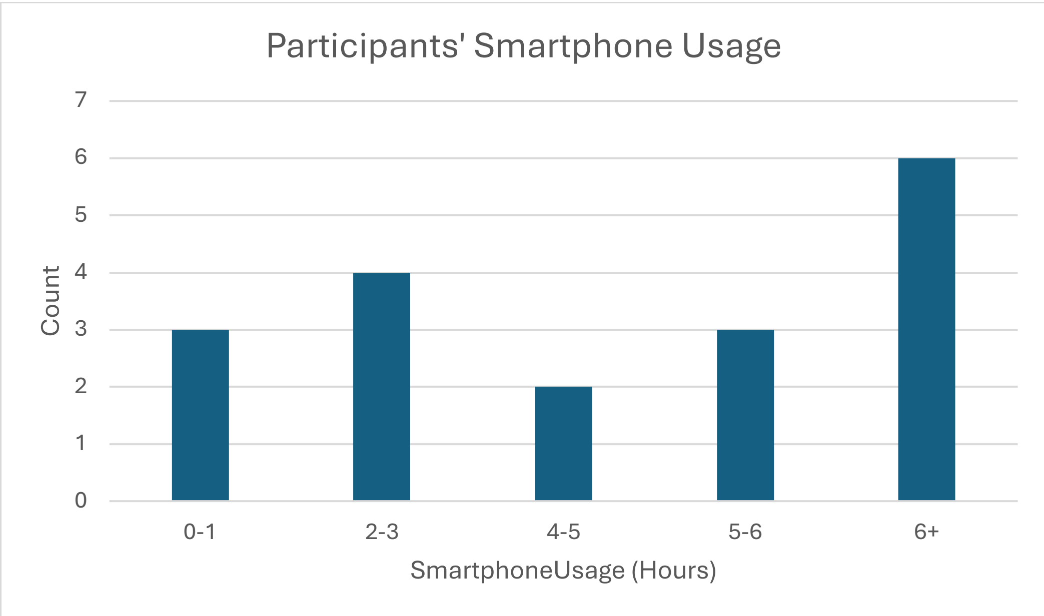


Figure 20 - Participants self-reported timescales of personal computer

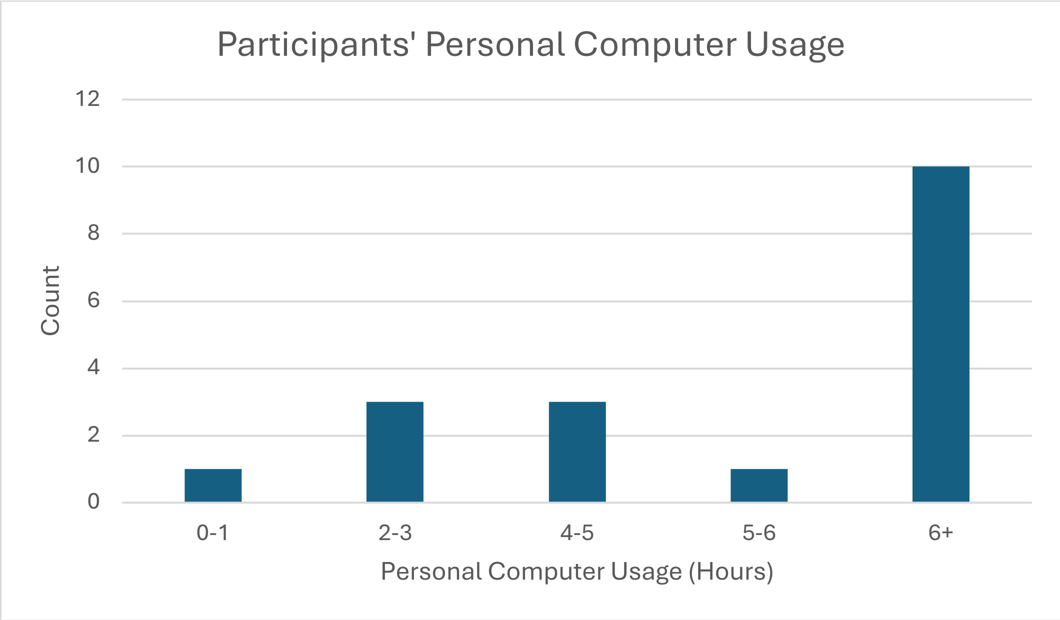


Figure 21 - Participants' self-reported timescales of work-based automation usage

Examples of automated system types were given to participants to provide a base definition of what was meant by the term 'automation' in this study. An example of 'Office automation' found in Foley Curley (1984) was included: [...] this could be anything from a word processor or excel sheet to assisted driving.

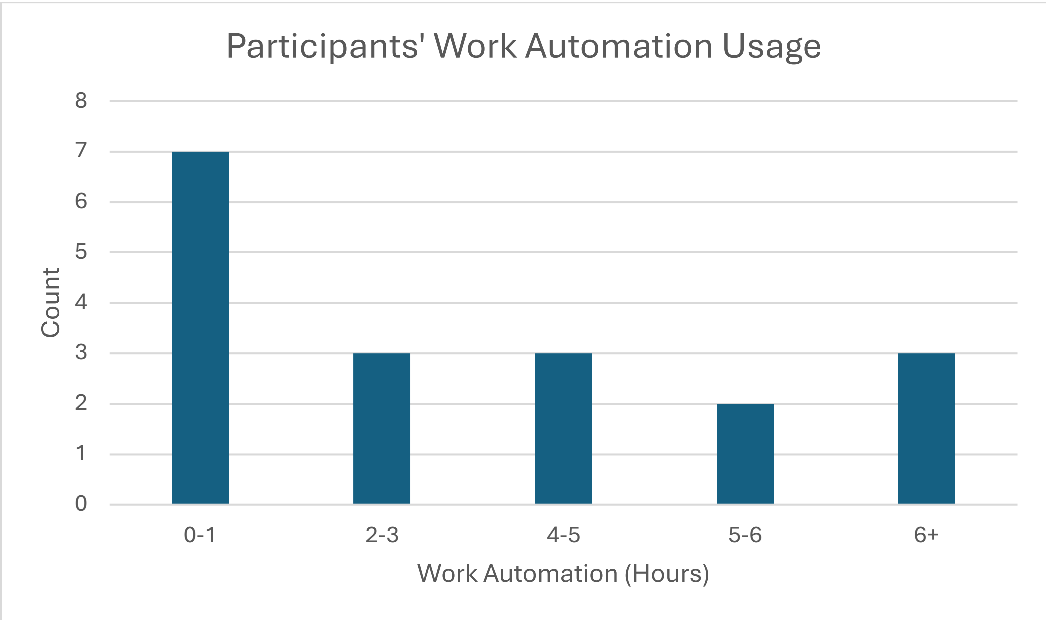
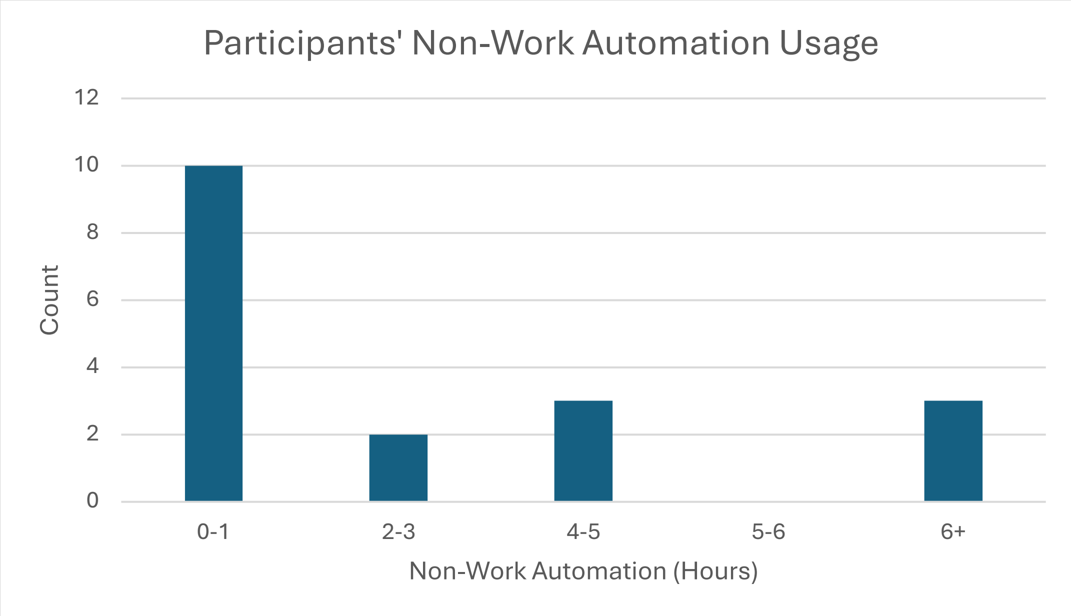


Figure 22 - Participants' self-reported timescales of non-work-based automation usage

Examples of automated system types were given to participants to provide a base definition of what was meant by the term 'automation' in this study. An example of 'Office automation' found in Foley Curley (1984) was included: [...] this could be anything from a word processor or excel sheet to assisted driving.



Materials

As the participants were computer users, they provided their own device. A total of 300 round-edged square cork tiles (originally captured as part of a prior student project, featured in other work, such as Argyle et al., 2021; Houghton et al., 2016) were duplicated or altered to form a set of 1200 which acted as stimuli. These were photographed from a close distance using a digital camera (resolution = 2085x2088) and compressed into a small size for online usability (resolution = 730x731). Of the 1200, 200 tiles were physically marked in one of five ways to imitate a defect: dented, glued, a large surface scratch, had the corner removed (named flat), or featured a small cut from the tile's edge towards the centre of the tile. Each fault had a numeric key between 1-5 allocated and 0 for no defect (Examples featured in the guide for participants Figure 23, Figure 24). By separating the defect keys and no defect key, it would reduce participant error by requiring multiple hands (or significant hand movement) to activate the keys. The task itself was self-paced.

Figure 23 – Guide presented to Participants in the automated assistance first experiment

Figure 24 – Example of a "normal" tile from the supplementary guide

The study utilised the PsychoPy platform (Peirce et al., 2019) to create the identification task (Figure 25). It was also used to display a trust scale (Jian et al., 2000) and a demographic questionnaire (Appendix H – Assistive Automation Experiment One Demographic Questions) using PsychoPy's Form functionality. To display the study, the Pavlovia platform was utilised to host the experiment online. The participant numbers were incremented using the VESPR platform (Morys-Carter, 2021). The data was output as a .CSV file reformatted using Microsoft Excel and analysed using IBM's SPSS Statistics software package. PsychoPy was used to create a Wizard-of-Oz automation by placing green borders around tiles, which the automation believed (manipulated by the researcher) to have a defect. The green borders were a transparent image file overlaid over the tile (resolution = 2085x2088) with a green-bordered hollow square placed away from the edge (Figure 25). A fully transparent image (resolution = 2085x2088) was overlaid in cases where the automation detected no defect.

Figure 25 - Screenshot of one trial of the fault identification task in PsychoPy

The automation's reliability was set to one of three levels for each of the three experimental conditions (Automation Reliability = 100%, 75% or 50%). In this instance, the reliability is the correctness of the identification of a fault (Whether the green border was overlaid on a tile with a fault). The experiment was of a between-subjects design; thus, each participant only experienced one level of reliability. The participants were not informed of the reliability of the automation.

Procedure

The participant started the study on the Prolific platform, which then redirected to the study in Pavlovia, which loaded the PsychoPy experiment in the participant's browser window. The participant then had to complete the consent forms and a demographic questionnaire. A tutorial of the task was then completed by participants (including information about the automation).

The main task required the participant to identify visual defects in images of cork tiles. The tiles were displayed one at a time to the participant in a random sequence. Once displayed, a participant had to select which type of defect they could visually identify or no defect using the number row on their keyboard. A message was displayed if the participant took longer than five seconds, which asked them to select a response faster. The automation was also present during this stage.

Upon main task completion, the participant completed an existing trust scale (Jian et al., 2000). This would then progress them onto a section that debriefed them on the deception of the automation's reliability. PsychoPy then allowed participants to revisit Prolific for payment using a study completion link.

Hypotheses

When describing the current literature on assistive decision support automation, it was apparent automation and experience should be a varying factor in the results. A set of hypotheses were designed to understand how assistive automation can be utilised in manufacturing environments:

> H_1: Participant detection times should decrease with higher reliabilities of automation

> H_2: Trust will be affected by a participant's prior experience of similar systems (automated or not)

By understanding these hypotheses, it should be possible to determine whether automation reliability could affect how users perceive automation. This may have an effect on an implementation of personalisation, which could be built upon an existing automated system (in this case, the defect identification automation).

In addition to the hypothesis, the literature review suggests that certain effects will not be present in the current study. As these relate to the aims of the study, these factors should still be investigated within the data:

- Trust will not be affected by changes in automation reliability
- Automation Reliability should not have an effect on the participants' mean response time for inspections

5.3.2 Results

The results presented utilise two-tailed p values. Results that aim to resolve a directional hypothesis will have an accompanying one tailed value.

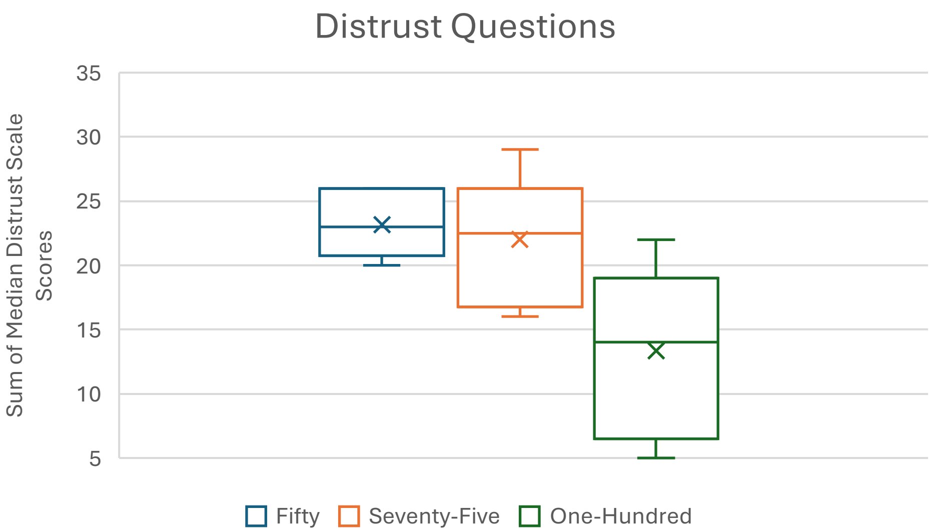
Automation Reliability on Self-Reported Trust

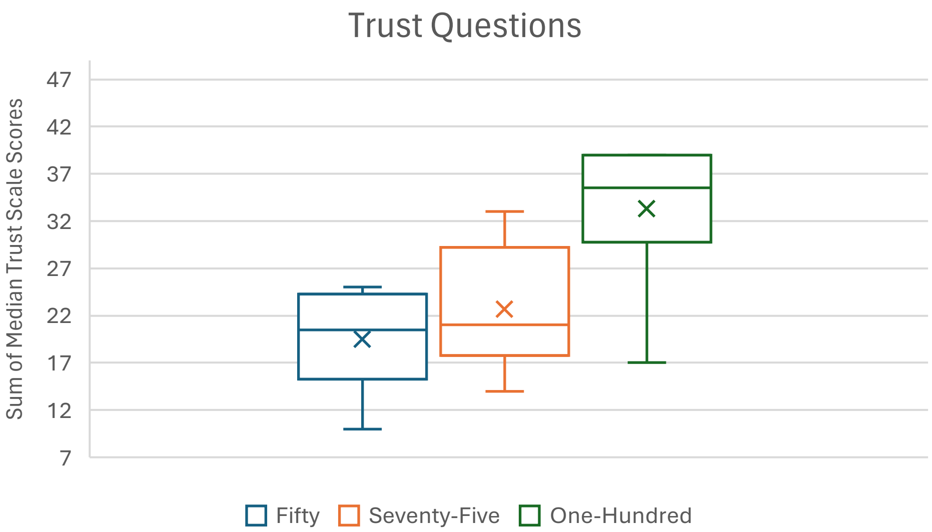
The trust scale questions were split into two categories: questions of distrust (questions 1 - 5) and trust (questions 6 - 12). The numerical Likert value for the two categories were formed from the sum of the individual questions. Analysis was performed on all conditions of Automation Reliability (50%, 75% and 100%) (Figure 26). The following analyses demonstrated an inconsistent effect of the reliability conditions on self-reported trust. Any found effects were located between the 50% and 100% reliability conditions.

A Kruskal-Wallis One-Way ANOVA for the sum of distrust questions showed the main effect of automation reliability was significant, $X^2 = 7.166$, $p = .028$. This was only found when the pairwise comparisons were performed and adjusted on the 100% and 50% reliability conditions, X^2 = -7.75, p = .035. Thus, there is a difference in the results between the 100% and 50% automation reliability conditions when looking at distrust.

A Kruskal-Wallis One-Way ANOVA for the sum of trust questions showed the main effect of automation reliability was significant, X^2 = 6.391, p = .041. This was not found when the pairwise comparisons were performed and adjustments made at the standard level (100% and 50% reliability reported a p = .52, X^2 = 2.383). There is thus no difference between any of the automation reliability conditions which can be proved at the standard level.

Figure 26 - Graphs of sums of the medians of the trust scale questions by automation reliability condition



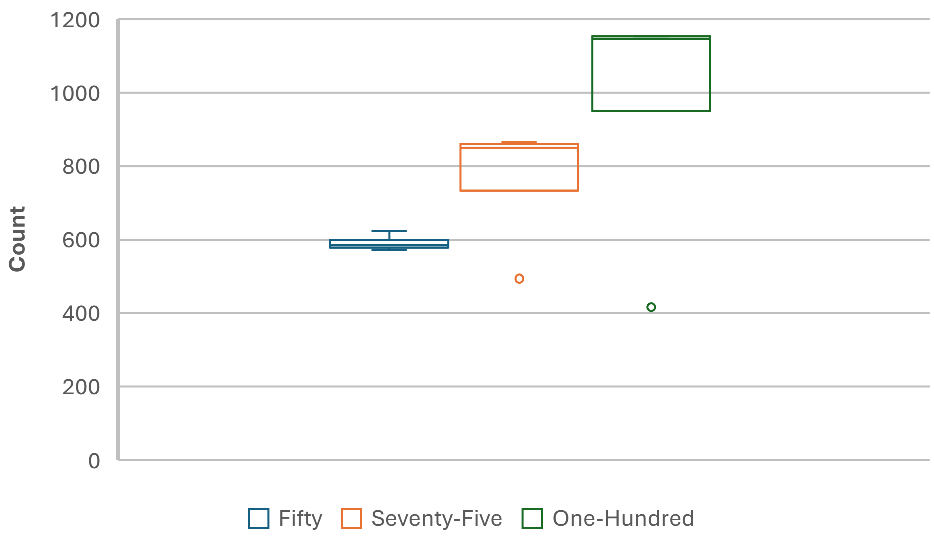


Participants Trust Shown Through Actions

While not an exact ‘trust fall’ as described in previous work (Miller et al., 2016) the level that the participant agreed with the automation was analysed to demonstrate trust in the automated system. “Agreed” is defined in this work as the participant selecting a defect option when the automation indicates a defect is present or the participant selects no defects when the automation does not signal a defect. Analysis was performed for each condition of automation reliability (50%, 75% and 100%).

A Univariate ANOVA analysis observed significance for the amount participants agreed with automation based on the effect of automation reliability, $F(2,15) = 7.719$, $p = .005$. This significant difference was able to be found only when the pairwise comparison was performed on automation reliability of 50% compared with 100%, $p = .004$. Automation reliability of 75% compared with either automation reliability of 100% or 50% was nonsignificant. These results show that 100% and 50% automation reliability conditions were noticeably different when looking through the lens of agreement with automation.

Figure 27 – Participants “Agreement” with the automated systems decisions



Average Response Time per Type of Response

Performance was measured by the mean response times of participants for each of the aforementioned eight potential types of responses to a given inspection (Table 29). 100% automation reliability contains null values for ‘false alarm’ and ‘missed alarm’ as it is 100% correct in all trials and has no ability to obtain those response types. Thus, four Univariate ANOVAs on the four response combinations containing no null values (all combinations of participant response with the No Alarm and Correct Alarm automation response). These analyses found that the main effect of automation reliability on the mean response time of four types of response was nonsignificant at the $p = .05$ level (Participant Hit, Automation Correct Alarm did report a weak link with a one-tailed $p = 0.088$).

Table 29 - Mean response times for a given type of participant and automation responses

			Automation Action					
			Correct Alarm			Missed Alarm		
		Automation Reliability	50%	75%	100%	50%	75%	100%
Participant Action	Defect	Hit	1.486	1.591	1.912	1.451	1.448	n/a
		Miss	2.421	1.382	1.5	1.232	1.168	n/a
			No Alarm			False Alarm		
		Automation Reliability	50%	75%	100%	50%	75%	100%
	No Defect	Correct Reject	1.331	1.023	0.99	1.503	1.197	n/a
		Incorrect Reject	2.199	2.214	3.249	1.885	2.879	n/a

Total Participant Completion Time

Using a Univariate ANOVA, the main effect of automation reliability on the total participant completion time of all trials was found to be nonsignificant, $F(2,15) = .126$, $p = .883$. An ANOVA on the total number of correct inspections by a participant, in which the automation reliability was also found to have a nonsignificant main effect, $F(2,15) = .59$, $p = .943$. For these performance metrics, it appears there is not a significant impact of automation reliability on completion time.

Participant Demographics Effect on Trust

A Two-Way ANOVA analysis was used to analyse the demographic and trust data (justified by a discussion into statistical methods in Yandi, 2020). The system usage demographic data for participants was analysed against the automation reliability for total distrust scores in automation. There was no main effect found of the participants' smartphone, computer or home automation usage with the automation reliability condition on the trust the user held in automation. A significant main effect of work automation usage and automation reliability on distrust in automation was found, $F(5,6) = 4.399$, $p = .05$, All of the four usage demographic data metrics, when compared with the automation reliability, had a nonsignificant main effect on the total trust scores in the automation.

5.3.3 Discussion

The current experiment was positioned to explore three factors that could affect automation usage in real-world environments. The first is the linking of performance to automation reliability and how this compares with existing theory (Dixon & Wickens, 2006). As systems in the real world often vary in reliability, understanding how the users' performance changes with reliability is vital (especially if reliability is dynamic). In addition, the trust of participants may also vary with automation reliability and this needs not only to be understood with self-report methods (Jian et al., 2000) but how participants actually use the automated system (Miller et al., 2016).

Performance

The current study revealed no statistical difference in the reliability of automation on the multiple metrics of performance in the visual inspection task. This is not in line with the hypothesis that detection times should decrease with increasing levels of automation reliability (H_1). The participants' total time to complete all trials was not shown to be different across the automation reliability levels. This finding could be expected as differences between the types of response could be minimised by the combination of all response types. Further, multiple participants had completion times for a singular inspection that were outliers compared with other results (perhaps indicating a break), which could have skewed the results.

To further evaluate participant response time, each of the types of response were individually analysed. This more closely matches the existing work in the area (Dixon & Wickens, 2006). The authors were able to find a difference between types of automation reliability. In the experiment, this did not materialise, with no significant differences between any of the types of response. The lack of difference may be due to the type of decision support automation; in the UAV scenario (Dixon & Wickens, 2006), the automation only signalled for one type of error. In the current quality control study, there were five potential defects to find. Thus, if the automation signalled a defect, the participant still needed to identify the type of defect. This may have taken time of a length that the difference in the initial search would be meaningless by comparison.

Trust

As discussed, participant actions have been used to signify trust in systems (Miller et al., 2016; Salem et al., 2015). There was an effect of the automation's reliability on whether they agree with the automation that is consistent with the findings in an existing study into the trust of robots (Salem et al., 2015). In the robotic study, the participants did not differ in their actions between conditions of an automation displaying 'faulty' behaviour (Salem et al., 2015). In the current study, the participants' agreement with the automation in terms of whether there was a defect was significant across conditions. The results suggest that the participant was not following the advice of the automation in cases it was incorrect. Combining this with the lack of difference in the performance of the participant may describe the participant ignoring the automation in many cases. Although this would be a less concrete conclusion to draw. Nonetheless, it is possible to confirm the previous suggestion that participants' responses were not altered by automation reliability.

In keeping with the theme of trust, it is possible to discuss the trust questionnaire scores. The results were able to show a main effect difference of automation reliability on the scores for both trust and distrust. However, pairwise comparisons were not able to confirm the differences between any conditions. The pairwise results would be in line with the suggestion that trust is not altered by automation reliability from the previous work in using this trust scale with different 'capabilities' of automation (Miller et al., 2016). This result may have occurred due to the limitations in population size. However, it can be argued that for a pilot of this population size, a main effect difference is enough to accept that there is an effect of automation reliability on the participants' trust in automated systems. If replicated and significant differences were found in pairwise comparisons, it could better alert for the potential for participants to be 'implicit[ly] learning [...] independent of conscious attempts to learn' the reliability of the system (Reber, 1993). Looking at the current trends in the medians for the trust and distrust questions, there appears to be a difference in how participants perceived the 100% reliability condition and the less-than-perfect reliability conditions (50%, 75%). Both conditions would have the automation being less reliable than the participant's own ability in achieving the correct answer in the trials (Mean per cent of correct trials across all conditions = 85.02%). There may be a situation occurring when automation reliability is below the participants own ability, they distrust the automation, and when the automation reliability is above their own reliability, they trust the automation. Further work is required to understand this potential situation.

Demographic Data

The participants' responses to the demographic questionnaire and subsequent completion of the study reveals insights into their understanding of automation. It appears that a broad definition of automation (Baldwin & Rouleau, 1981; Foley Curley, 1984) may not be perceived as automation to the participants. Most participants reported more than an hour of smartphone and personal computer usage, which these devices are full of the 'technologies' (or like 'personal computers' are a technology listed) which form automation (Foley Curley, 1984). They also mostly reported low non-work-based automation usage. It appears clear that participants are utilising a different internal definition of what automation is, or what the minimum technical capability of an automation is. Thus, concepts like 'word-processing', which previously would classify as a type of automation (Foley Curley, 1984), are now not regarded as such by participants, perhaps due to widespread use.

The comparison of the demographic data with the automation reliability was able to show (in a limited way) how demographics can change how participants react towards automation. It was possible to show an effect of the participants' reported work-based automation usage on how much they distrusted the automation. The effect of the participants' work-based automation experience appeared varied and it is difficult to draw a conclusive statement on the direction of the effect. This makes accepting the hypothesis of the participants' prior experience affecting their trust in automated systems (H_2) challenging. Further work should examine this further to identify the direction of the effect to ascertain if the existing literature's results can be replicated (Riley, 1994b in Parasuraman and Riley, 1997).

As the current study is attempting to pilot this type of study design, validation should be discussed. The results point to the conclusion that the study design is represented through participants' responses. The trust scale results were able to show a main effect difference of the automation reliability, relaying those participants were aware of the automation's reliability (as stated above as due to 'implicit learning' (Reber, 1993)). While there are no significant differences between each condition, the small study population is likely to be a key factor in this. Further, there is a slow emergence of trends regarding response time per type of inspection. Especially regarding 50% automation reliability, which appears, at face value, to be closer bunched than the other automation reliability conditions, perhaps signifying the participant not acknowledging the automation in a meaningful way (but also that the reverse is true for the other automation reliability conditions). Additionally, Jian et al. (2000) specified that trust and distrust are 'opposites, lying along a single dimension of trust'. The results from the trust appear to flip when compared to the distrust graph, showing that they are 'opposites' and are consistent with the trust scale's intended outcomes (Jian et al., 2000). This is a positive outcome for the study in terms of the validation of the study's design.

5.4 Experiment Two: Personalised Automated Assistance

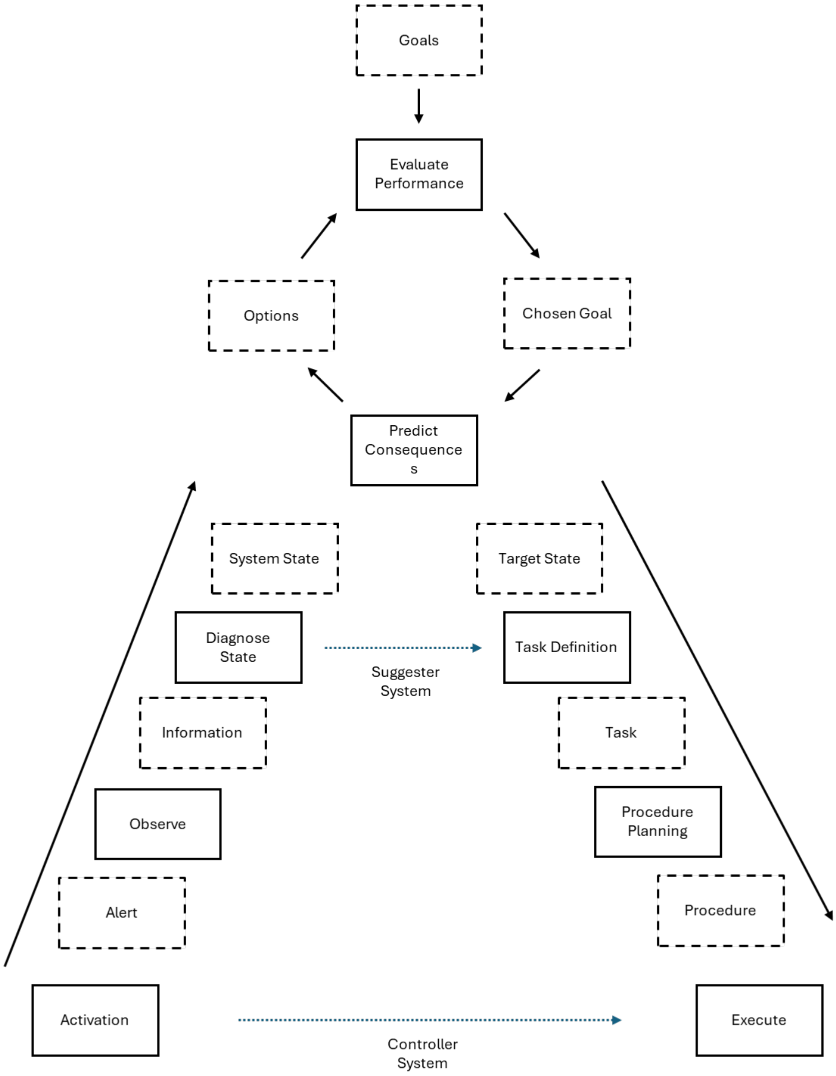
5.4.1 Designing Work-Based Personalisation

The current work features the use of two different types of personalisation systems. These systems are based on an early work into the classification of personalisation systems (Chapter 2: Taxonomy of Personalisation Systems). However, when implementing a personalisation system, it could be better thought of as being related to classical human factors literature. This section will illustrate the chosen personalisation systems and how they are able to change the structure of potential tasks.

In the current study, the focus was on two types of personalisation systems. The Suggesters condition presented the unique colour border around a tile for each defect (when a defect is present) when active. The guide displayed to participants showed each colour border around the associated defect to act as a reminder. The borders were predefined and were not an actual machine learning system. By allowing the user the final decision, the system is similar to previous decision support systems (e.g. Yang et al., 2020) by providing a recommendation rather than completing the problem. The Controllers condition takes the opposite approach and “completed” the next set of tiles for participants. The participants were able to see each tile appear on the screen for a very short amount of time and then it would disappear. The participants were informed that there were completed but these were just skipped and did not feature again.

The choice of personalisation functionality can be related to existing theoretical literature. The Decision Ladder is an approach used to model the potential structure that people use to complete tasks (Figure 28). It has received numerous variations from the original ‘developed by Jens Rasmussen’ (Jenkins et al., 2010), but they are often visualised using the same vertical ascent and descent form (Frank Gleeson & Vincent Hargaden, 2014; Jenkins et al., 2010). The “rungs” on the ladder have been suggested could formed by using ‘automation’ (Frank Gleeson & Vincent Hargaden, 2014) and reference situations where certain elements do not need to be completed (the authors refer to ‘cognitive leaps’ which are human-based). In the current work, it is proposed that personalisation automation can be conceptualised to represent different rungs based on the type. For example, the Suggesters system minimises the state diagnosis and presents a task definition for the user. It achieves this by presenting the task definition for the user by presenting the correct decision to be made. In contrast, the Controller system moves straight to the end of the ladder, thus completing the whole process for the user. The rationale behind having multiple types of personalisation is the implementation. There are situations where a Suggesters system is appropriate for a user. For example, where data quality is low and users can be trusted to make the correct decision. Controllers’ systems are best suited when the user’s personal data indicates they may not be as performative as a digital system. The systems in the present work both represent their respective archetypes from prior work (Chapter 2: Taxonomy of Personalisation Systems).

Figure 28 – Decision Ladder to account for personalisation systems. Built from figures in Frank Gleeson & Vincent Hargaden, 2014; Jenkins et al., 2010



The Cold Start Problem is a personalisation performance problem related to a 'lack of historical data' (Yuan & Hernandez, 2023) when a user first uses a system. To mitigate this, a simple algorithm was implemented. The personalisation system would activate if the average completion time of the last inspection was higher than the average of the previous ten inspections by a buffer (20%). In the grand scale of the utilised experiment (over 700 inspections), the lack of personalisation for the first ten inspections should not be overtly negative for users. If used in an industry setting, this type of 'short-term' personalisation would allow for a fast uptake of the technology in tasks, even with new users. The downside of this approach being the lack of long-term trend modelling. While not implemented here, the data could be collected to be used to understand users' performance over weeks or years. To avoid constant activation, the current system does not activate for the next ten inspections. This allows the personalisation system to gather another set of ten inspections that are not biased by the personalisation system's activation. The algorithm was the same for both personalisation systems in the experiment.

5.4.2 Method

Participants

A new set of participants were recruited on the Prolific platform in multiple waves. The Prolific platform provided the functionality to apply a filter, which were refined using sub-filters. In this case, applied filters were the Work, Industry then Manufacturing. By using the filters, it is possible to make sure our participants have an understanding of manufacturing environments. This allows the results to be a closer match to the population of potential real-world users than a non-filtered population.

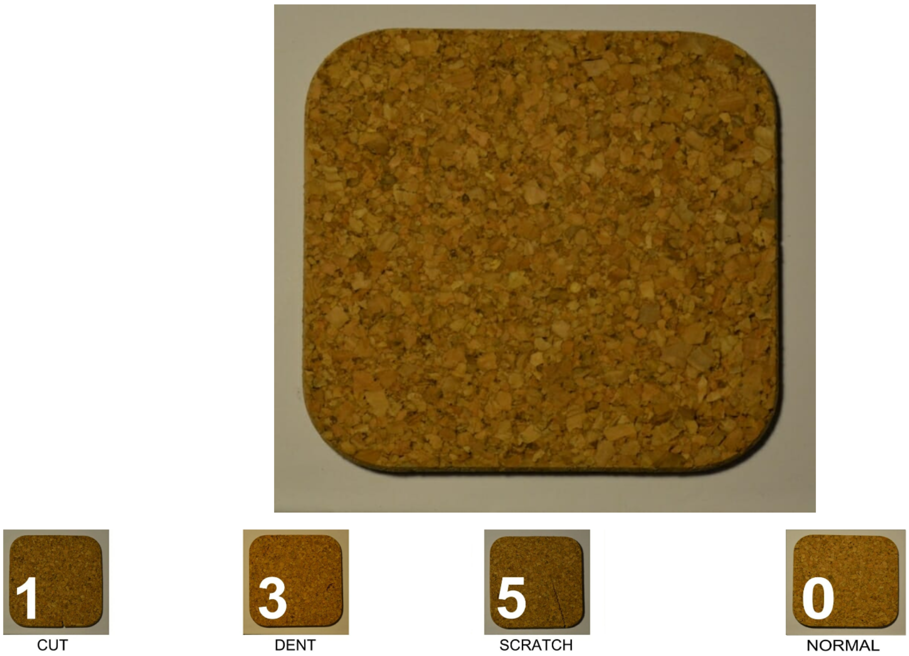
Participants who could be verified to have finished the study were compensated approximately £2.85 (Prolific adjusts this based on median completion time of a wave of participant recruitment). The majority of the participants who joined the study but did not finish were paid a fixed £2.85. Due to the nature of the Prolific platform, participants' who did not provably finish the study were able to have their allocated space in the experiment replaced by new participants.

Excluding those who did not finish, the experiment consisted of 90 Participants. Three participants were removed from the analysis due to incomplete data, duplicate completion of part of the experiment or responding in a different language than the study itself which may have affected their participation in the study. The age ranges skewed towards lower options (18-24 = 15, 25-34 = 28, 35-44 = 24, 45-54 = 11, 55-64 = 8, 65+ = 1, Prefer not to say = 0), degree qualified (GCSE or lower (or equivalent) = 4, A-Level (or equivalent) = 10, Foundation Degree (or equivalent) = 13, Undergraduate Degree (or equivalent) = 43, Master's Degree (or equivalent) = 17, PhD or higher (or equivalent) = 0, Prefer not to say = 0) male (Male = 67, Female = 20, Non-binary = 0, Prefer not to say = 0) and held manufacturing experience (Yes = 67, No = 20). From the 87, two participants in the Controllers condition were excluded from the quantitative analysis. One due to a technical issue and the other for unusually low performance, which, upon closer inspection of data, revealed unexpected identification behaviour.

Materials

For the experiment, 720 non-unique Images of cork tiles were used for the identification task. The images were a subset of those in the prior experiment (and found in other published work e.g. Argyle et al., 2021; Houghton et al., 2016), at a compressed 730x731 resolution. 120 images had one of three defects present: Scratch, Dent or Cut. Scratch has a surface mark in an approximately straight or curved line. Dent has a small, approximately circular mark on the surface. Cut has an approximately triangular removal of part of the tile from an edge. The participants completed the study on their own computer. The PsychoPy/Pavlovia (Peirce et al., 2019) (Figure 29) and MS Forms platforms were utilised for the task and data collection. By utilising an online experiment, it was possible to have a population which was more relevant and in greater numbers. A NASA TLX (S. G. Hart & Staveland, 1988; NASA, n.d.) was utilised due to its widespread usage in understanding a participant's current physical and mental position. A Trust scale (Jian et al., 2000) (Appendix I – Trust Scale) was included to examine the difference in trust of the inclusion of different types of personalisation. The NASA TLX was recreated in PsychoPy with the Trust Scale and demographic questions being recreated in MS Forms. A 'free text' question was included: 2. What are your thoughts towards the system you just used?

Figure 29 – An image showing the participants' perspective of a defect identification trial



Study Design

A between-subjects design was utilised for the three conditions (Baseline, Suggesters and Controllers). The baseline condition had no active personalisation element. The Suggesters suggested a coloured border based on the defect type. The Controllers took the control away from the user.

Procedure

It can be safely assumed participants were directed from Prolific to the PsychoPy study (this was completed by the Prolific platform). The participant was then shown the ethics documentation and asked to complete the consent form. The participant was then provided information about the task (and personalisation system if present) and asked to complete a set of demonstration inspections of the three defects and non-defect tiles. During this section, an incorrect guide was shown, displaying an additional two defects, but this guide was not present during the task itself (this was noted by one participant). Next, the participants completed the task, which consisted of 720 inspections (unless the Controller system was activated). The participants had to press a key associated with the type of tile to complete the inspection and move on. The order was randomised for each participant. If present, during the task the personalisation system could activate in instances of low performance. After the approximately 720 inspections, the participants were then asked to complete the NASA TLX. This completed the PsychoPy section of the experiment. Participants were then directed to MS Forms to complete a Trust Scale, open-ended question and Demographic questions (Appendix J - Assistive Automation Experiment Two Demographic Questions). To complete the study, participants were shown a debrief document detailing the pseudo-like elements of the personalisation system.

Hypotheses

Even with the literature on personalisation in manufacturing being limited, it is possible to generate a range of hypotheses that represent how personalisation should affect the completion of a manufacturing task:

> H_3: Personalisation should reduce the trust held by users compared with the baseline
>
> H_4: Personalisation reduces completion time
>
> H_5: Negative feelings towards Suggester system in the Trust scale

In addition to the hypothesis, there is also the potential for negative thoughts towards the Suggester system with the qualitative data. While this can not be tested hypotheses, it is possible to utilise a qualitative analysis to differentiate the conditions and participant responses.

5.4.3 Results

Quantitative Analysis

The dataset was analysed using Python in the Spyder IDE and additional modules (SciPy, NumPy and pandas). 85 participants were included for the quantitative results (N = 30 Suggesters, N = 29 Baseline, N= 26 Controllers). The response time and percent correct (taken as 1 being 100%) were rounded to three decimal places before analysis. The percentage was utilised to avoid the Controllers participants from having less correct results due to the automation being active. Any data from the period of Controller systems activation was removed. Even in the very limited time a tile appeared on screen, Participants could still respond to the inspection during the Controller activation. Thus, in instances where it was unclear where the system was active, a best assumption was made to determine if the system was active and the participant's response should be removed. Results are presented as two-tailed p values unless stated.

Subjective Measures

A One-way ANOVA was utilised to analyse the NASA TLX (S. G. Hart & Staveland, 1988; NASA, n.d.) and (Jian et al., 2000) trust scale (Table 30). Neither the NASA TLX nor the Jian et al. (2000) Trust scale produced any significant results for main effects. It is possible to suggest that the personalisation conditions were not able to affect the responses of the chosen subjective measures.

Table 30 – ANOVA analysis of the subjective self-report methods. NASA TLX titles from selected variant (NASA, n.d.)

*One tailed, p = .382

**One tailed, p = .234

Measure	Suggesters $\bar{X}$	Baseline $\bar{X}$	Controllers $\bar{X}$	F	p
NASA TLX					
Mental Demand	13.333	12.69	13.385	.162	.850
Physical Demand	10.6	7.966	9.885	1.345	.266
Temporal Demand	13.067	10.931	13.077	1.739	.182
Performance	14.533	15.414	14.231	.852	.430
Effort	15	14.414	14.231	.229	.795
Frustration	12.367	10.586	9.5	1.536	.221
Sum Of Trust Questions	25.267	24.724	23.846	.272	.763*
Sum Of Distrust Questions	14.867	12.793	13.962	.768	.467**

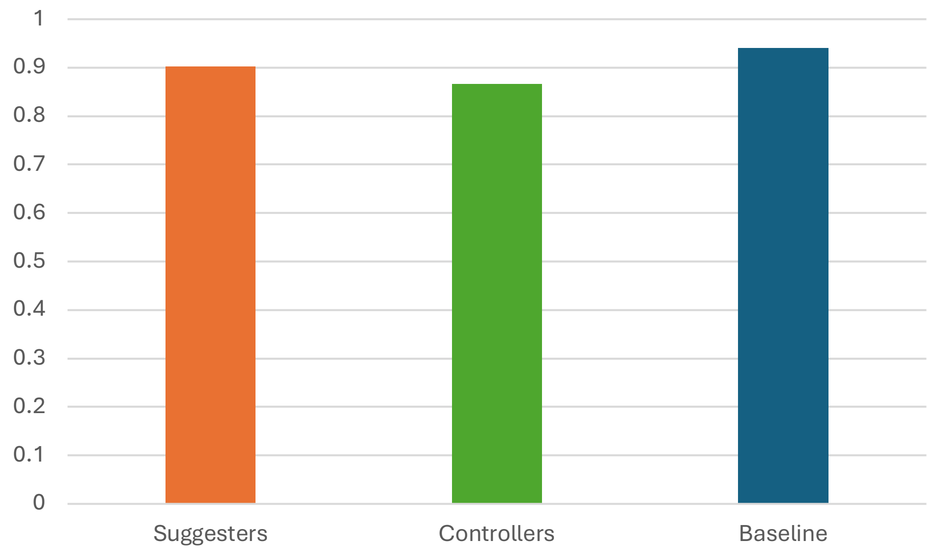
Performance Measures

The mean response time and percentage of correct inspections was also analysed utilising a One-way ANOVA. This analysis revealed no significant results for main effects at the standard, p = .05, level for either Response Times (as a metric of completion time), F = .473, p = .625 (One-tailed p = .313). Suggesters $\bar{X}$ = 1.335, Baseline $\bar{X}$ = 1.206, Controllers $\bar{X}$ = 1.296 or the percentage of correct inspections, F = 3.027, p = .054. Means, Suggesters $\bar{X}$ = .902, Baseline $\bar{X}$ = .941, Controllers $\bar{X}$ = .866. The percentage of correct inspections was significant at the p = .1 level, which provides a small amount of evidence for an effect of personalisation (

Figure 30). A t-test for the number of times the personalisation was activated was nonsignificant, t = 1.144, p = .258. Means, Suggesters $\bar{X}$ = 7.133, Controllers $\bar{X}$ = 5.577. Due to the lack of significance at the standard level, it seems unlikely that the personalisation condition had an effect on performance.

Figure 30 – Bar chart showing the percentage of correct inspections during the task

Values 0.0 represent 0% and 1.0 represent 100%; the bars represent the mean for each condition.



Qualitative Data – Thematic Analysis

The experiment contained one qualitative question after the main task to gain deeper insight than is possible with established scales. An adapted form of the Braun and Clarke (2006) thematic analysis was utilised. The adaptations were required due to the experimental conditions of the study conditions. After the coding the codes were separated by condition, with similar codes being grouped. Further, the 'Story' elements defined by Braun and Clarke (Braun & Clarke, 2006) were completed by condition and contrasted and related by still kept mostly separate. Some of the fourth step defined in the method was completed earlier. Due to the responses given by participants being short in length, the themes consisted of no subthemes; however, one theme was found throughout two of the conditions with a slight modification in the remaining condition (Table 31). This reoccurring theme is separated in the following theme descriptions to provide appropriate consideration.

Table 31 – Themes and quotes from participants are represented for each condition

Quotes presented as supplied by participants.

Condition	Theme	Participant Quotes
Suggesters	Strong Usability	Participant 6: Easy to use, kinda demanding in terms of speed.

		Participant 28: It was easy to learn how to use it.
	System Activation	Participant 8: The low performance detection induced stress and hurry, which might have compromised the quality of the assessments. Participant 11: It made absolutely no sense, the assist would kick in randomly and a lot of the times kick in when I got the correct answer.
	Task/System Expected A lot from Participants	Participant 16: DENT was difficult to identify as it was hard to distinguish between dents and colour difference of the cork. Participant 27: it was a little bit difficult to follow because of the textures, the defects blended with the texture.
Controllers	Personalisation System had High Expectations	Participant 62: it pushed me to be more effective in my task as time went on and occasionally it was helpful when I slowed down, although it put some pressure on my performance. I think it is a good system. Participant 79: I thoughts I would never reach the required speed and that it was not possible to reach in the reality
	Difficult Task	Participant 72: The difference in exposure between photos and different placements of the subjects makes the evaluation unecessarily dificult. The working method is tiresome for extended periods of time. Participant 82: It was really a huge mental task so the system must have been programmed like an AI
	Positive About The System	Participant 66: I find the system easy to use and quite fast. Participant 63: it pushed me to be more effective in my task as time went on and occasionally it was helpful when I slowed down, although it put some pressure on my performance. I think it is a good system.
Baseline	Task Difficulty	Participant 40: It was very easy to use, even though the task was a bit challenging.

		Participant 51: The dent was difficult to identify.
	Questioning the Experiment	Participant 32: I have the impression that the system tried to lull my vigilance at times Participant 36: It makes the user go faster with each image making him open to mistakes. Also I had a weird feeling that the images were getting bigger the more I did them.

Suggesters

In the Suggesters conditions, three themes were identified. These relate to the acknowledgement of strong usability of the system, the insecurity about why and effect of the activation of the personalisation system and the final theme was based on the difficulty and demanding nature of the task and system.

The first theme, Strong Usability, related to a general sense that the system was beneficial to the end users. One participant noted the ability of the personalisation system to aid in decision making '[...] in addition to being of great help in case of lowering performance, helping to increase it or in case of doubt in selecting the options.' (Participant 26). The participant's quote identifies the potential usefulness of a personalisation system can aid in instances where participants may require more aid in completing the task. This is partly expected as the Suggesters system is made to leave the user in control and aid in their choices, in comparison with the Controllers condition.

The theme of System Activation refers to the way in which the personalisation system is activated. Even with the Suggester implementation participants did voice their potential concerns or questions. From this short quote, 'The low performance detection induced stress and hurry, which might have compromised the quality of the assessments' (Participant 8), it is possible to discern that the prospect of a system taking a "Big Brother" type approach may be a net negative for users. Further, the suggestion that 'It made absolutely no sense, the assist would kick in randomly and a lot of the times kick in when I got the correct answer' (Participant 11) provides a sense that participants did not understand the system. In this case, the participant assumed it was their correct identification of defects (or rather the lack thereof) which caused the personalisation system to activate. In actuality it was a decrease in their identification speed. Both of the theme's comments form the basis for the theme relating to the systems activation over the task.

Controllers

As with the Suggesters condition, three themes were identified. Participants identified the high expectations placed on them by the personalisation system, the difficulty of the task and a general sense of positivity surrounding the system.

The Controllers theme: Personalisation system had High Expectations of Users is one that demonstrates the similar feeling felt by some in the Suggesters condition. However, this theme places full emphasis on the personalisation system's weight on the participant during the task. The statement, 'I thoughts I would never reach the required speed and that it was not possible to reach in the reality' (Participant 79), shows how the participant felt unfairly treated by the personalisation system. It is worth reiterating that the 'speed' was judged solely on past performance. Participants could reduce performance over multiple inspections but sudden drops resulted in the personalisation systems activation.

The next identified theme, Positive Views of the system, highlights a general view of the system from participants. This can be seen with a combination of the previous theme from the quote, 'it pushed me to be more effective in my task as time went on and occasionally it was helpful when I slowed down, although it put some pressure on my performance. I think it is a good system.' (Participant 62). Participants suggested different reasons why the system was good, for example, 'this is a very effective system which is also reliable with the results.' (Participant 71). This theme is intentionally wide-ranging as the specific focus on a topic, such as usability in the Suggesters theme, was not as clear.

Baseline

The Baseline condition only located two themes: Questioning the Experiment and Difficult Task. The former of these themes picked up on the concept that participants seemed to question the nature of the experiment. In one such case, a participant's response, 'It makes the user go faster with each image making him open to mistakes. Also I had a weird feeling that the images were getting bigger the more I did them.' (Participant 36) provides an almost conspiratorial stance. The images were not designed to get larger or speed up the user in the baseline condition. This theme may be the result of there not being any manipulation and participants thought there may be secret conditions being observed.

Task Difficulty Reoccurring Theme (with Suggester Variant)

In the Baseline and Controllers conditions, the task's difficulty was noted by participants. A similar finding was located in the Suggester condition but it was wider in scope, covering the system and demands placed by this (and the task). As the theme was found in approximately all conditions, it must be a topic participants felt strongly about no matter the personalisation present.

In the Suggester Condition, an example can be seen: 'It was very fast, so it was rather impossible to think about something else while doing this task.' (Participant 30). A similar idea was conveyed by another participant, 'It becomes difficult to keep concentration and quickness after a lot of tiles showed, especially when asked to identify the dented ones' (Participant 87). Both quotes highlight the high mental demand placed on them by the task. The latter participant described the visual difficulty of one specific type of defect identification.

5.4.4 Discussion

With the analysis completed, it is possible to evaluate the proposed alternative[3] hypotheses for the current experiment:

> H_3: Personalisation should reduce the trust held by users compared with the baseline
>
> H_4: Personalisation reduces completion time
>
> H_5: Negative feelings towards Suggester system in the Trust scale

The first alternative hypothesis suggested that personalisation may lower the trust held by users in the conditions where present (H_3). The current study was not able to locate a difference between the conditions and thus, the null hypothesis cannot be ruled out. This could be due to the personalisation system not being present for the majority of the time and it can be suggested they were not interacting with the personalisation enough to alter their trust in the system at a level perceptible in the experiment.

The completion time of the participants should also have been lowered in keeping with the alternative hypothesis (H_4). In terms of the response time for a single defect, no difference was found (with the null hypothesis being unable to be rejected). These two results could present the lack of utility of the performance-based personalisation system for the current task. This itself may be a continuation of the task/automation discussion found within the pilot in that the automation's assistance is not beneficial enough to affect time-based metrics. Tasks that are simplistic in nature, such as the one included here, may not be ideal use cases for decision support automation of any type. The work by Gajos et al. (2006) used 'Microsoft Word XP' as a task of sorts in which users had agency over how they utilised the personalisation. In the current study, the agency is in the "hands" of automation. Thus, the lack of user agency and the relatively low amount of inspections the personalisation system was active within is the most likely not an effective way of engaging users with personalisation. Dixon and Wickens (2006) were able to demonstrate the effectiveness of non-personalised automation but the decision support was for one task, which was part of multiple simultaneous tasks, thus adding complexity.

[3] Terminology used as shown in figure (StudyCorgi: https://images.app.goo.gl/uCnzL3xTLAhgDYEG6)

There was an expectation of negative feelings towards users in the Suggesters personalisation condition against the non-personalised condition. The Suggester-type personalisation had a specifically positive response in terms of usability. Further, the Controllers condition provides insight into how participants may view these types of personalisation system. The participants' qualitative responses indicated that the Controller had high expectations of them, which may relate to the feeling of being controlled. This irony may provide the logic as to the Suggester condition's unexpectedly strong positive feelings. The allowance for users to choose their own defect allows them to manage at their own discretion, the opposite being a criticism of automated systems (Agnew et al., 1997).

The Trust Scale results (or NASA TLX) were not able to locate any differences of personalisation condition (H_5) and thus, the null hypothesis cannot be ruled out. It was found in the qualitative data that the task was difficult. There may be an effect of the difficult task desensitising the participants from the negative feelings of using the personalisation systems (or the Peak-End Rule (Kahneman, 2012, p380)). For example, if the participants were almost at a maximum level of negativity completing the task, they may not be as affected either way by a change in the automation which is -comparably to the task- in the background. This does, however, suggest that the automated personalisation systems of either type were unable to actively improve the trust held in the system, which may be a vital factor in certain scenarios.

To contextualise these findings, it is worth considering what the personalisation systems are achieving in their attempt to aid users. Firstly, it has been said that people like 'certainty' (Sutherland, 2019, p232). In the case of the Suggesters system, the system would engage in ways that the participant did not understand. Thus, the system could have been placing additional mental stress on participants as they were unsure as to when the system would activate. This would contribute to the pressure felt by participants in the Controller condition as they may have felt intense pressure due to not understanding what level of performance the system expected them to meet. Personalisation systems which react in real-time with the users may be at risk of placing additional stresses on users regarding their performance. The Suggesters noted high usability, may be attributed to its removal of uncertainty for the identification task, as it provided ten inspections where the participant knows if they are correct or incorrect.

On a similar topic, it is known that people generally 'avoid unpleasant surprises' (Sutherland, 2019, p266). In the current study, the participants were often surprised by the Controllers activation and the reduction in their agency by the system taking away control may have been perceived negatively. In contrast, the Suggesters system retains user agency, as they have the final decision on the inspection while being aided in the form of being provided the answer.

While none of the proposed alternative hypotheses were able to be confirmed, the knowledge of how the system did not significantly (at the standard level) change performance could suggest the utility of personalisation. Not in a tangible performance improvement, but linked with the results of the thematic analysis, the Suggesters system may improve usability without compromising performance. The inability to find different subjective measures suggests that further work is necessary to ascertain in what situations personalisation could be used effectively. The specificity of the positive comments may indicate a preference for this type of system over the Controllers.

5.5 Limitations

The experiments' findings should be taken with a number of considerations. Firstly, the task design could have contributed to the results found (or not found). The task was chosen to represent a potential task found in manufacturing settings. This may just be a use case where Personalisation is not effective. Further, this experiment included only two types of personalisation system. There may have been a larger difference with another Personalisation system. The types were chosen as they felt most appropriate to each other, without requiring the task to be redesigned for the system. Further, while most systems would be integrated into one workplace, the population consisted of a broad range of mainly manufacturing users. For a real-world system, the results may not transfer in terms of thoughts towards the system.

5.6 Conclusion

In conclusion, a quality control task was utilised as the platform in which the assistive automation could attempt to aid the users. In the first experiment there were a number of hypotheses specified which related to automation trust and task performance. The second experiment altered the study design slightly to match the need to understand a personalisation implementation within a manufacturing scenario (to answer the third thesis research question).

The second experiment was the main driver behind providing research towards the objectives of thesis RQ3. The first objective was: to understand if the type of personalisation systems effects user performance. For a quality control environment, automated decision support systems of this nature would not provide significant improvements in task performance. The next objective was: to identify users perception of personalisation systems is varied dependant on the type implemented. The participants noted the usability of the Suggester system and seemed, at times, puzzled by the activation. For Controllers system, the high expectations of the system were a negative aspect of what was thought to be a generally positive system.

Additions findings were uncovered within the current chapter. From the experiment one results, it was not possible to accept that automation reliability affected automation performance. It was discussed this could be due to the type of task and the aid provided by the automation would not identify the defect, just the presence of a defect. It was also possible to accept that the automation would not be able to affect the behaviour of the participants with varying levels of automation reliability, in line with previous work in this area (Salem et al., 2015). In experiment two, the numerical analysis was mostly unable to decipher any differences between conditions. This was suggested to be related to the low activation of the system, such that for the vast majority of the task participants would not be engaging with the personalisation element.

Chapter 6: Discussion and Conclusion

6.1 Introduction

The current thesis presented an examination of personalisation within automated manufacturing environments. This started with the justification of the work through the current interest in Industry 4.0/5.0 paradigms. Personalisation was suggested to be an effective way of incorporating these paradigms into potential manufacturing environments. However, the context of what personalisation currently entails has varied over time and modern personalisation is almost always computational. While the thesis did not revolve around the technical approach to personalisation, this topic was summarised to provide additional insight into how personalisation functions. The methodology of the thesis to describe the use case of personalisation was discussed, and the use of three main empirical chapters was detailed: a codesign workshop study, a survey of end-users' acceptance of personalisation and the role of automated assistance in manufacturing tasks. These were majorly built upon using a novel Taxonomy of Personalisation systems. The initial chapter sets out the structured approach of the thesis, that being as a set of principles for system integrators or developers to use in their implantation of personalisation systems in the workplace.

The inclusion of a chapter featuring motivating work enhances the importance of the thesis. Working with a group of researchers (Marinescu et al., 2022) (Chapter 1: Introduction), the study consisted of a ContraVision method (Mancini et al., 2010), featuring two opposing viewpoints on future Digital Manufacturing Technologies. This study revealed that there are concerns around personal data, including how it is used or captured. It also described a wide array of negative feelings towards these types of technology. This forms the major emphases points of the thesis: how can personalisation be used to better improve the work lives of people rather than fulfilling their negative connotations and how user data should play a role in personalisation.

The motivating work is linked to the research questions devised for the thesis. These were: how would codesign be used to design personalisation within automated manufacturing environments, how accepting of personalisation are end-users and can personalisation impact an end-user's interaction within an automated system. Each of these questions is used as a template for the main experiments in the empirical chapters. The questions revolve around two components: the automated manufacturing sector and the person behind the system. This naturally links the thesis to the motivating work and the Industry 5.0 paradigms (European Commission. Directorate General for Research and Innovation., 2021).

The first empirical chapter described the codesign workshops. To study personalisation, the workshops were infused with a novel set of ideation cards which featured all the components to build personalisation systems (Personalisation Design Cards) (Duvnjak et al., 2024) (Chapter 3: Codesign with the Personalisation Design Cards). The cards featured the taxonomy categories as one type of card. The people who participated in the study utilised the cards to create personalisation systems, which they then described. The outcome of this study was a greater understanding of how people felt about personalisation, some of these findings matching those found in the motivating work. There were types of personalisation they felt would be best used in systems and concerns around data and the societal impact of these systems.

The second empirical chapter studied how end-users feel about acceptance of personalisation in potential workplace scenarios (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). To achieve this, an online experimental survey was used. This provided the benefits of being able to effectively reach a large number of people while simultaneously providing a way of gathering end-users' thoughts. The survey consisted of a vignette/s which described personalisation scenarios, characteristic scales and prior metrics of acceptance, which could then be statistically compared to other users or their own responses on other systems. The first experiment provided a number of insights into how the study design could be altered to provide more effective results. The second experiment showed that users feel more strongly about personalisation when different data types are utilised. Out of the three personalisation types represented from the taxonomy (Suggesters, Swappers, Controllers), Swappers harboured higher levels of acceptance, which were more prominent when heart rate was the chosen data type.

In automated manufacturing environments, it can be suggested that automated assistance can play a key role. Thus, this notion forms the basis for the last empirical chapter (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). The first experiment takes a representative manufacturing task and provides a simplistic form of automated assistance. This was a visual mark around artefacts to signify defects in a quality control task. As systems cannot always be fully reliable, the study's condition hinges on three reliability levels of the assistance. This revealed how users felt about the assistance. The second experiment expanded upon the automated assistance to provide greater functionality and tie this assistance to the users' own performance. One assistance system provided an upgraded (in comparison to the first experiment) visual mark (a Suggester system), with the other "completing" a set number of inspections for the user (a Controller system). The statistical data was largely inconclusive, showing that personalisation has minimal impact on the numerical completion of the task. However, the qualitative data from the experiment showed that users thought the Suggester had good usability and the Controllers had a strong judgemental element when the system provided assistance.

The role of data in the thesis must be emphasised. As aforementioned the placement of data in the structure of the empirical chapters. However, the results also reflected this. The motivating work and codesign study indicated a concern for certain data types (Chapters 1, 3) (Marinescu et al., 2022). The data types featured in the survey study appeared to affect how the systems were viewed (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) and the users of the personalisation implementations were not always certain of how their data was affecting the system (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). Data was then featured as a design topic in the discussion (Chapter 6: Discussion). While the thesis uncovered a number of factors of data, how these findings could change should be monitored.

With the empirical chapters detailed, there can be discussion with the perspective of answering the central questions of the thesis. As the target audience is system integrators, the discussion will accommodate this viewpoint. There will be a set of design insights into personalisation which can enable future development and implementation of these systems. To provide further insight, responses to the questions will be presented using the research generated by the studies completed throughout the thesis. In addition, over the PhD project, a number of limitations were identified. While these were mitigated where possible and this was detailed in the relevant chapter, there remain potential avenues to resolve the limitations and build upon the findings from the thesis. The future work is separated into time periods to provide discussion into how the personalisation may need to be studied across a period longer than a single follow up experiment. First, it is important to provide academic-focused answers to the three specified research questions. While these were answered in greater detail within each empirical chapter, a revaluation after all the studies have been described can provide additional context.

6.2 Research Questions

Beyond the overarching thesis question, the thesis specified three research questions (Chapter 1: Introduction). These were related to how end-users would design systems, discovering stakeholders' acceptance of personalisation and how personalisation would affect a manufacturing task. While each chapter provided its own discussion on the topic, each empirical chapter can be described in comparison to the other chapters. It should be possible to provide answers to each of the questions and this can be shown through the described objectives (Table 32).

Table 32 – How the research Questions and objectives were explored

Research Question	Objectives	Achieved Through
RQ1	To understand how users design personalisation systems within manufacturing environments.	Chapter 3, 4, 5
	To use the created designs to learn what types of personalisation is acceptable.	Chapter 3
	To develop a set of ideation cards (similar to prior work; Lucero and Arrasvuori, 2010).	Chapter 3
RQ 2	To explore how acceptance factors from prior literature	Chapter 4
	To understand if the type of personalisation system has an effect on acceptance.	Chapter 3, 4, 5
	To identify a link between acceptance and choice to use a system.	Chapter 4
RQ3	To understand if the type of personalisation systems effects user performance.	Chapter 5
	To identify users perception of personalisation systems is varied dependant on the type implemented.	Chapter 5

6.2.1 Research Question: Co-design of Personalisation

The current thesis aimed to find out how people would design personalisation systems, the theory being that the participants would design systems they themselves would want to use. This was mainly achieved through the third chapter (Chapter 3: Codesign with the Personalisation Design Cards) and this can be evidenced through the lens of the stated objectives. The first and second objectives related to design of personalisation and the participants of the workshops expressed a preference for 'dynamic' systems, which, in theory, could be represented by many types of personalisation system. In the case of the selected quotes, these arguably referred to Suggesters, Swappers and Controllers systems. In one example, the Controllers automation is provided with more autonomy in situations where users are new. Design insights were also found in other chapters such as the in the representative task (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). The participants, who can be generally assumed novices to the task, were critical of the Controllers system's engagement in the task. Thus, it can be suggested that not every design described in the workshops is directly implementable. The Personalisation Design Cards were a newly developed tool used to achieve the previous objectives in the codesign workshops (Chapter 3: Codesign with the Personalisation Design Cards) and their successful implementation completes the final objective of the research question. The answer to the overall research question can be stated: personalisation is wished to be designed to use certain paradigms (e.g. shortcuts) and use only certain non-invasive personal data types.

6.2.2 Research Question: Acceptance of Personalisation

The second research question to understand acceptance and personalisation can be discussed through the objectives. The first objective relating to acceptance factors was studied in the acceptance survey (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). The study found that the participant characteristics studied, Need for Cognition (Cacioppo & Petty, 1982) and Affinity for Technology (Edison & Geissler, 2003), did not have an effect on acceptance. This is an interesting finding, as one might expect this to be the case.

In terms of the type of personalisation and acceptance (the second objective), it found that users held higher acceptance for the Swappers example presented to them (in one data condition). The Swappers system having a higher acceptance is interesting, due to the fact that Controllers and Suggesters both fell below, creating a need for balance between types of personalisation. This - in terms of the taxonomy (Chapter 2: Taxonomy of Personalisation Systems)- presents the notion that the amount of intervention can affect acceptance. In this instance, too much intervention is a negative and too much can also be a hindrance for users. As aforementioned, the codesign workshop study (Chapter 3: Codesign with the Personalisation Design Cards) also contributed to answering this question by the discussion around data, finding that users did not like the idea of using all available data forms (this is also present in the motivating work (Chapter 1: Introduction)).

The final objective was an investigation into acceptance and choice to use a system. The experimental survey (Chapter 4: Survey to Measure Stakeholder

Acceptance of Personalisation Systems) was only able to link the ATTAA Usefulness (Van Der Laan et al., 1997) and the acceptance question (based on (Abraham et al., 2019). Thus, to answer the second thesis research question: users vary their acceptance (academically considered “Usefulness”) based on the type of personalisation and the implementation.

6.2.3 Research Question: Understanding Personalisation Implementation in Automated Systems

The final question was attempting to understand if personalisation would materially change how users completed tasks. The final study (Chapter 5: Implementing Automated Assistance in a Manufacturing Task) linked this question by allowing sets of participants to use two different types of personalisation system (with another group not having access to personalisation). This study was not able to find a concrete difference between how users performed in the tasks (objective one), with most of the difference lying in the qualitative answers participants provided. It appeared that users found the Controllers and Suggesters systems engaging with users sporadically a poor experience (objective two). This was not picked up on in prior studies and emphasises how important testing systems are. It did, however, appear as though Suggesters had a positive comment regarding usability compared to the Controllers condition’s general sense of positivity in regards to the system as a whole. To answer the research question, it is possible to state that while there is limited performance increase/decrease found in the current studies, the type of personalisation did appear to have a small impact on perception.

6.3 Design Recommendations for Personalisation Usage in Automated Manufacturing Environments

The title of the thesis presents the notion of using personalisation within automation manufacturing environments. Each of the studies addresses a different element of this concept. To best understand personalisation’s potential use case in manufacturing, a discussion around how to best use personalisation is warranted. The perspective taken will be that of a system integrator who is looking to incorporate new technologies into their workplace. The findings of the thesis can be placed into three categories, those related to personal data usage in workplace personalisation, how procedures could be affected and how the technology may benefit the end-users. There may be concerns from users about how this data is used (Chapters 1, 3). This is linked to the design of personalisation systems, as data forms a key component in how a system should be implemented. Personalisation may also affect procedures currently present in workplace environments and how people respond to them. This is built around acceptance of personalisation, as there are many factors that affect how users view personalisation that are outside of the immediate use of the technology for a task. The benefits of the use of personalisation can be defined from the completed research. With the last research question aiming to understand how personalisation can affect tasks, these final insights will resolve this question.

6.3.1 Data in the Workplace

Personalisation requires personal data to function correctly. Thus, an audit of what personal data could be available in the intended environment must be completed. It is no surprise that different workplaces may have different data types available for personalisation. A workplace that uses physical records may not track metrics in a form that could enable effective personalisation. In these cases, the first step would be suggesting basic forms of data capture that may lead to useful personalisation, or just provide a basic data-driven approach to the organisation. This could be the requesting of performance data capture, but this may be a significant hurdle to overcome. In contrast, if an organisation already captures some potentially usable personal data, understanding whether this is useable for the type of personalisation intended to be implanted should be a priority. In either case, the document of all data (currently available or potentially available) should be the first step in the design requirements process.

As discovered, users may not view all data types equally (Chapters 1, 3, 4). This will have implications for implementation. Physiological data types appear more controversial for users (Chapter 1, 3). This may be due to these data types are often not easily controllable by users themselves. For example, one may not always be aware that their heart rate is increasing or understand the reason why this is the case. If personalisation systems are being used based on physiological data, they may be adjusting the system in ways the user does not expect. The approach should be to discuss with potential users how they feel about current or planned data types. Ensuring the dialogue is open an honest will ensure that users feel comfortable in stating if they feel a data type oversteps what is felt to be appropriate. This should be included with a mindset to remove any data types that could be controversial. In total, this should lead to a set of available data types to include in a system's user requirements. Further, the use of certain data types affects how they view an implemented system. The surveys on acceptance showed that end-users felt more strongly towards the different personalisation systems when heart rate was the data type used (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). The use of heart rate may be a concern of users, who may feel they need to take more notice of how the system is using this data. It may also be the case that -as a system integrator- less care needs to be taken when using conventionally accepted data types (like performance) and more attention can be attributed towards effective implementation rather than user-held beliefs about the system.

There are, however, data types which users feel are acceptable and should be greatly featured. Participants of the workshops suggested using 'task experience' as a viable data type for personalisation (Chapter 3: Codesign with the Personalisation Design Cards). This was echoed with performance data, which lessened the impact of different types of personalisation on users' acceptance (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). Users may feel comfortable with personalisation, which makes use of experience and performance, as these types of data can be found in current workplaces and are -for the most part- directly controlled by users. A user could engage in training activities to boost their knowledge and experience while improving their performance through practice. In the implementation study (Chapter 5: Implementing Automated Assistance in a Manufacturing Task), the data type was not a main feature of the participants' discussion; most appeared to be accepting and it did not raise concerns (the bigger concern was in the understanding of the system). For the design of personalisation, using data types which users are generally more fond of, will make implementation of these systems easier. However, as some types of personalisation require more data (Chapter 2: Taxonomy of Personalisation Systems), attempting to use accepted forms of data may not be possible. Thus, it is important to balance potential concerns with the choice of system to be implemented. Failing this, communication with the workers about how the data will be chosen and aiming to resolve any potential concerns could mitigate future problems.

6.3.2 Workplace Procedures

End-users could feel sceptical about the insights generated from the systems (Chapter 3: Codesign with the Personalisation Design Cards). This is a tricky challenge to solve, as users may not have the technical knowledge to understand the system and even abstractions could confuse users or provide misleading impressions. Learning from a high-profile current mainstream system (BBC, 10 April), the approach should be the fair investigation of any potential problems with the outcomes of the investigation made available to end-users. As users have been shown to work out automation reliability (Chapter 5: Implementing Automated Assistance in a Manufacturing Task), users may be the first people to interact with a system they believe to be a misfunctioning system. By promoting open investigations in the case of issues, users will feel more confident in the system if it is working correctly and will trust the organisation more if they freely admit the system may make mistakes (the latter being the topic of discussion in mainstream news (BBC, 10 April)).

The current research found that the accuracy of the system is a potential concern (Chapter 3: Codesign with the Personalisation Design Cards). The utilisation of personalisation does rely on a level of data analysis. The reliability of data analysis can vary and often, not every variable can reasonably be analysed, meaning that all types of analysis could be incorrect in some form. The approach should be similar to that suggested above, allowing users to report problems easily and thoroughly working through these problems. In one study (Chapter 5: Implementing Automated Assistance in a Manufacturing Task), even though the system was functioning correctly, a user felt it was malfunctioning as they misunderstood how the system worked. It appeared that they believed the 'performance' being measured was their accuracy at the task and not their speed. However, another participant described the correct performance metric being measured. For system integrators, in situations when engaging with users to describe the functionality of the system, care should be taken to ensure a correct picture of the system is conveyed and potential irregularities are resolved. It is important to remember that not all users will be experts in systems design and will immediately understand the presented information.

Automated systems and their effect on work have been featured in mainstream media (Cellan-Jones, 2019) and it is not hard to imagine that certain people might concerned of 'job insecurity' (Dodel & Mesch, 2020). The current thesis was able to locate this notion from the participants featured. In the motivating work (Chapter 1: Introduction), the participants provided statements around the effect of Digital Manufacturing Technologies on Job Security. In instances where employees feel this may be a problem, the mitigating strategy would be to provide adequate training programs in order to alleviate these fears (one 'SME owner[~~s~~]' did use 'upskill[ing]' in Waldman-Brown, 2020). Further, personalisation has its own challenges regarding 'compensation' (Chapter 3: Codesign with the Personalisation Design Cards). This study revealed that 'tiredness' may provoke businesses to reduce the amount of compensation given to an employee. With a large amount of personal data collected, it could be a case that this links to the inaccuracy highlighted (Chapter 3: Codesign with the Personalisation Design Cards), where this compensation is based on faulty data. This is a situation where some employees may be more heavily impacted than others. For example, a participant referred to '[HGV] drivers' (Chapter 3: Codesign with the Personalisation Design Cards), which may not be a salaried position and thus, more susceptible to a decrease in pay. The resolve this potential issue may be difficult for an individual business and may rely on increased political regulation or current legislation around equality.

6.3.3 Benefits for Users

The importance of personalisation is its potential positive impact on users. As previously specified, the current thesis is motivated by Industry 5.0 paradigms and, thus, aims to make use of digital technologies to provide benefits to users (European Commission. Directorate General for Research and Innovation., 2021). With the current research completed, it is possible to answer how the use of personalisation can affect tasks in automated manufacturing environments. The viewpoint taken is that of how personalisation can be made to be beneficial to users. While not every research finding was positive, it is still possible to discuss how best these should be addressed in an organisation.

The first empirical chapter allowed participants to design their own personalisation systems (Chapter 3: Codesign with the Personalisation Design Cards), this showed how interactions with systems can be refined through the technology. One participant's suggestion of 'shortcuts' shows that commonly utilised functions for users can be made easier to access based on their personal data. This would reduce the time taken for users to reach key functions and improve their productivity. The participants also suggested making use of Swapper systems. In the provided examples, users would have the content changed to represent written or visual forms. 'Experience' was also portrayed as a potential personalisation data type. The work describes this as a 'Dynamic' system, in one system, the participant has created a Controllers system. By providing multiple thresholds on the "experience" of the user, the system can step in for beginners and the opposite for experienced users. This could be instrumental for organisations who may not have the capacity to monitor new employees but with a personalisation system, can have the redundancy of the system being able to minimize the risk of errors. As with educational personalisation systems (e.g. Maaliw, 2021), workplace systems could better represent the way in which users prefer content to be presented to them or what would make them more effective in their role. Further, the described systems include changing the 'language' for different users. This would aid in situations where users may have a greater understanding in one language than another. While this may appear minimal or inconsequential, as language choices already exist in many systems, using personal data to change this automatically could reduce aggravation caused by misinterpretations. This can have great positive benefits for safety or manufacturing tasks where there is a low tolerance for mistakes.

Personalisation was shown to have an almost imperceivable impact on performance (Chapter 5: Implementing Automated Assistance in a Manufacturing Task); while this could be portrayed as a negative, personalisation's impact could be in improving other aspects of the user experience to be an overall benefit to users. For example, the Swappers system was shown to have 'Strong Usability' (Chapter 5: Implementing Automated Assistance in a Manufacturing Task) in relation to how the assistance is provided to participants in manufacturing. Designers should make the best use of the low impact on productivity to utilise personalisation to create the best experience for users. Further, this may be a situation where the chosen study's design was not of the type that would best provide a detectable numerically different response. Future work may be able to further clarify this area.

6.4 Future Work

There were a number of potential avenues of exploration which were unable to be covered with the current project resources or limitations with the included studies. As the project focused on the user experience within the implementation, there is scope for future projects to continue pushing the boundaries of personalisation. For the sake of clarity, these can be thought of in terms of the Short, Medium and Long term. The short-term considerations directly follow the current research. As the research focused on three distinct questions and attempts to answer each, there is room to diverge into three separate paths of future work. These would cover the topics of design, user acceptance and development to mimic the research questions. The medium term will present a practical evolution of the topic and the long term will focus on a future in which personalisation is commonplace or widely accepted in the workplace.

6.4.1 Short Term

The current work looked at design from the perspective of a general population using a set of ideation cards (Personalisation Design Cards, PDCs) (Chapter 3: Codesign with the Personalisation Design Cards). There was a planned study to use design experts to use the cards to understand how they would design personalisation. In contrast to participatory design theory, which uses 'developers hav[ing] a disproportionate design impact' as a rationale for the approach (Muller, 1991). The use of design experts in this case would allow contribution in potential avenues for design even without familiarity of personalisation. As design is a diverse field, it is expected that not all would be familiar with personalisation and fewer would be familiar with personalisation in manufacturing settings. For example, a User Interface designer may have worked with prior systems that could benefit from a personalised approach that they would bring with them into a workshop. The PDC approach would -as with non-designers- provide the necessary information for them to generate innovative personalisation systems (bridging potential gaps in their knowledge). While the study was planned, a lack of resources meant this study fell out of the boundary of the current thesis. Future work should look to take this pathway to push forward the design side of personalisation.

The PDCs could also be used with industry professionals. While the majority of the included experiments are utilised with end-users, the original codesign workshops population was that of convenience; it would be interesting to understand how the stakeholders of the project would feel about personalisation. It can be assumed that the stakeholders who have worked in manufacturing environments can better visualise places where personalisation could be useful. This will generate greater levels of insight into the way personalisation can be designed. The PDCs may also be used by organisations looking to integrate personalisation. As a first step into how personalisation will function in their workplace, the PDCs could be an invaluable tool in gathering user requirements from potential end users. There would also be an interest in understanding how a general population varies in their designs of personalisation compared with end-users. These users may be more sceptical of certain types of personalisation systems being implemented as they are more likely to be impacted if personalisation systems were widely utilised.

While the current thesis examined the acceptance of personalisation within a range of vignettes acceptance, this provides only a broad understanding of the topic (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems). With the finding that Swappers system was the most accepted in the Heart Rate condition, there is the potential to explore why this was the case. A future study could allow a user to use a personalisation system in a manufacturing setting and discover whether they would continue to use it when provided the choice (similar to the acceptance question (Abraham et al., 2019) (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems)). With a limited level of resources for a project such as this, only three types of personalisation were featured. As current work is trending towards Generators systems (Chapter 2: Taxonomy of Personalisation Systems), a future study could examine this type of personalisation in comparison to others. Additionally, it may be possible to replace the fictional vignettes with existing systems to provide further validity to the results.

The final study looked at personalisation within a Quality Control task. In the included experiments, personalised assistance was not found to provide performance improvements. However, the different types of personalisation were viewed differently in terms of usability. The Quality Control task was conducted remotely, thus providing an interesting discussion point: how the personalisation would be best represented in this case. An in-person task allows for a different set of options for personalisation. For example, it would be more practical to collect data such as heart rate in person, as it can be expected that not many potential participants have the equipment or consistent methods to collect this data. The usage of in-person experiments would allow a wide range of study designs while also enabling a wider range of data types for personalisation.

As with the prior study, resource limitations meant it was only possible to represent a limited number of personalisation systems. One planned experiment would utilise a Swapper system to change how the tiles were presented to users by featuring more tiles in the defect condition they were stronger at and the inverse. Performance has already been used in systems similar to the suggested personalised-based task scheduling. Khazankin *et al.* (2011) present a ‘Skill-Aware’ system that optimises for skill-to-task matching, presenting tasks to users of an online ‘platform’ while conforming to this principle. The system gathers personal data from users themselves and through user performance. The authors were able to show that 'average job outcome quality' was stronger within the Skill-Aware approaches. The video game Tetris has an implemented system that makes the next available pieces related to a player's skill (Lora et al., 2016). By changing the order of the pieces, the experiment was able to show an increase in player performance. For this future work, the expectation should be that -by following a similar concept of changing the future science of work- the participants' performance could be increased.

With one study utilising personalisation, it was possible to discover a potential weakness of the technology. As discussed prior, participants noted that they were confused by the way the system was triggered during the quality control task (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). Future work could alleviate this by providing a warning before the personalisation system activates the assistance. As users felt like they were observed by the system, the additional warning may provoke these feelings further. Another direction could be in the amount of times the system activates. Additional experimentation should vary how and when the system activates to provide the best experience for the user.

6.4.2 Medium Term

In the medium term, collaboration with real-world industry professionals or organisations will be key to advancing the field of personalisation. There may be manufacturing organisations that already have an implementation of personalisation. Research can then be conducted to understand how their implementation functions, the impact on the users and what practical constraints were acknowledged in the development of their system. Where this is not viable, future work should look to find a viable business for the technology that is willing to provide resources for a small-scale trial. This would need to be a real situation or task but could be replicated outside of the business. For example, an organisation could provide the task and host a set of people to study how they use their existing tools (e.g. 'usability testing' (Krug, 2014, Chapter 9)). This would provide insight into the user requirements, which would lead to a personalisation prototype.

While the inclusion of industry would provide great strides for developing personalisation, the use of the technology is still relatively unexplored and may not be at a stage where a business wants the cost of -what is essentially-testing technology. To address this, future work in the longer term should look at demonstrating the effectiveness of personalisation in more complex tasks. It is unlikely that every task will be as simplistic as the quality control task (Chapter 5: Implementing Automated Assistance in a Manufacturing Task). Developing a complex task that users can complete would provide a concrete example of the use case of personalisation.

There is an interest in understanding how the type of implementation changes over time. As found in the systematic review (Chapter 2: Taxonomy of Personalisation Systems), there is an increasing amount of Generators-type systems appearing in academic work. There is the potential for a future review to take a view over the upcoming decade to understand if this shift towards one type of personalisation continues. It may be possible for Controller systems to become more popular with the evolution of technology or the availability of more data types. Existing work has discussed how a specific person’s feelings towards data ‘privacy’ may affect how systems are integrated (Chen et al., 2022) and changes in the broad opinions may cause the amount of one type of system to increase or decrease. For example, the current survey (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems) suggested Swapper systems being the most acceptable, and these were shown to be common in academic work (Chapter 2: Taxonomy of Personalisation Systems). If this changes, then there may be a shift towards a different type of personalisation.

Future evaluation of acceptance of personalisation should be considered. With prior work looking at personalisation in a general sense over a wide population (Kozyreva et al., 2021) and the current survey (Chapter 4: Survey to Measure Stakeholder Acceptance of Personalisation Systems), it is possible to describe how personalisation has or has not moved in the viewpoint of the people. Over the medium term, there may be an expectation that as personalisation becomes more commonplace, how it is perceived will differ.

6.4.3 Long Term

It is not possible to ascertain what the future of a technology will be without knowing what potential future breakthroughs could radically change the landscape. The technology may become standardised to such an extent that many “mainstream” products or services allow for personalisation as standard (imagine reaching the peak section of the Technology Adoption Life Cycle (Moore, 2006, p.12)). For example, computer-aided design software may provide personalised elements for users. However, it is possible to extrapolate current trends to provide a level of understanding as to what developments could occur moving into the future.

The future case could revolutionise training in organisations. For a User Experience designer, this may take the form of understanding a user’s current task and completing parts of the design in advance. By taking their behaviour data at previous times, the system could learn when to intervene to help the user before assistance is required (in an instance where a user forgets how to use a tool, they may have not used recently). This can be extrapolated across an entire team or organisation to provide guidance to users based on how their performance compares to others. Users may be able to be trained without requiring as much formal training by the system, suggesting the next completion steps.

The use of personalisation affects the future of the design side of Human Factors (HF). It has been previously stated that '[...] the designer, who usually must come up with a single design for everyone' is not a good or realistic approach (Norman, 2013, p.243). Personalisation reduces the need to use an 'average' user (Norman, 2013, p.243) for the design requirements, as each system will change to fit (theoretically) any user it encounters. In its place, the design focus should revolve around what elements of the interface can be changed with data. For example, does the interface need to be in a different hue to account for users who struggle to perceive certain colours or what performance data needs to be taken for automation to take over from a user in an instance of low performance? For HF designers, the role of the designer will need to be familiar with using and manipulating data to inform design in ways not expected in the current climate.

References

Abraham, M., Niessen, C., Schnabel, C., Lorek, K., Grimm, V., Möslein, K., & Wrede, M. (2019). Electronic monitoring at work: The role of attitudes, functions, and perceived control for the acceptance of tracking technologies. *Human Resource Management Journal*, *29*(4), 657–675. https://doi.org/10.1111/1748-8583.12250

Adomavicius, G., & Tuzhilin, A. (2005). Personalization technologies: A process-oriented perspective. *Communications of the ACM*, *48*(10), 83–90. https://doi.org/10.1145/1089107.1089109

Aggarwal, R., & Ranganathan, P. (2017). Common pitfalls in statistical analysis: Linear regression analysis. *Perspectives in Clinical Research*, *8*(2), 100. https://doi.org/10.4103/2229-3485.203040

Agnew, A., Forrester, P., Hassard, J., & Procter, S. (1997). Deskilling and reskilling within the labour process: The case of computer integrated manufacturing. *International Journal of Production Economics*, *52*(3), Article 3. https://doi.org/10.1016/S0925-5273(97)00092-3

Aguirre, E., Roggeveen, A. L., Grewal, D., & Wetzels, M. (2016). The personalization-privacy paradox: Implications for new media. *Journal of Consumer Marketing*, *33*(2), 98–110. https://doi.org/10.1108/JCM-06-2015-1458

Aheleroff, S., Mostashiri, N., Xu, X., & Zhong, R. Y. (2021). Mass Personalisation as a Service in Industry 4.0: A Resilient Response Case Study. *Advanced Engineering Informatics*, *50*, 101438. https://doi.org/10.1016/j.aei.2021.101438

Ahmed, R., Toscos, T., Rohani Ghahari, R., Holden, R. J., Martin, E., Wagner, S., Daley, C., Coupe, A., & Mirro, M. (2019). Visualization of Cardiac Implantable Electronic Device Data for Older Adults Using Participatory Design. *Applied Clinical Informatics*, *10*(04), 707–718. https://doi.org/10.1055/s-0039-1695794

Ahmed, S. K. (2025). Sample size for saturation in qualitative research: Debates, definitions, and strategies. *Journal of Medicine, Surgery, and Public Health*, *5*, 100171. https://doi.org/10.1016/j.glmedi.2024.100171

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Alshammari, M., Anane, R., & Hendley, R. J. (2016). Usability and Effectiveness Evaluation of Adaptivity in E-Learning Systems. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2984–2991. https://doi.org/10.1145/2851581.2892395

AMFG. (2019, March 28). *Industry 4.0: 7 Real-World Examples of Digitial Manufacturing in Action* [Business]. AMFG. https://amfg.ai/2019/03/28/industry-4-0-7-real-world-examples-of-digital-manufacturing-in-action/

An, S., Moore, R., Liu, E. Y., & Ren, G.-J. (2021). Recipient Design for Conversational Agents: Tailoring Agent's Utterance to User's Knowledge. *CUI 2021 - 3rd Conference on Conversational User Interfaces*, 1–5. https://doi.org/10.1145/3469595.3469625

Analyttica Datalab. (2021, November 13). *Understanding—'Wald Test'*. https://medium.com/@analyttica/understanding-wald-test-2e3fa7723516

Araujo, P., Costa, C., Viana, W., De Lima Veras, N., & Farias, E. J. P. (2020). Automatic Personalisation of Study Guides in Flipped Classroom: A Case Study in a Distributed Systems Course. *2020 IEEE Frontiers in Education Conference (FIE)*, 1–9. https://doi.org/10.1109/FIE44824.2020.9274186

Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., & Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. *International Journal of Human-Computer Studies*, *145*, 102522. https://doi.org/10.1016/j.ijhcs.2020.102522

Asaithambi, S. P. R., Venkatraman, S., & Venkatraman, R. (2021). Big Data and Personalisation for Non-Intrusive Smart Home Automation. *Big Data and Cognitive Computing*, *5*(1). https://doi.org/10.3390/bdcc5010006

Ashtari, N., Alamzadeh, P., Ganapathy, G., & Chilana, P. (2022). PONI: A Personalized Onboarding Interface for Getting Inspiration and Learning About AR/VR Creation. *Nordic Human-Computer Interaction Conference*, 1–14. https://doi.org/10.1145/3546155.3546642

Asthana, S., Singh, P., & Singh, A. (2013). Design and evaluation of adaptive interfaces for IVR systems. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, 1713. https://doi.org/10.1145/2468356.2468663

Awad, N. F., & Krishnan, M. S. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. *MIS Quarterly*, *30*(1), 13–28.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), Article 6. https://doi.org/10.1016/0005-1098(83)90046-8

Baldwin, S. L., & Rouleau, G. P. (1981). Office automation: From barriers to gateways. *Telecommunications Policy*, *5*(4), Article 4. https://doi.org/10.1016/0308-5961(81)90040-9

Barker, L. M., & Nussbaum, M. A. (2011). The effects of fatigue on performance in simulated nursing work. *Ergonomics*, *54*(9), 815–829. https://doi.org/10.1080/00140139.2011.597878

BBC. (10 April). Post Office scandal explained: What the Horizon saga is all about. *BBC*. https://www.bbc.co.uk/news/business-56718036

Beers, B. (2024, July 31). *Regression: Definition, Analysis, Calculation, and Example*. Investopedia. https://www.investopedia.com/terms/r/regression.asp

Belo, O., Correia, H., Rodrigues, P., & Barros, R. (2016). A personalization system for data visualization platforms. *2016 Sixth International Conference on Innovative Computing Technology (INTECH)*, 162–167. https://doi.org/10.1109/INTECH.2016.7845024

Benyon, D., & Murray, D. (1988). Experience with Adaptive Interfaces. *The Computer Journal*, *31*(5), 465–473. https://doi.org/10.1093/comjnl/31.5.465

Bernstein, M. S., Tan, D., Smith, G., Czerwinski, M., & Horvitz, E. (2010). Personalization via friendsourcing. *ACM Transactions on Computer-Human Interaction*, *17*(2), 1–28. https://doi.org/10.1145/1746259.1746260

Billsus, D., Brunk, C. A., Evans, C., Gladish, B., & Pazzani, M. (2002). Adaptive interfaces for ubiquitous web access. *Communications of the ACM*, *45*(5), Article 5. https://doi.org/10.1145/506218.506240

Biswas, P., Langdon, P., Sarangi, S., Mimrot, V., Kar, S., Ribeiro, V. J., Umadikar, J., & Prashant, S. (2014). Interface Personalization through Inclusive User Modelling Web Service. *Proceedings of the XV International Conference on Human Computer Interaction*, 1–5. https://doi.org/10.1145/2662253.2664632

Blackburn, N. N., Gardone, M., & Brown, D. S. (2023). Player-Centric Procedural Content Generation: Enhancing Runtime Customization by Integrating Real-Time Player Feedback. *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 10–16. https://doi.org/10.1145/3573382.3616069

Blom, J. (2000). Personalization: A taxonomy. *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, 313–314. https://doi.org/10.1145/633292.633483

Brand, W. (2018). *Visual doing: Applying visual thinking in your day to day business*. BIS Publishers.

Brandt, E., Binder, T., & Sanders, E. B.-N. (2012). Tools and techniques. In *Routledge International Handbook of Participatory Design*. Routledge. https://doi.org/10.4324/9780203108543.ch7

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), Article 2. https://doi.org/10.1191/1478088706qp063oa

Buganza, T., Trabucchi, D., & Pellizzoni, E. (2020). Limitless personalisation: The role of Big Data in unveiling service opportunities. *Technology Analysis & Strategic Management*, *32*(1), 58–70. https://doi.org/10.1080/09537325.2019.1634252

Bunt, A., Conati, C., & McGrenere, J. (2010). Mixed-Initiative Interface Personalization as a Case Study in Usable AI. *AI Magazine*, *30*(4), 58. https://doi.org/10.1609/aimag.v30i4.2264

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Canhoto, A. I., Keegan, B. J., & Ryzhikh, M. (2023). Snakes and Ladders: Unpacking the Personalisation-Privacy Paradox in the Context of AI-Enabled Personalisation in the Physical Retail Environment. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-023-10369-7

Cellan-Jones, R. (2019, June 26). Robots 'to replace up to 20 million factory jobs' by 2030. *BBC News*. https://www.bbc.com/news/business-48760799

Chang, J. C., Zhang, A. X., Bragg, J., Head, A., Lo, K., Downey, D., & Weld, D. S. (2023). CiteSee: Augmenting Citations in Scientific Papers with Persistent and Personalized Historical Context. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3544548.3580847

Chen, X., Sun, J., & Liu, H. (2022). Balancing web personalization and consumer privacy concerns: Mechanisms of consumer trust and reactance. *Journal of Consumer Behaviour*, *21*(3), 572–582. https://doi.org/10.1002/cb.1947

Cialdini, R. B. (2007). *Influence: The psychology of persuasion* (Rev. ed., [Nachdr.]). Collins.

Cingil, I., Dogac, A., & Azgin, A. (2000). A Broader Approach to Personalization. *Commun. ACM*, *43*(8), 136–141. https://doi.org/10.1145/345124.345168

Clayton, J. (2024, May 26). *'I was misidentified as shoplifter by facial recognition tech'*. BBC News. https://www.bbc.co.uk/news/technology-69055945

Corno, F., De Russis, L., & Monge Roffarello, A. (2020). TAPrec: Supporting the composition of trigger-action rules through dynamic recommendations. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 579–588. https://doi.org/10.1145/3377325.3377499

Correia, N., & Boavida, M. (2002, May). Towards an Integrated Personalization Framework: A Taxonomy and Work Proposals. *Proceedings of the AH'2002 Workshop on Personalization Techniques in Electronic Publishing*.

Costa, D., & Duarte, C. (2019). Personalized and Accessible TV Interaction for People with Visual Impairments. *Proceedings of the 16th International Web for All Conference*, 1–4. https://doi.org/10.1145/3315002.3317566

Crowson, M. (Producer). (2020, March 2). *Two-way repeated measures ANOVA in SPSS: one-within, one-between (March 2020)* [YouTube Video]. https://www.youtube.com/watch?v=KZHHAuTj8GU

Davis, F. (1985). *A Technology Acceptance Model for Empirically Testing New End-User Information Systems*.

Dede, M. İ. C., Büyüköztekin, T., Hanalioğlu, Ş., Işıkay, İ., & Berker, M. (2022). Enabling personalization of a robotic surgery procedure via a surgery training simulator. *Robotica*, 1–16. https://doi.org/10.1017/S0263574722001023

Demirbilek, O., & Demirkan, H. (2004). Universal product design involving elderly users: A participatory design model. *Applied Ergonomics*, *35*(4), 361–370. https://doi.org/10.1016/j.apergo.2004.03.003

Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, *5*(1), Article 1. https://doi.org/10.1007/s007790170019

Dixon, S. R., & Wickens, C. D. (2006). Automation Reliability in Unmanned Aerial Vehicle Control: A Reliance-Compliance Model of Automation Dependence in High Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(3), Article 3. https://doi.org/10.1518/001872006778606822

Dodel, M., & Mesch, G. S. (2020). Perceptions about the impact of automation in the workplace. *Information, Communication & Society*, *23*(5), Article 5. https://doi.org/10.1080/1369118X.2020.1716043

Duvnjak, J., Kucukyilmaz, A., & Houghton, R. (2024). *Understanding user needs of personalisation-based automated systems with development and application of novel ideation cards*. 15th International Conference on Applied Human Factors and Ergonomics (AHFE 2024). https://doi.org/10.54941/ahfe1005032

EBSE. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering* (EBSE Technical Report No. Version 2.3). Keele University and University of Durham.

Edison, S. W., & Geissler, G. L. (2003). Measuring attitudes towards general technology: Antecedents, hypotheses and scale development. *Journal of Targeting, Measurement and Analysis for Marketing*, *12*(2), 137–156. https://doi.org/10.1057/palgrave.jt.5740104

European Commission. Directorate General for Research and Innovation. (2021). *Industry 5.0: Towards a sustainable, human centric and resilient European industry.* Publications Office. https://data.europa.eu/doi/10.2777/308407

Faria, A. L., Paulino, T., & Badia, S. B. I. (2019). Comparing adaptive cognitive training in virtual reality and paper-pencil in a sample of stroke patients. *2019 International Conference on Virtual Rehabilitation (ICVR)*, 1–7. https://doi.org/10.1109/ICVR46560.2019.8994746

Foley Curley, K. (1984). Are there any real benefits from office automation? *Business Horizons*, *27*(4), Article 4. https://doi.org/10.1016/0007-6813(84)90055-7

Frank Gleeson & Vincent Hargaden. (2014, May). *A task based framework to improve knowledge work efficiency*. POMS 25th Annual Conference 2014, Atlanta, GA. https://www.researchgate.net/publication/279689894_A_task_based_framework_to_improve_knowledge_work_efficiency

Gajos, K., Christianson, D., Hoffmann, R., Shaked, T., Henning, K., Long, J. J., & Weld, D. S. (2005). Fast and Robust Interface Generation for Ubiquitous Applications. In M. Beigl, S. Intille, J. Rekimoto, & H. Tokuda (Eds.), *UbiComp 2005: Ubiquitous Computing* (Vol. 3660, pp. 37–55). Springer Berlin Heidelberg. https://doi.org/10.1007/11551201_3

Gajos, K. Z., Czerwinski, M., Tan, D. S., & Weld, D. S. (2006). Exploring the design space for adaptive graphical user interfaces. *Proceedings of the Working Conference on Advanced Visual Interfaces - AVI '06*, 201. https://doi.org/10.1145/1133265.1133306

Gasteiger, N., Hellou, M., & Ahn, H. S. (2021). Factors for Personalization and Localization to Optimize Human–Robot Interaction: A Literature Review. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-021-00811-8

Glöggler, M., & Ammenwerth, E. (2021). Development and Validation of a Useful Taxonomy of Patient Portals Based on Characteristics of Patient Engagement. *Methods of Information in Medicine*, *60*(S 01), e44–e55. https://doi.org/10.1055/s-0041-1730284

Goldsmith, R. E., & Freiden, J. B. (2004). Have it your way: Consumer attitudes toward personalized marketing. *Marketing Intelligence & Planning*, *22*(2), 228–239. https://doi.org/10.1108/02634500410525887

Greenfield, A. (2006). *Everywhere: The Dawning age of Ubiquitous Computing*. New Riders.

Gullà, F., Cavalieri, L., Ceccacci, S., Germani, M., & Bevilacqua, R. (2015). Method to Design Adaptable and Adaptive User Interfaces. In C. Stephanidis (Ed.), *HCI International 2015—Posters' Extended Abstracts* (Vol. 528, pp. 19–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-21380-4_4

Haas, M. W., & Hettinger, L. J. (2001). Current Research in Adaptive Interfaces. *The International Journal of Aviation Psychology*, *11*(2), 119–121. https://doi.org/10.1207/S15327108IJAP1102_01

Halskov, K., & Dalsgård, P. (2006). Inspiration card workshops. *Proceedings of the 6th ACM Conference on Designing Interactive Systems - DIS '06*, 2. https://doi.org/10.1145/1142405.1142409

Hancock, P. A. (2019). Some pitfalls in the promises of automated and autonomous vehicles. *Ergonomics*, *62*(4), Article 4. https://doi.org/10.1080/00140139.2018.1498136

Hart, C. (2018). *Doing a literature review: Releasing the research imagination* (2nd edition). Sage.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology* (Vol. 52, pp. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Heeger, D. (2006). *Signal Detection Theory*. https://www.cns.nyu.edu/~david/courses/perception/lecturenotes/sdt/sdt.html

Heerwegh, D., & Loosveldt, G. (2006). An experimental study on the effects of personalization, survey length statements, progress indicators, and survey sponsor logos in web surveys. *Journal of Official Statistics*, *22*, 191–210.

Ishikawa, C., Sakamura, K., & Maekawa, M. (1981, December). Adaption and personalization of VLSI-based computer architecture. *MICRO 14: Proceeding of the 14th Annual Workshop on Microprogramming*. MICRO 14. https://doi.org/10.5555/800075.802435

Jeevan, V. K. J., & Padhi, P. (2006). A selective review of research in content personalization. *Library Review*, *55*(9), 556–586. https://doi.org/10.1108/00242530610706761

Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H., & Rafferty, L. (2010). Using the Decision-Ladder to Add a Formative Element to Naturalistic Decision-Making Research. *International Journal of Human-Computer Interaction*, *26*(2–3), 132–146. https://doi.org/10.1080/10447310903498700

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, *4*(1), Article 1. https://doi.org/10.1207/S15327566IJCE0401_04

Jimison, H., Pavel, M., & Pavel, J. (2003). *Adaptive interfaces for home health*.

Kaaniche, N., Laurent, M., & Belguith, S. (2020). Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey. *Journal of Network and Computer Applications*, *171*, 102807. https://doi.org/10.1016/j.jnca.2020.102807

Kahneman, D. (2012). *Thinking, fast and slow*. Penguin Books.

Karen Grace-Martin. (n.d.-a). The difference between logistic and probit regression. *The Analysis Factor*. https://www.theanalysisfactor.com/the-difference-between-logistic-and-probit-regression/

Karen Grace-Martin. (n.d.-b). *When the hessian matrix goes wacky*. https://www.theanalysisfactor.com/wacky-hessian-matrix/

Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW)*, *7*(3–4), 167–185. https://doi.org/10.1023/A:1008689307411

KERIN, R. A. (1974). PERSONALIZATION STRATEGIES, RESPONSE RATE AND RESPONSE QUALITY IN A MAIL SURVEY. *Social Science Quarterly*, *55*(1), 175–181.

Khazankin, R., Psaier, H., Schall, D., & Dustdar, S. (2011). QoS-Based Task Scheduling in Crowdsourcing Environments. In P. P. Maglio, M. Weske, J. Yang, & M. Fantinato

(Eds.), *Service-Oriented Computing* (Vol. 6470, pp. 297–311). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-25535-9_20

Kim, H., Jung, S., & Ryu, G. (2020). A Study on the Restaurant Recommendation Service App Based on AI Chatbot Using Personalization Information. *International Journal of Advanced Culture Technology*, *8*(4), 263–270. https://doi.org/10.17703/IJACT.2020.8.4.263

Kobsa, A. (2007). Privacy-enhanced personalization. *Communications of the ACM*, *50*(8), 24–33. https://doi.org/10.1145/1278201.1278202

Kozyreva, A., Lorenz-Spreen, P., Hertwig, R., Lewandowsky, S., & Herzog, S. M. (2021). Public attitudes towards algorithmic personalization and use of personal data online: Evidence from Germany, Great Britain, and the United States. *Humanities and Social Sciences Communications*, *8*(1), 117. https://doi.org/10.1057/s41599-021-00787-w

Krug, S. (2014). *Don't make me think, revisited: A common sense approach to Web [and mobile] usability* (third edition). New Riders.

Kucirkova, N. (2018). A taxonomy and research framework for personalization in children's literacy apps. *Educational Media International*, *55*(3), 255–272. https://doi.org/10.1080/09523987.2018.1512446

Kulviwat, S., Bruner II, G. C., Kumar, A., Nasco, S. A., & Clark, T. (2007). Toward a unified theory of consumer acceptance technology. *Psychology and Marketing*, *24*(12), 1059–1084. https://doi.org/10.1002/mar.20196

Lallé, S., & Conati, C. (2019). The role of user differences in customization: A case study in personalization for infovis-based content. *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 329–339. https://doi.org/10.1145/3301275.3302283

Lammert, F. B. (2021). *Comparison of strategies in the organisation of visual inspection work* [MSc Thesis]. University of Nottingham.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, 6(4), Article 4. https://doi.org/10.1007/s12599-014-0334-4

Lavie, T., Sela, M., Oppenheim, I., Inbar, O., & Meyer, J. (2010). User attitudes towards news content personalization. *International Journal of Human-Computer Studies*, *68*(8), 483–495. https://doi.org/10.1016/j.ijhcs.2009.09.011

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), Article 1. https://doi.org/10.1518/hfes.46.1.50_30392

Leesakul, N., Oostveen, A.-M., Eimontaite, I., Wilson, M. L., & Hyde, R. (2022). Workplace 4.0: Exploring the Implications of Technology Adoption in Digital Manufacturing on a Sustainable Workforce. *Sustainability*, *14*(6), 3311. https://doi.org/10.3390/su14063311

Leggett, T. (2019, October 25). Boeing 737 Max Lion Air crash caused by series of failures. *BBC*. https://www.bbc.co.uk/news/business-50177788

Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., Wuest, T., Mourtzis, D., & Wang, L. (2022). Industry 5.0: Prospect and retrospect. *Journal of Manufacturing Systems*, *65*, 279–295. https://doi.org/10.1016/j.jmsy.2022.09.017

Letsu-Dake, E., & Ntuen, C. A. (2009). A case study of experimental evaluation of adaptive interfaces. *International Journal of Industrial Ergonomics*, *40*(1), 34–40. https://doi.org/10.1016/j.ergon.2009.10.002

Li, L. (2024). Reskilling and Upskilling the Future-ready Workforce for Industry 4.0 and Beyond. *Information Systems Frontiers*, *26*(5), 1697–1712. https://doi.org/10.1007/s10796-022-10308-y

Lochner, M., Duenser, A., & Sarker, S. (2019). Trust and Cognitive Load in semi-automated UAV operation. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 437–441. https://doi.org/10.1145/3369457.3369509

Lora, D., Sánchez-Ruiz, A., Gonzalez-Calero, P., & Gómez-Martín, M. (2016, March). Dynamic Difficulty Adjustment in Tetris. *Proceedings of the Twenty-Ninth International Florida Artifical Intelligence Research Society Conference*. International Florida Artifical Intelligence Research Society Conference. https://cdn.aaai.org/ocs/12824/12824-57660-1-PB.pdf

Lucero, A., & Arrasvuori, J. (2010). *PLEX Cards*: A source of inspiration when designing for playfulness. *Proceedings of the 3rd International Conference on Fun and Games - Fun and Games '10*, 28–37. https://doi.org/10.1145/1823818.1823821

Lyu, S., & Bidarra, R. (2023). Procedural generation of challenges for personalized gait rehabilitation. *Proceedings of the 18th International Conference on the Foundations of Digital Games*, 1–11. https://doi.org/10.1145/3582437.3582468

Maaliw, R. R. (2021). A Personalized Virtual Learning Environment Using Multiple Modeling Techniques. *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0008–0015. https://doi.org/10.1109/UEMCON53757.2021.9666645

Madeira, R. N., Santos, P. A., Vieira, A., & Correia, N. (2014). Model-Based Solution for Personalization of the User Interaction in Ubiquitous Computing. *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*, 311–318. https://doi.org/10.1109/UIC-ATC-ScalCom.2014.139

Malik, Z. K., & Fyfe, C. (2012). Review of Web Personalization. *Journal of Emerging Technologies in Web Intelligence*, *4*(3), 285–296. https://doi.org/10.4304/jetwi.4.3.285-296

Manber, U., Patel, A., & Robison, J. (2000). Experience with personalization of Yahoo! *Communications of the ACM*, *43*(8), 35–39. https://doi.org/10.1145/345124.345136

Mancini, C., Rogers, Y., Bandara, A. K., Coe, T., Jedrzejczyk, L., Joinson, A. N., Price, B. A., Thomas, K., & Nuseibeh, B. (2010). Contravision: Exploring users' reactions to futuristic technology. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 153–162. https://doi.org/10.1145/1753326.1753350

Marinescu, A., Argyle, E. M., Duvnjak, J., Wilson, M. L., Lawson, G., Sharples, S., Hubbard, E., & Justham, L. (2022). The future of manufacturing: Utopia or dystopia? *Human Factors and Ergonomics in Manufacturing & Service Industries*, hfm.20976. https://doi.org/10.1002/hfm.20976

Martin, J.-C., & Clavel, C. (2018). Tailored, Multimodal and Opportune Interactions on a Wearable Sport Coach: The WE-nner Framework. In T. Clemmensen, V. Rajamanickam, P. Dannenmann, H. Petrie, & M. Winckler (Eds.), *Global Thoughts, Local Designs* (Vol. 10774, pp. 24–32). Springer International Publishing. https://doi.org/10.1007/978-3-319-92081-8_3

Mesquita, C., Diniz, S., Barbosa, S., Jos, C., & Lucena, P. (2002). *Towards the Identification of Concerns in Personalization Mechanisms Via Scenarios*.

Miller, D., Johns, M., Mok, B., Gowda, N., Sirkin, D., Lee, K., & Ju, W. (2016). Behavioral Measurement of Trust in Automation: The Trust Fall. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 1849–1853. https://doi.org/10.1177/1541931213601422

Mitchell, C., Cler, G., Fager, S., Contessa, P., Roy, S., De Luca, G., Kline, J., & Vojtech, J. (2022). Ability-based Keyboards for Augmentative and Alternative Communication: Understanding How Individuals' Movement Patterns Translate to More Efficient Keyboards: Methods to Generate Keyboards Tailored to User-specific Motor Abilities. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1–7. https://doi.org/10.1145/3491101.3519845

Mobasher, B., Cooley, R., & Srivastava, J. (2000). Automatic personalization based on Web usage mining. *Communications of the ACM*, *43*(8), 142–151. https://doi.org/10.1145/345124.345169

Montague, K., Hanson, V. L., & Cobley, A. (2011). Adaptive Interfaces: A Little Learning is a Dangerous Thing… In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Design for All and eInclusion* (pp. 391–399). Springer Berlin Heidelberg.

Montgomery, A. L., & Smith, M. D. (2009). Prospects for Personalization on the Internet. *Journal of Interactive Marketing*, *23*(2), 130–137. https://doi.org/10.1016/j.intmar.2009.02.001

Monzer, C., Moeller, J., Helberger, N., & Eskens, S. (2020). User Perspectives on the News Personalisation Process: Agency, Trust and Utility as Building Blocks. *Digital Journalism*, *8*(9), 1142–1162. https://doi.org/10.1080/21670811.2020.1773291

Moore, G. A. (2006). *Crossing the chasm: Marketing and selling disruptive products to mainstream customers* (rev. ed., first Collins business essentials ed). Collins Business Essentials.

Morys-Carter, W. L. (2021). *Participant IDs for Pavlovia* [VESPR]. https://moryscarter.com/vespr/pavlovia.php

Muller, M. J. (1991). PICTIVE---an exploration in participatory design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Reaching through Technology - CHI '91*, 225–231. https://doi.org/10.1145/108844.108896

Mulvenna, M. D., Anand, S. S., & Büchner, A. G. (2000). Personalization on the Net using Web mining: Introduction. *Communications of the ACM*, *43*(8), 122–125. https://doi.org/10.1145/345124.345165

Nahavandi, S. (2019). Industry 5.0—A Human-Centric Solution. *Sustainability*, *11*(16), 4371. https://doi.org/10.3390/su11164371

NASA. (n.d.). *NASA TLX: PAPER/PENCIL VERSION*. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. Retrieved 18 July 2023, from https://humansystems.arc.nasa.gov/groups/TLX/tlxpaperpencil.php

National Careers Service (UK Gov). (n.d.). *Car manufacturing worker* [Goverment]. National Careers Service. https://nationalcareers.service.gov.uk/job-profiles/car-manufacturing-worker

Nederhof, A. J. (1983). Effects of repetition and consistency of personalization treatments on response rate in mail surveys. *Social Science Research*, *12*(1), 1–9. https://doi.org/10.1016/0049-089X(83)90006-6

Ni, A., & Kehtamavaz, N. (2023). A Real- Time Smartphone App for Field Personalization of Hearing Enhancement by Adaptive Dynamic Range Optimization. *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1–6. https://doi.org/10.1109/HORA58378.2023.10156710

Niu, J., Geng, H., Zhang, Y., & Du, X. (2018). Relationship between automation trust and operator performance for the novice and expert in spacecraft rendezvous and docking (RVD). *Applied Ergonomics*, *71*, 1–8. https://doi.org/10.1016/j.apergo.2018.03.014

Nodder, C. (2013). *Evil by design: Interaction design to lead us into temptation*. Wiley.

Nordhoff, S., de Winter, J., Madigan, R., Merat, N., van Arem, B., & Happee, R. (2018). User acceptance of automated shuttles in Berlin-Schöneberg: A questionnaire study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 843–854. https://doi.org/10.1016/j.trf.2018.06.024

Norman, D. (1990). The ‘problem ’ with automation: Inappropriate feedback and interaction, not ‘over-automation’. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), Article 1241. https://doi.org/10.1098/rstb.1990.0101

Norman, D. A. (1998). *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution*. MIT Press.

Norman, D. A. (2013). *The Design of Everyday Things* (Revised and expanded editon). The MIT Press.

Office for National Statistics. (2013). *Internet Access—Households and individuals, Great Britain: 2012 part 2*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2013-02-28#internet-activities

Office for National Statistics. (2019). *Internet Access—Households and individuals, Great Britain: 2019*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2019

Okoli, C., & Schabram, K. (2010). *A Guide to Conducting Systematic Literature Review of Information Systems Research* (Sprouts Working Papers on Information Systems).

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Pardos, Z. A., Tang, S., Davis, D., & Le, C. V. (2017). Enabling Real-Time Adaptivity in MOOCs with a Personalized Next-Step Recommendation Framework. *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale*, 23–32. https://doi.org/10.1145/3051457.3051471

Pashler, H. E., Wixted, J., & Stevens, S. S. (2004). *Stevens' handbook of experimental psychology* (3rd ed). Wiley. https://books.google.co.uk/books?id=ptbcCBSWvvQC&dq=signal+detection+theory&lr=&source=gbs_navlinks_s

Pavlich-Mariscal, J., Uribe, Y., Barrera-León, L. F., Mejia-Molina, N. A., Carrillo-Ramos, A., Pomares Quimbaya, A., Vicari, R. M., Fabregat, R., & Baldiris, S. (2015). An Architecture for Dynamic and Adaptive User Activity Planning Systems: *Proceedings of the 11th International Conference on Web Information Systems and Technologies*, 228–235. https://doi.org/10.5220/0005430302280235

Pawar, U., O'Shea, D., Rea, S., & O'Reilly, R. (2020). Explainable AI in Healthcare. *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, 1–2. https://doi.org/10.1109/CyberSA49311.2020.9139655

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Postma, O. J., & Brokke, M. (2002). Personalisation in practice: The proven effects of personalisation. *Journal of Database Marketing & Customer Strategy Management*, *9*(2), 137–142. https://doi.org/10.1057/palgrave.jdm.3240069

R. A. Bailey & Peter J. Cameron. (2003, May 30). *Latin squares in experimental design*. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwinxdPmr7H-AhVThlwKHXheBQsQFnoECCcQAQ&url=https%3A%2F%2Fwebspace.maths.qmul.ac.uk%2Fl.h.soicher%2Fdesigntheory.org%2Flibrary%2Fencyc%2Ftopics%2Flsed.pdf&usg=AOvVaw1u50ZPfhwSv6_eRjlIBFfv

Realmuto, J., Klute, G., & Devasia, S. (2019). Preliminary Investigation of Symmetry Learning Control for Powered Ankle-Foot Prostheses. *2019 Wearable Robotics Association Conference (WearRAcon)*, 40–45. https://doi.org/10.1109/WEARRACON.2019.8719630

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford University Press. http://site.ebrary.com/id/10212166

Regulation (EU) 2016/679, European Parliament and of the Council (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Reza, M., Kim, J., Bhattacharjee, A., Rafferty, A. N., & Williams, J. J. (2021). The MOOClet Framework: Unifying Experimentation, Dynamic Improvement, and Personalization in Online Courses. *Proceedings of the Eighth ACM Conference on Learning @ Scale*, 15–26. https://doi.org/10.1145/3430895.3460128

Rieth, M., & Hagemann, V. (2022). Automation as an equal team player for humans? – A view into the field and implications for research and practice. *Applied Ergonomics*, *98*, 103552. https://doi.org/10.1016/j.apergo.2021.103552

Robert J. Houghton, Jessica Wardlaw, & James Sprinks. (2016). *D7.3 Report on Data Validation Tests by Citizen Scientific Users* (p. 73) [Project Deliverable]. UNOTT (UCL). http://www.i-mars.eu/documents/d7-3.pdf

Robson, C. (2011). *Real world research: A resource for users of social research methods in applied settings* (3. ed). Wiley.

Roser, M., Ritchie, H., & Mathieu, E. (2023). What is Moore's Law? *Our World in Data*.

Rudian, S., & Pinkwart, N. (2021). Generating adaptive and personalized language learning online courses in Moodle with individual learning paths using templates. *2021 International Conference on Advanced Learning Technologies (ICALT)*, 53–55. https://doi.org/10.1109/ICALT52272.2021.00024

Sailaja, N., Crabtree, A., Colley, J., Gradinar, A., Coulton, P., Forrester, I., Kerlin, L., & Stenton, P. (2019). The Living Room of the Future. *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, 95–107. https://doi.org/10.1145/3317697.3323360

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 141–148. https://doi.org/10.1145/2696454.2696497

Schiaffino, S., & Amandi, A. (2004). User – interface agent interaction: Personalization issues. *International Journal of Human-Computer Studies*, *60*(1), 129–148. https://doi.org/10.1016/j.ijhcs.2003.09.003

Schnabel, T., Amershi, S., Bennett, P. N., Bailey, P., & Joachims, T. (2020). The Impact of More Transparent Interfaces on Behavior in Personalized Recommendation. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 991–1000. https://doi.org/10.1145/3397271.3401117

See, J. E., Drury, C. G., Speed, A., Williams, A., & Khalandi, N. (2017). The Role of Visual Inspection in the 21st Century. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 262–266. https://doi.org/10.1177/1541931213601548

Severance, C. R. (2016). *Python for everybody: Exploring data using Python 3* (First complete Python 3.0 version). Charles Severance.

Shen, J., Hu, J., Dudley, J. J., & Kristensson, P. O. (2022). Personalization of a Mid-Air Gesture Keyboard using Multi-Objective Bayesian Optimization. *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 702–710. https://doi.org/10.1109/ISMAR55827.2022.00088

Shinohara, A. C., da Silva, E. H. D. R., de Lima, E. P., Deschamps, F., & da Costa, S. E. G. (2017). Critical Success Factors for Digital Manufacturing Implementation in the Context of Industry 4.0. *IISE Annual Conference. Proceedings*, 199–204. ProQuest Central; SciTech Premium Collection.

Shinohara, K., Jacobo, N., Pratt, W., & Wobbrock, J. O. (2020). Design for Social Accessibility Method Cards: Engaging Users and Reflecting on Social Scenarios for Accessible Design. *ACM Transactions on Accessible Computing*, *12*(4), 1–33. https://doi.org/10.1145/3369903

Silva, K. G. G. H., Abeyasekare, W. A. P. S., Dasanayake, D. M. H. E., Nandisena, T. B., Kasthurirathna, D., & Kugathasan, A. (2021). Dynamic User Interface Personalization Based on Deep Reinforcement Learning. *2021 3rd International Conference on Advancements in Computing (ICAC)*, 25–30. https://doi.org/10.1109/ICAC54203.2021.9671076

Skatova, A., Johal, J., Houghton, R., Mortier, R., Bhandari, N., Lodge, T., Wagner, C., Goulding, J., Crowcroft, J., & Madhavapeddy, A. (2013). Perceived risks of personal data sharing. *Proc. Digital Economy: Open Digital*.

Spinuzzi, C. (2005). The Methodology of Participatory Design. *Technical Communication*, *52*, 163–174.

Statistics How To. (n.d.). *Wald Test: Defintion, Examples, Running the Test*. https://www.statisticshowto.com/wald-test/

Stohr, M., Schneider, M., & Henkel, C. (2018). Adaptive Work Instructions for People with Disabilities in the Context of Human Robot Collaboration. *2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*, 301–308. https://doi.org/10.1109/INDIN.2018.8472070

Strousopoulos, P., Papakostas, C., Troussas, C., Krouska, A., Mylonas, P., & Sgouropoulou, C. (2023). SculptMate: Personalizing cultural heritage experience using fuzzy weights. *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, 397–407. https://doi.org/10.1145/3563359.3596667

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, *36*, 298–322. https://doi.org/10.1111/j.1468-2958.2010.01377.x

Sutherland, R. (2019). *Alchemy: The surprising power of ideas that don't make sense*. W H Allen, an imprint of Ebury Publishing.

Talkad Sukumar, P., Avellino, I., Remy, C., DeVito, M. A., Dillahunt, T. R., McGrenere, J., & Wilson, M. L. (2020). Transparency in Qualitative Research: Increasing Fairness in the

CHI Review Process. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3334480.3381066

Tam & Ho. (2006). Understanding the Impact of Web Personalization on User Information Processing and Decision Outcomes. *MIS Quarterly*, *30*(4), 865. https://doi.org/10.2307/25148757

Taviss, I. (1972). A Survey of Popular Attitudes toward Technology. *Technology and Culture*, *13*(4), 606. https://doi.org/10.2307/3102838

Teo, C. H., Nassif, H., Hill, D., Srinivasan, S., Goodman, M., Mohan, V., & Vishwanathan, S. V. N. (2016). Adaptive, Personalized Diversity for Visual Discovery. *Proceedings of the 10th ACM Conference on Recommender Systems*, 35–38. https://doi.org/10.1145/2959100.2959171

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness* (Revised edition, new international edition). Penguin Books.

The Role of Automation and Humans in Nuclear Power Plants (No. IAEA-TECDOC-668; p. 176). (1992). International Atomic Energy Agency. https://www-pub.iaea.org/MTCD/Publications/PDF/te_668_web.pdf

Tiihonen, J., & Felfernig, A. (2017). An introduction to personalization and mass customization. *Journal of Intelligent Information Systems, 2017-08, Vol.49 (1), p.1-7*. https://doi.org/10.1007/s10844-017-0465-4

Toch, E., Wang, Y., & Cranor, L. F. (2012). Personalization and privacy: A survey of privacy risks and remedies in personalization-based systems. *User Modeling and User-Adapted Interaction*, *22*(1–2), 203–220. https://doi.org/10.1007/s11257-011-9110-z

Torn, I. A. R., & Vaneker, T. H. J. (2019). Mass Personalization with Industry 4.0 by SMEs: A concept for collaborative networks. *Procedia Manufacturing*, *28*, 135–141. https://doi.org/10.1016/j.promfg.2018.12.022

van de Garde-Perik, E., Markopoulos, P., de Ruyter, B., Eggen, B., & Ijsselsteijn, W. (2008). Investigating Privacy Attitudes and Behavior in Relation to Personalization. *Social Science Computer Review*, *26*(1), 20–43. https://doi.org/10.1177/0894439307307682

Van Der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, *5*(1), 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3

Visual Journalist Team. (2023, July 10). A Simple Guide to help you understand AI. *BBC*. https://www.bbc.co.uk/news/resources/idt-74697280-e684-43c5-a782-29e9d11fecf3

Volokh, E. (2000). Personalization and Privacy. *Commun. ACM*, *43*(8), 84–88. https://doi.org/10.1145/345124.345155

Waldman-Brown, A. (2020). Redeployment or robocalypse? Workers and automation in Ohio manufacturing SMEs. *Cambridge Journal of Regions, Economy and Society*, *13*(1), Article 1. https://doi.org/10.1093/cjres/rsz027

Weinmann, M., Schneider, C., & Robra-Bissantz, S. (2013, August 15). A Taxonomy of Web Personalization. *Proceedings of the Nineteenth Americas Conference on Information Systems*.

Weiser, M. (1991, September). The Computer for the 21st Century. *Scientific American*. https://moodle.nottingham.ac.uk/pluginfile.php/7815753/mod_resource/content/1/Weiser-SciAm.pdf

Wesson, J. L., Singh, A., & van Tonder, B. (2010). Can Adaptive Interfaces Improve the Usability of Mobile Applications? In P. Forbrig, F. Paternó, & A. Mark Pejtersen (Eds.), *Human-Computer Interaction* (pp. 187–198). Springer Berlin Heidelberg.

Wetzel, R., Rodden, T., & Benford, S. (2017). Developing Ideation Cards for Mixed Reality Game Design. *Transactions of the Digital Games Research Association*, *3*(2). https://doi.org/10.26503/todigra.v3i2.73

Xiong, X., Wang, Y., & Beck, J. B. (2015). Improving students' long-term retention performance: A study on personalized retention schedules. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, 325–329. https://doi.org/10.1145/2723576.2723636

Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. *Journal of Manufacturing Systems*, *61*, 530–535. https://doi.org/10.1016/j.jmsy.2021.10.006

Yamamoto, Y., & Yamamoto, T. (2020). Personalization Finder: A Search Interface for Identifying and Self-controlling Web Search Personalization. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*, 37–46. https://doi.org/10.1145/3383583.3398519

Yandi, L. (2020, April 21). [Online post]. ResearchGate. https://www.researchgate.net/post/Is_there_a_non-parametric_equivalent_of_a_two_way_ANOVA#:~:text=For%20nonparametric%20data%20(without%20normal,you%20should%20use%20Friedman%20test.)

Yang, F., Kalloori, S., Chalumattu, R., & Gross, M. (2022). Personalized Information Retrieval for Touristic Attractions in Augmented Reality. *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 1613–1616. https://doi.org/10.1145/3488560.3502194

Yang, S., Page, T., Zhang, Y., & Zhao, Y. F. (2020). Towards an automated decision support system for the identification of additive manufacturing part candidates. *Journal of Intelligent Manufacturing*, *31*(8), 1917–1933. https://doi.org/10.1007/s10845-020-01545-6

Yao, Y., Basdeo, J. R., Kaushik, S., & Wang, Y. (2019). Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3290605.3300428

Yeung, K. (2018). Five Fears About Mass Predictive Personalisation in an Age of Surveillance Capitalism. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3266800

Yin, Y., Stecke, K. E., & Li, D. (2018). The evolution of production systems from Industry 2.0 through Industry 4.0. *International Journal of Production Research*, *56*(1–2), 848–861. https://doi.org/10.1080/00207543.2017.1403664

Yu, P., Ahmad, W. U., & Wang, H. (2018). Hide-n-Seek: An Intent-aware Privacy Protection Plugin for Personalized Web Search. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 1333–1336. https://doi.org/10.1145/3209978.3210180

Yu, P. S. (1999). Data mining and personalization technologies. *Proceedings. 6th International Conference on Advanced Systems for Advanced Applications*, 6–13. https://doi.org/10.1109/DASFAA.1999.765731

Yuan, H., & Hernandez, A. A. (2023). User Cold Start Problem in Recommendation Systems: A Systematic Review. *IEEE Access*, *11*, 136958–136977. https://doi.org/10.1109/ACCESS.2023.3338705

Zakoldaev, D. A., Korobeynikov, A. G., Shukalov, A. V., Zharinov, I. O., & Zharinov, O. O. (2020). Industry 4.0 vs Industry 3.0: The role of personnel in production. *IOP Conference Series: Materials Science and Engineering*, *734*(1), 012048. https://doi.org/10.1088/1757-899X/734/1/012048

Zepf, S., Gupta, A., Krämer, J.-P., & Minker, W. (2020). EmpathicSDS: Investigating Lexical and Acoustic Mimicry to Improve Perceived Empathy in Speech Dialogue Systems. *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–9. https://doi.org/10.1145/3405755.3406125

Zhang, B., & Sundar, S. S. (2019). Proactive vs. reactive personalization: Can customization of privacy enhance user experience? *International Journal of Human-Computer Studies*, *128*, 86–99. https://doi.org/10.1016/j.ijhcs.2019.03.002

Zheng, P., Yu, S., Wang, Y., Zhong, R. Y., & Xu, X. (2017). User-experience Based Product Development for Mass Personalization: A Case Study. *Procedia CIRP*, *63*, 2–7. https://doi.org/10.1016/j.procir.2017.03.122

Appendices

Appendix A – Motivating Work Method and Results

1 Method

The qualitative data was gathered from the included questionnaire's open-ended questions. Of the total participants, N = 165, not all answered or responded adequately and were removed, N = 66. Four were removed due to the aforementioned lack of time spent answering the question. The remaining participants, N = 62 (Marinescu et al., 2022), had their data analysed per condition of video content (Utopic or Dystopic).

Participants were requested to watch one of two videos in a between-subjects design (Figure 31). The videos were based on a ContraVision approach (Mancini et al., 2010), which presented the Utopic and Dystopic variants of how future Digital Manufacturing Technologies could be used. The Utopic video is more positive in tone, with the narrator stating that "a fair amount of rest" can be provided by DMTs. The Dystopic video contained content more negative in tone, with comment about incorrectly set up machinery due to poor transposing of data. Participants were then required to answer multiple questions, the latter of these were two "free text" questions that allowed participants to provide additional detail to their thoughts about the video they viewed. These free text questions form the basis for the thematic analysis featured.

Figure 31 – Image of the Dystopian Video showing a manufacturing environment

The videos were created by the DigiTOP team and Albino Mosquito.

2 Results: Thematic Analysis

In addition to the quantitative analysis completed by the collaborators (Marinescu et al., 2022), the data was further analysed with a qualitative analysis. Guided by Braun and Clarke's (2006) approach to thematic analyses, data was coded, the codes were examined and where similarities were found, sorted into themes and accompanying sub-themes. For clarity, a tabular form rather than a visual diagram was utilised (Table 33, 3) with examples of quotes which match the respective sub-theme. The thematic analyses for both conditions were then able to be discussed individually and in comparison.

2.1 Utopic Condition

The analysis of the Utopic video condition produced three major themes. Participants showed themes of apprehension for Personal Data and the effect on People. The final theme presents a more neutral outlook for the Future with DMTs, perhaps reflecting the nature of future-based video content.

Table 33 – Utopic Tabular Thematic Analysis

Three themes are presented with accompanying sub-themes. Quotes (or relevant segments) from participants are included in the form they were captured. Participant numbers are reset for each condition; thus, a participant can have the same number as a different participant in the alternative condition. Quotes are presented in form provided by participants, unless indicated.

Themes	Sub-Themes	Quotes
Personal Data	Capture	Participant 22: What kind of dystopian ideal is this? Forcing workers to be biometrically monitored during their time at work?

		Participant 62: OK with heart and breathing rate being recorded, find brain activity a bit creepy
	Usage	Participant 59: If the data is shared with an employer, i think its too much data shared with others. Participant 19: How are our own personal (out of work) factors taken in to consideration in a way that does not prejudice the analysis.
People	Possibility of Discrimination	Participant 48: I think if u measure people`s heart rate etc in jobs, some people will never be able to get a job. Participant 63: These technologies are not for everyone so some talent might be lost if this is not recognised
	Degrading of Human Values	Participant 57: It scares me that we are becoming a society that will depend mainly on data. I fear that it will remove any connection to reality. Participant 7: I think it could be quite intrusive, and means that the person is almost like a robot
Future	Inevitable push towards this technology	Participant 37: I think there's no doubt that the technology is here to stay and how it can botost [boost] production[...] Participant 56: It clearly is, and has been for some time, the future direction of manufacturing in most major industries[...]
	Positive Outlook on Technology	Participant 50: I am sure it is a very good thing in terms of the advancement of technology. Participant 49: amazing how it can improve productivity
	Negative outlook on Technology	Participant 37: [...]but i wonder how much robots impact people's jobs, especially when so many more people have been made redundant. Participant 55: I think this technology will be used to exploit workers rather than

	improve their workplace. Why would technology make the workplace better if the people in charge have the same ethics codes and interests (i.e., profit) as before?

2.1.1 Personal Data

The theme of Personal Data was identified from the data. This theme initially appears broad, but it was chosen to represent an individual employee's personal data collected by the systems in the utopic vision. The inclusion of "personal" helps emphasise the type of data that is being collected about the user, such as age or heart rate. This is in comparison to the wider idea of data, which is prevalent in the topic of digital manufacturing. The sub-themes highlight the key concerns being Data Capture and Data Usage.

The participants presented an interest in the topic of data capture. Data in these instances would be captured in the workplace which is a cause of anguish for participants with one quote: 'What kind of dystopian ideal is this? Forcing workers to be biometrically monitored during their time at work?' (Participant 22). This relates to the Utopic video's suggestion that employees can have their data captured throughout the day, for example, with task performance (whether they are at work or not). This sub-theme also demonstrates a strong revulsion to data capture with participants using phases such as 'dystopian ideal' or that physiological data capture is 'a bit creepy' (Participant 62). This was located within the utopic vision, which suggested that even when presented with a positive view, the concept of constant data capture evokes negative thoughts in participants. The cause of this could be the Utopic video having dystopic elements or that just the thought of the topic produces negative thoughts for participants.

The participants disclosed their thoughts around the usage of the captured data. When users have full control over their personal data, they are able to keep it secure and can ensure appropriate usage. However, as one participant suggested, 'If the data is shared with an employer, i think its too much data shared with others.' (Participant 59), there is concern around when 'data' (Participant 59) leaves the user's control. An employer who is looking to implement a data system in this way should be wary of how it uses the data it captures, even avoiding this practise if data cannot be correctly handled.

Further, usage also can pertain to how elements of the data are being used; the quote, 'How are our own personal (out of work) factors taken into consideration in a way that does not prejudice the analysis.' (Participant 19), demonstrates that whoever ends up using the data needs to be careful to avoid incorrect assumptions from the data. A second insight from this quote is slightly nuanced, there is the assumption of the participant (19) that this data will be used on an individual basis. This also seems to be one of the points of contention with the data usage, that it could identify one person for a specific reason. Additionally, it could be suggested that this participant may feel differently if the data was aggregated over a large dataset.

2.1.2 People

The next theme identified was People, concerning mainly the idea of many individuals as a collective rather than specific individuals or the society on a large scale. This can be built from the sub-themes of possible discrimination and the degrading of human values.

The first branch of the theme is focused on the discrimination of potential workers. Participants picked up on the idea that certain workers may not be able to acquire work if their performance is being assessed. The sub-theme is similar to the Personal Data theme in the idea of data runs throughout. The discrimination aspect looks at the wider implications of data rather than the specific thoughts on how an individual's personal data is being utilised. The examples of quotes provided highlight this to a greater extent, with one participant noting how 'some people will never be able to get a job.' (Participant 48). From this, it appears the comment references how certain members of society will struggle to find meaningful employment if they do not match the industry's view of the ideal employee. With the future use of these technologies, it appears to be, at least in the view of the participants, a future which will cause discrimination against members who cannot match the data outputs of other members of the workforce.

The next sub-theme was that of the changing of how humans are perceived. Participants noted a disconnect between how people are perceived currently and how they could be perceived with the increased use of these digital technologies. One participant identified '[...]that the person is almost like a robot' (Participant 7). This example presents the "degrading of human values" by comparison against a machine, which are synonymous with constant repeated actions and showing little emotion. The participant could be further suggesting that in the utopic vision, humans would be expected to act in a more calculated methodological way. This is echoed in the other selected quote for this sub-theme, which suggested 'that we are becoming a society that will depend mainly on data' (Participant 57). The data referred too could also be linked to the degradation of human values, in the sense that rather than making personal decisions based on intuition, people could start making all decisions based on data. The data driven decision making linking with the calculated machine-like view of society in the previous quote.

2.1.3 Future

The final theme located within the data was that of the future. The "future" in the context of the subtheme will relate closely with the direction of travel and how participants feel. This differs from the previous themes, which looked more closely at specific aspects or elements of the utopic vision. There also appeared to be a divide between thoughts about the technology from participants; some appeared to be looking forward to the implementation of the technology while others were concerned with the potential pitfalls of the technology. Both were covered as separate sub-themes to give an impression of the arguments for and against digital manufacturing in the eyes of the participants.

The first sub-theme identified was around the nature of the current trend towards technology. This was often taken as a belief that technological progress is ever moving forward. One participant describes the idea that 'It [digital manufacturing] clearly is, and has been for some time, the future direction of manufacturing' (Participant 56). Their feeling of the direction of travel of the industry is unambiguous, and here lies one of the key aspects, they feel as if there is no alternative to this ever-present moving forwards of technological advancement. This further lends itself to the nature of the with the technologies featured being 'Industry 4.0' (Nahavandi, 2019); the idea of it being the 4th generation is an indication of an update to match current trends.

The next sub-theme was that of a positive outlook on the technologies presented by the video content. This sub-theme, along with the following sub-theme, both seem almost contradictory in the inclusion of both; there is an argument to be made that the themes themselves are not mutually exclusive. As found within the data there was scope for participants to be positive about the future use of the technology. One comment focused on the potential to affect tasks by stating, '[it is] amazing how it can improve productivity' (Participant 49). As a note, it is difficult to ascertain from participant's comment whether the technology's perceived benefits to "productivity" are in or out of work contexts. One could suggest that productivity is referring to work environments; the video's content showing a gym scene may allow the argument that the participant is referring to fitness goals or targets. Further, this widening of the scope allows for the theme to not limit itself to only industry contexts and suggests that the technology could be perceived to be beneficial to outside of work activities, which can be benefited by the technology.

The final sub-theme is related to how technology is perceived in a negative way as it moves into the future. One participant's response to the question was around the consequences of moving towards technological advancement with the comment, '[...] i wonder how much robots impact people's jobs, especially when so many more people have been made redundant.' (Participant 37). One idea from this to consider further is the idea of 'wonder', which presents a slight uncertainty of the future, with the comment around job losses adding a negative context to this uncertainty. While this specific comment mentions 'how much robots impact people's jobs', there is not a set topic to this theme of the outlook, as another quote suggest that 'this technology will be used to exploit workers rather than improve their workplace' (Participant 55). It is difficult to be sure what element of the technology will be used to what end, but there is this perceived sense that the technology will generally be used in a negative way, which is one of the key elements of this sub-theme.

2.2 Dystopic Condition

The Dystopic Condition was formed of two themes: Wider Society and Individual Workspaces. These refer to participant-identified ideas that discuss how these technologies will lead to changes in how they currently experience the day-to-day aspects of employment or their lives.

Table 34 – Dystopic Thematic Analysis in Tabular Format

Quotes are presented as formatted in Utopic table.

Themes	Sub-Themes	Quotes
Wider Society	Positive Outlooks for Industry	Participant 75: This sort of technology is certainly exciting and offers benefits to productivity and wealth creation[...] Participant 57: I believe digital technology brings improvement for businesses
	Degradation of Human Values	Participant 76: Physiological data collection turns the human into an ineffecient machine' Participant 74: [...]but humans should not be treated as robots, and should fill roles that cannot be done as well as by robots.
Individual Workspaces	Less Socialising	Participant 62: It is a very controlling environment which encourages isolation between workers thus not encouraging interaction and therefore not benefiting by group learning. Participant 66: it eliminates human interaction and places robots in place of the workers.
	Job Security	Participant 21: [...] but it is also important to introduce them in the right way to avoid mass redundancies without other jobs being available or a universal basic income like alternative Participant 61: Learning how robots and humans can work together so that humans don't loose their jobs to robots.

2.2.1 Wider Society

The first theme uncovered refers to a Wider Society. This was devised from participant responses which related to the broader aspect of society, rather than a more individualistic or small-scale basis and is represented by the two sub-themes chosen. The use of Industry and Human Values both refer to the larger aspect of society. With industry focusing on the collective total of businesses, workplace environments or organisations which could potentially utilise the technologies in the visions. The Human Values section of the other sub-theme encompasses the collective individuals of a society and how the technologies could impact their lives.

The first sub-theme revolved around a perceived Positive Outlook for Business. Participants were noted to have taken a more industry-focused perspective when discussing the effects of the technology moving into the future. There was also a positive spin; one participant commented, ‘This sort of technology is certainly exciting and offers benefits to productivity and wealth creation[...]’ (Participant 75). From the response, it is possible to reinforce the idea of the positives for businesses. ‘Productivity’ could be suggested to be referring to employees being more productive in the workplace. This can be suggested links to the ‘wealth creation’ with the increased productivity leading to great output, in turn, leading to greater revenues or profits. The next quote further illustrates the positive aspect for businesses with the statement: ‘I believe digital technology brings improvement for businesses’ (Participant 57). As a precise comment, this participant suggests that digital manufacturing technologies provide a benefit and thus, they have a positive view of the technology.

The final sub-theme for the section is related to how digital manufacturing technologies could affect and alter how society views people, with the title Degradation of Human Aspect. An example of this is found within the participants’ questionnaire responses with one quote of, ‘Physiological data collection turns the human into an ineffecient machine’ (Participant 76), showing ideas of this degradation. The notion of ‘the human’ in this case would be a worker chasing the ‘ideal’ performance metrics present within certain ‘physiological data’ types. It is possible to suggest that the participant is implying that the constant striving for these goals will make the work rigid and methodical, such as found within automation, the ‘ineffecient machine’. This is clearer with a part of another comment, ‘[...]but humans should not be treated as robots [...]’ (Participant 74), echoing the point of the previous participant. The comment continues, ‘[...], and [humans] should fill roles that cannot be done as well as by robots.’ (Participant 74). Presents the idea that people should not be made to act as an automation, as people are better suited to different job roles. This links to the wider society theme of how peoples' roles in society could change with the implementation of this sort of technology.

2.2.2 Individual Workspaces

Individual Workspaces is the second theme identified within the Dystopic condition. This theme contrasts the alternative theme in this condition by looking towards a small-scale industry setting, such as localised effects found within a business, instead of describing the societal implications of the technology. The sub-themes of Less Socialising demonstrate this theme, as this would affect the working groups within an individual business, defined here as a "workspace". The second sub-theme of Job Security is another effect which could be more localised to specific workspaces, as it could potentially vary across different businesses.

One of the sub-themes was that of the potential to have Less Socialising at workspaces. To best explain this theme, one participant's comment was that '[i]t [digital manufacturing technologies] is a very controlling environment which encourages isolation between workers thus not encouraging interaction and therefore not benefiting by group learning.' (Participant 62). From this, it is possible to suggest that the lack of socialising results in 'isolation'. The isolation, such as workers meeting during break hours, will affect their work. The participant presents the idea of 'group learning', which is suggested to be a positive effect on the work produced by the employees. This is further emphasised with a second participant's response that 'it eliminates human interaction and places robots in place of the workers.' (Participant 66). As with the previous comment, there is the running theme that there will be less socialising, which could affect workspaces in how employees act. There is a greater sense of perceived changes in team working dynamic as the technology 'places robots in place of the workers'. As implied by the participant (66), this change in the team working could also be suggested to have an effect on the ability for workers to socialise.

Another sub-theme that was identified was that of the effect on Job Security in an individual workspace. This was related to the responses by participants which focused arounds concerns about how secure their job would be, as a result of implementing digital manufacturing technologies. One response '[...] but it is also important to introduce them in the right way to avoid mass redundancies without other jobs being available or a universal basic income like alternative' (Participant 21) captures not only this concern, but further suggesting that there needs to be a safeguarding process for employees. The two present within this example are 'other jobs being available or a universal basic income like alternative'. The central sub-theme is also identified within another response to the survey which states '[l]earning how robots and humans can work together so that humans don't loose their jobs to robots.' (Participant 61). This comment further reinforces the idea job security is a potential issue for digital manufacturing technology. Additionally, there is the idea that human-robot collaboration could be used to avoid issues with job security.

Appendix B – Chapter Three Ethics Documentation

(Formatting changed where required)

Ethics Forms Additional Information

Contents

Number And Type Of Participants

General Adult Population, three workshops consisting of a maximum of 6 participants (Maximum 18 in total).

Number And Duration Of Activities Participants Will Be Involved In

Participants will be involved in a single day "event", consisting of two workshops, introduction and debrief sessions, and a refreshment break. The event is semi-structured meaning that while a schedule is provided there is freedom in how this is completed by participants (as in, they will be instructed to complete in a certain way, but they may make their own decisions as to which parts they complete or how they complete the activities).

Further, changes may be made to the in schedule/activities dependant on how previous workshops are completed. For example, workshop group 2's schedule/activities may change dependant on how group 1's workshop went. This is to ensure that any problems or potential for problems can be resolved and not repeated to improve factors like participant experience and/or research outcomes.

The schedule/activities are as follows:

Set up/ Introduction (10-15 Minutes)

1. Participants will receive printed handouts of the ethics forms (including an information sheet which acts as a schedule)
2. Participants will be asked to complete a demographic questionnaire (either printed or online version)
3. Participants will be asked if they have any questions about the schedule or other general questions
4. 'ice breakers' (perhaps a 'round-table introduction' like in Yao *et al.* (2019), but there may be many activities which can be used as ice breakers)
5. Explanation of the 'design items' (Muller, 1991) available for use
6. Explanation of the purpose of the workshops

Workshop 1 (30 minutes)

(Combination of Wetzel, Rodden and Benford, 2017 and Yao et al., 2019 methods)

- Participants will be given:
    - Sets of cards: Yellow are "Task" card, Green are "Automation" cards and blue are "Data" cards (These are based on Wetzel, Rodden and Benford (2017) cards)
    - Access to 'Design Objects' (Muller, 1991)

8. Shuffle all three sets of cards (and keep the sets separate)
9. The participant will be given/asked to draw a "Task" card and a "Automation" card

10. The participant will follow a modified Wetzel, Rodden and Benford (2017) ‘limited choice’ method, in which they will draw a “Data” card
11. Then ‘brainstorm’ (as suggested by Yao et al. (2019)) a way in which the automated system can use the data (in the “Data” card) in that system for that task (described in the “Automation” and “Task” cards)
    a. Participants will be asked to think about how the system is designed and their attitudes towards this system if they had to use it (Further, clarifications and questions may be asked by the researcher in an unstructured way)
12. They will then draw another “Data” card and repeating step 4, but with two “Data” cards to consider
13. They will then draw another “Data” card and repeating step 4, reaching the ‘three [...] cards’ described in Wetzel, Rodden and Benford (2017) and consider them all
14. The participant should have three data infused systems, A discussion should be had about their ideas where participants ‘tell’ (Brandt et al., 2012) their ideas to other participants and the researcher.
    a. Other participants and the researcher are allowed to ask questions about their ideas in an unstructured way

Break (10 minutes)

- Participants are allowed to take a break for approximately 10 minutes (if they wish for a longer or shorter break this should be possible)

Workshop 2 (30 minutes)

(Repeat of workshop 1 with scenario 2)

- Participants will be given:
    - Sets of cards: Yellow are “Task” card, Green are “Automation” cards and blue are “Data” cards (These are based on Wetzel, Rodden and Benford (2017) cards)
    - Access to ‘Design Objects’ (Muller, 1991)

1. Shuffle all three sets of cards (and keep the sets separate)
2. The participant will be given/asked to draw a “Task” card and a “Automation” card
3. The participant will follow a modified Wetzel, Rodden and Benford (2017) ‘limited choice’ method, in which they will draw a “Data” card
4. Then ‘brainstorm’ (as suggested by Yao et al. (2019)) a way in which the automated system can use the data (in the “Data” card) in that system for that task (described in the “Automation” and “Task” cards)
    a. Participants will be asked to think about how the system is designed and their attitudes towards this system if they had to

use it (Further, clarifications and questions may be asked by the researcher in an unstructured way)

5. They will then draw another "Data" card and repeating step 4, but with two "Data" cards to consider
6. They will then draw another "Data" card and repeating step 4, reaching the 'three [...] cards' described in Wetzel, Rodden and Benford (2017) and consider them all
7. The participant should have three data infused systems, A discussion should be had about their ideas where participants 'tell' (Brandt et al., 2012) their ideas to other participants and the researcher.
    a. Other participants and the researcher are allowed to ask questions about their ideas in an unstructured way

Finish (5 Minutes)

- Recap the day's events
- Participants will provide their email address to receive an Amazon Voucher at a later date.

Equipment And Procedures To Be Applied

Materials:

- Room(s) for workshop to take place
- Printed forms (Consent forms, information sheet, privacy notice, MS Forms Demographic Questionnaire)
- 'video equipment' / devices (Muller, 1991)
- Audio equipment / devices
- 'The "design objects" mentioned above fall into two categories. The first category is simple office materials. These include pens, high-lighters, papers, Post-It™ notes of various sizes, stickers and labels, and paper clips — all in a range of bright colors.' (Muller, 1991)
- Multiple copies (enough for each participant and spares, to show examples/replace damage and significant wear and tear) of the sets (the three described below) of personalisation based "Ideation Cards" (please see following section and accompanying card drafts for further information):
    - A set of yellow themed cards which are "Task" cards (cards in the form of Wetzel, Rodden and Benford's (2017) 'theme card')
    - A set of green themed cards that are "Automation" cards (Each card describing a type of Personalisation system)
    - A set of blue themed cards which are "Data" cards (cards in the form of 'opportunity cards' (Wetzel et al., 2017)) About the different personal data types available

- University of Nottingham Automated transcription services for audio data transcription (from audio and video data sources)

"Ideation cards"

The cards/game in use are loosely based on the work by Wetzel, Rodden and Benford's (2017) and an informal discussion with Prof Benford. They were also helped by discussions with my principal supervisor who has worked with another card set. The cards shown in the accompanying sheet are drafts and may change between the ethics forms submission and the running of the workshops. The cards may also change with use of the workshops (for example changing the cards for group 2 as a result of how the cards were used by group 1).

Some of the ideas on the cards themselves are based on previous studies within the DigiTOP project and in my PhD. The cards may be "dystopic" and are used to create thought into these topics.

The images on the cards may have been edited for use on the cards

The cards are physical cards and may be reprinted between workshop groups/sessions.

Information About How Participants Will Be Recruited

Note: As on ethics forms

Participants will be recruited through email adverts and word of mouth, participants may be colleagues at the University of Nottingham (work/research) and/or friends.

Whether Participants Will Be Paid (State How This Will Be Done)

Participants will receive £20 in amazon vouchers. This will be given after the study at a later date through the email.

Plans To Ensure Participant Confidentiality And Anonymity

Note: As on ethics forms

Participant personal information and data collected from the study will be anonymised where possible in the release of findings. With the nature of qualitative data capture it may not be possible to fully anonymise every piece data/information, but this will be kept confidential where possible.

Plans For Storage And Handling Of Data

Note: As on ethics forms

During the workshops voice data from the groups will be captured (including things like discussions/ explanations of the participant's workshop activities outputs and answers to questions from the researcher about participant's workshop activities outputs). Video data of the workshops which will have the participants in view will be captured. Any drawings or written material will also be captured and physical items kept. Use of the any materials during the workshops (use/placement/"playing" of "ideation cards") will be captured.

The researcher will take written (physically or digitally and stored in this way) notes about the workshops, what happened, participants participation and about the outcomes.

Necessary demographic data will be captured using MS forms either digitally or printed (Age, Self identity and education).

Digital data from the workshops will be stored on the researcher's computer after capture. The researcher's computer will be password protected (and/or pin code). Some of the data (MS forms demographic data, and audio/video recordings which are transcribed) may be uploaded and present on external platforms, such as: MS Forms, the University of Nottingham transcription platform, MS OneDrive, University of Nottingham, this data may or may not be stored on the researcher's computer.

The email addresses captured will be stored digitally (in a word document) only for allowing participants to receive their Amazon Voucher.

Due to the nature of a workshop task (one example could be a participant may be in the background of a video or audio source) it may not be considered reasonable to remove all traces of a participants involvement in a study. Although it will be removed where considered appropriate and/or reasonable.

Information About What Will Happen To The Data After The Study

After the reasonable use of the study data (analysis, reporting), it will be passed on to my "principal" supervisor (Dr Robert Houghton).

Information About How Any Data And Images May Be Used

Audio data (from audio recordings and video recordings) will be transcribed using an automated service (The University of Nottingham's transcription service). Transcriptions may be created/completed/edited/verified by in a manual way (for example, by hand by the researcher) if it is deemed reasonable.

Images/videos of the study will be viewed and may be described to better understand the content created by the participant/ how they participated in the study. These may appear in the reported results.

Both audio (transcriptions) and video data (images, video clips, descriptions) may be combined/ synthesised to better understand the how the participants participated in study. Both audio (transcriptions) and video data (images, video clips, descriptions) will be used in analysis by the researcher.

The researcher's written notes, will be used in analysis, synthesis and reported results.

Note: as above

The email addresses captured will be stored digitally (in a word document) only for allowing participants to receive their Amazon Voucher.

State Whether It Will Be Possible To Identify Any Individuals

Note: As on ethics forms and above

Participant personal information and data collected from the study will be anonymised where possible in the release of findings. With the nature of qualitative data capture it may not be possible to fully anonymise every piece data/information, but this will be kept confidential where possible.

Aims Of The Study

The aims of this study are to understand: How would end-users design personalisation within automated manufacturing environments and How accepting of personalisation are end-users. A further aim is similar to other works (Lucero & Arrasvuori, 2010) which aims to development a set of cards for participatory design.

References

Brandt, E., Binder, T. and Sanders, E.B.-N. (2012) 'Tools and techniques', in Routledge International Handbook of Participatory Design. Routledge. Available at: https://doi.org/10.4324/9780203108543.ch7.

Lucero, A. and Arrasvuori, J. (2010) 'PLEX Cards: a source of inspiration when designing for playfulness', in Proceedings of the 3rd International Conference on Fun and Games - Fun and Games '10. the 3rd International Conference, Leuven, Belgium: ACM Press, pp. 28–37. Available at: https://doi.org/10.1145/1823818.1823821.

Muller, M.J. (1991) 'PICTIVE---an exploration in participatory design', in Proceedings of the SIGCHI conference on Human factors in computing systems Reaching through technology - CHI '91. the SIGCHI conference, New Orleans, Louisiana, United States: ACM Press, pp. 225–231. Available at: https://doi.org/10.1145/108844.108896.

Wetzel, R., Rodden, T. and Benford, S. (2017) 'Developing Ideation Cards for Mixed Reality Game Design', Transactions of the Digital Games Research Association, 3(2). Available at: https://doi.org/10.26503/todigra.v3i2.73.

Yao, Y. et al. (2019) 'Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes', in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19: CHI Conference on Human Factors in Computing Systems, Glasgow Scotland Uk: ACM, pp. 1–12. Available at: https://doi.org/10.1145/3290605.3300428.

Participant Consent Form

Please read the following points carefully.

1. I understand that I am voluntary participating in this study and can withdraw anytime without penalty. ☐
2. I have read and understand the "Participant Information Sheet". ☐
3. I understand that to complete this study I need to participate in both workshop "tasks". ☐
4. I understand that on completion of the study I will be given a £20 Amazon Voucher. ☐
5. I understand that my participation in the study (audio and visual data capture, and researcher's notes about the workshop, outcomes and how I participated), any drawings, any written material and use of the "cards" will be captured and used. ☐
6. I understand that my demographic data will be captured and used. ☐
7. I understand that the use of my data (from points 4 and 5) will involve an automated and/or manual (by researcher) analysis and/or transcription. ☐
8. I understand that my data will be anonymised where possible, but it may be possible to identify me from the anonymised data. ☐

9. I understand that if I withdraw it may not be considered reasonable by the researcher to fully remove all of my data from use in analysis/results and my data may still be used as stated above. ☐
10. I have read and understand the "Privacy information for Research Participants". ☐

Participant Signature:______________________________ Date:_____________________

Researcher Contact Details:

- Joshua Duvnjak (Joshua.duvnjak@nottingham.ac.uk)
- Dr Robert Houghton (robert.houghton@nottingham.ac.uk)
- Engineering Ethics Email: ez-eng-ethics@nottingham.ac.uk

For Researcher Use:

Participant Number:_________________________________

Ethics Approval (from reviewer)

Ethics Committee Reviewer Decision

This form must be completed by each reviewer. Each application will be reviewed by two members of the ethics committee. Reviews may be completed electronically and sent to the Faculty ethics administrator from a University of Nottingham email address, or may be completed in paper form and delivered to the Faculty of Engineering Research Office.

Applicant full name Joshua Duvnjak

Reviewed by:

Name S11

Signature (paper based only) ……………………………………………………………………………………………

Date 18/10/2022

☐ Approval awarded - no changes required

■ Approval awarded - subject to required changes (see comments below)

☐ Approval pending - further information & resubmission required (see comments)

☐ Approval declined – reasons given below

Comments:

I approve with minor change stated below.

Is research's computer password protected? This needs to be stated.

Please note:

1. The approval only covers the participants and trials specified on the form and further approval must be requested for any repetition or extension to the investigation.
2. The approval covers the ethical requirements for the techniques and procedures described in the protocol but does not replace a safety or risk assessment.
3. Approval is not intended to convey any judgement on the quality of the research, experimental design or techniques.
4. Normally, all queries raised by reviewers should be addressed. In the case of conflicting or incomplete views, the ethics committee chair will review the comments and relay these to the applicant via email. All email correspondence related to the application must be copied to the Faculty research ethics administrator.

Any problems which arise during the course of the investigation must be reported to the Faculty Research Ethics Committee

Appendix C – Original Personalisation Design Cards

PERFORMANCE

Performance is the information collected about how well or poorly a task is completed.

Image Credit: Kolleen Gladden, 2017

HEART RATE

Heart rate is the hearts "Beats per Minute" of a person.

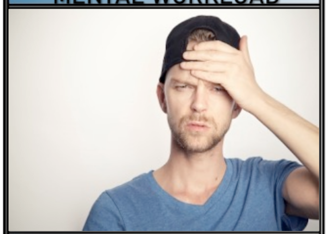
Image Credit: Jair Lázaro, 2017

HOBBIES

Hobbies are the interests that a person has.

Image Credit: Rifqi Ali Ridho, 2017

MENTAL WORKLOAD



Mental workload is how hard a persons mind is working.

Image Credit: Sander Sammy, 2021

fNIRS

fNIRS measures the blood flowing to parts of the head, often used to measure mental functions.

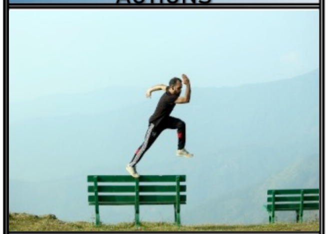
Image Credit: Josh Riemer, 2018

EYE TRACKING

Eye tracking captures where the eyes are currently looking.

Image Credit: Scott Webb, 2017

ACTIONS



Actions represents the capturing of a person's physical actions.

Image Credit: Niyas Khan, 2018

COMPLETION TIME

Completion time is the time taken by a person to complete a task.

Image Credit: Aron Visuals, 2017

TRUST IN AUTOMATION

Trust in Automation, is the trust a person would place in an automation in a situation.

Image Credit: Omar, 2021

TASK ERRORS

Task errors are the amount of errors a person makes while competing a task.

Image Credit: Kenny Eliason, 2017

TASK EXPERIENCE

Task experience relates to how much experience a person has on a task.

Image Credit: Mr Cup/ Fabian Barral. 2018

EMOTION TRACKING

Emotion tracking uses cameras to understand a person's emption through their facial features.

Image Credit: Tengyart, 2019

FATIGUE

Fatigue is how tired a person currently is.

Image Credit: Karollyne Hubert, 2022

DEMOGRAPHIC

Demographic relates to a person's background such as gender identity or age.

Image Credit: Timon Studler, 2016

DECISIONS

Decisions are the choices a person makes.

Image Credit: Jens Lelie, 2015

PHYSICAL CONSTRUCTION

Physical Construction is a task in which a person is building or making products.

Image Credit: Callum Hill, 2020

PRODUCT DESIGN

Product Design is a task in which a person is designing a product.

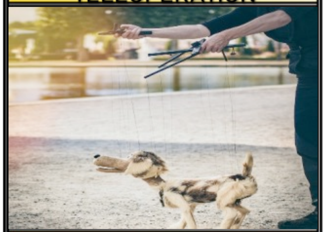
Image Credit: Kumpan Electric, 2020

QUALITY CONTROL

Quality Control is a task in which a person checks products for defects.

Image Credit: CDC, 2021

TELEOPERATION



Teleoperation is a task in which a person remotely controls another device, such as a robot.

Image Credit: Eric Masur, 2018

ROBOTIC SUPERVISION

Robotic Supervision is a task in which a person supervises a robot.

Image Credit: Simon Kadula, 2022

CONTROLLERS

"Controllers" are automated systems that take control away from a person based on personal data.

Image Credit: JESHOOTS.COM, 2017

GENERATORS

"Generators" are automated systems that create new things for a person based on personal data.

Image Credit: Sungrow EMEA, 2020

SUGGESTERS

"Suggesters" are automated systems that make suggestions to a person based on personal data.

Image Credit: Desola Lanre-Ologun, 2018

SWAPPERS

"Swappers" are automated systems that replace one type of content for another based on personal data.

Image Credit: HiveBoxx, 2020

REARRANGERS

"Rearrangers" change the order of content for a person based on personal data.

Image Credit: Maarten van den Heuvel, 2016

Appendix D – Revised Personalisation Design Cards and Game Mat

PERFORMANCE

Performance is the information collected about how well or poorly a task is completed.

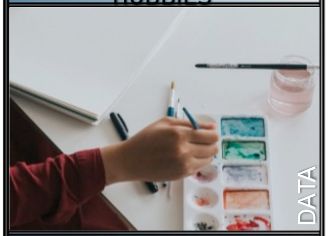
Image Credit: Kolleen Gladden, 2017

HEART RATE

Heart rate is the hearts "Beats per Minute" of a person.

Image Credit: Jair Lázaro, 2017

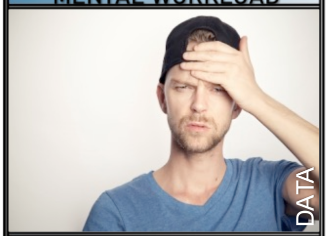
HOBBIES



Hobbies are the interests that a person has.

Image Credit: Rifqi Ali Ridho, 2017

MENTAL WORKLOAD



Mental workload is how hard a persons mind is working, often captured through asking a person.

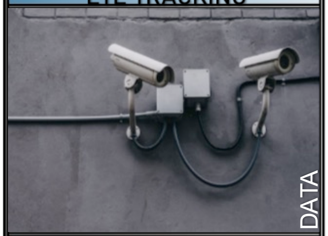
Image Credit: Sander Sammy, 2021

fNIRS

fNIRS headbands measure the blood flowing to parts of the head, often used to measure mental functions.

Image Credit: Josh Riemer, 2018

EYE TRACKING



Eye tracking captures where the eyes are currently looking.

Image Credit: Scott Webb, 2017

PHYSICAL CHARACTERISTICS

Physical Characteristics represents a person's physical properties, such as height.

Image Credit: Niyas Khan, 2018

COMPLETION TIME

Completion time is the time taken by a person to complete a task.

Image Credit: Aron Visuals, 2017

TRUST IN AUTOMATION

Trust in Automation, is the trust a person would place in an automation in a situation.

Image Credit: Omar, 2021

TASK ERRORS

Task errors are the amount of errors a person makes while competing a task.

Image Credit: Kenny Eliason, 2017

TASK EXPERIENCE

Task experience relates to how much experience a person has on a task.

Image Credit: Mr Cup/ Fabian Barral. 2018

EMOTION TRACKING

Emotion tracking uses cameras to understand a person's emption through their facial features.

Image Credit: Tengyart, 2019

FATIGUE

Fatigue is how tired a person currently is.

Image Credit: Karollyne Hubert, 2022

DEMOGRAPHIC

Demographic relates to a person's background such as gender identity or age.

Image Credit: Timon Studler, 2016

DECISIONS

Decisions are the choices and actions a person makes.

Image Credit: Jens Lelie, 2015

PHYSICAL CONSTRUCTION

Physical Construction is a task in which a person is building or making products.

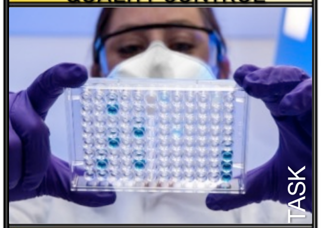
Image Credit: Callum Hill, 2020

PRODUCT DESIGN

Product Design is a task in which a person is designing a product.

Image Credit: Kumpan Electric, 2020

QUALITY CONTROL



Quality Control is a task in which a person checks products for defects.

Image Credit: CDC, 2021

ROBOTIC TELEOPERATION

Teleoperation is a task in which a person remotely controls another device, such as a robot.

Image Credit: Eric Masur, 2018

ROBOTIC SUPERVISION

Robotic Supervision is a task in which a person supervises a robot.

Image Credit: Simon Kadula, 2022

CONTROLLERS

"Controllers" are automated systems that take control away from a person based on personal data.

Image Credit: Taiki Ishikawa, 2020

GENERATORS

"Generators" are automated systems that create new things for a person based on personal data.

Image Credit: Sungrow EMEA, 2020

SUGGESTERS

"Suggesters" are automated systems that make suggestions to a person based on personal data.

Image Credit: Desola Lanre-Ologun, 2018

SWAPPERS

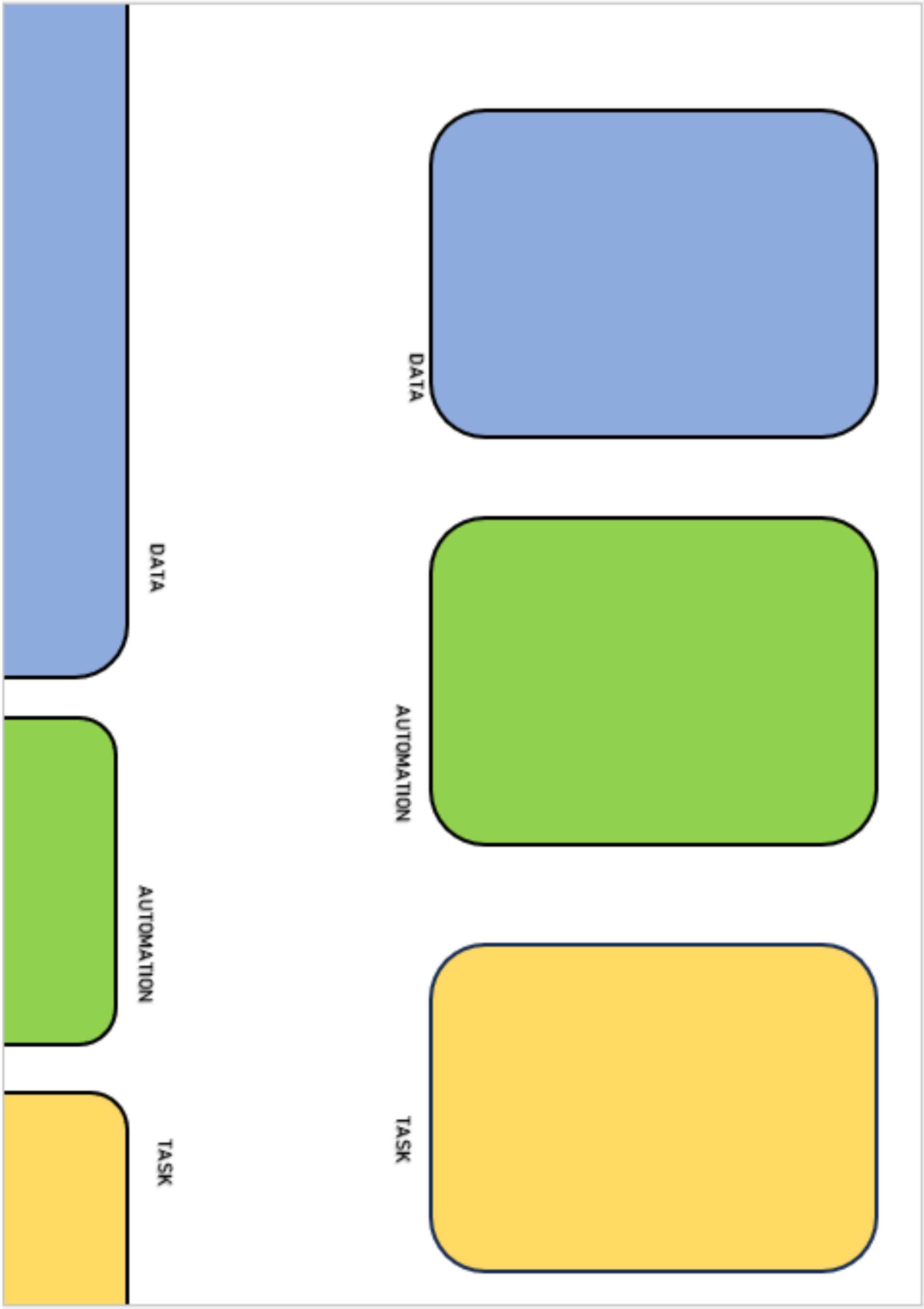
"Swappers" are automated systems that replace one type of content for another based on personal data.

Image Credit: HiveBoxx, 2020

REARRANGERS

"Rearrangers" change the order of content for a person based on personal data.

Image Credit: Maarten van den Heuvel, 2016


DATA
AUTOMATION
TASK
DATA
AUTOMATION
TASK

Appendix E – Workshop Demographic Questions

These questions were created in MS Forms and printed out for the in-person workshops. A typographic error was present on the last question, but there is a high chance this had no impact on the responses from the participants.

1. What is your age range?
   - 18-24
   - 25-34
   - 35-44
   - 45-54
   - 55-64
   - 65+
   - Prefer not to say

2. How do you identify?
   - Male
   - Female
   - Non-binary
   - Other
   - Prefer not to say

3. What is your highest education qualification
   - GCSE (Or equivalent)
   - A Level (Or equivalent)
   - Foundation Degree (Or equivalent)
   - Bachelor's Degree (Or equivalent)
   - Master's Degree (Or equivalent)
   - PhD (Or equivalent)
   - Other
   - Perfer not to say

Appendix F – Ethics Approval for Survey

Ethics Committee Reviewer Decision

This form must be completed by each reviewer. Each application will be reviewed by two members of the ethics committee. Reviews may be completed electronically and sent to the Faculty ethics administrator from a University of Nottingham email address, or may be completed in paper form and delivered to the Faculty of Engineering Research Office.

Applicant full name Joshua Duvnjak

Reviewed by:

Name C08

Signature (paper based only) ……………………………………………………………………………………………

Date 10/01/2023

☐ Approval awarded - no changes required

■ Approval awarded - subject to required changes (see comments below)

☐ Approval pending - further information & resubmission required (see comments)

☐ Approval declined – reasons given below

Comments:

I'm happy to approve this study. Only minor comment to student would be to store data on their Microsoft OneDrive Uni account, rather than on their laptop hard drive.

Please note:

5. The approval only covers the participants and trials specified on the form and further approval must be requested for any repetition or extension to the investigation.
6. The approval covers the ethical requirements for the techniques and procedures described in the protocol but does not replace a safety or risk assessment.
7. Approval is not intended to convey any judgement on the quality of the research, experimental design or techniques.
8. Normally, all queries raised by reviewers should be addressed. In the case of conflicting or incomplete views, the ethics committee chair will review the comments and relay these to the applicant via email. All email correspondence related to the application must be copied to the Faculty research ethics administrator.

Any problems which arise during the course of the investigation must be reported to the Faculty Research Ethics Committee

Appendix G – Survey Demographic Questions

The acceptance survey included a set of questions which related to acceptance. The numbers represent the order present on the Microsoft form, these were different between the first and second experiments (second experiments numbering included). The asterixis indicated a required completion by participants.

14. Please Insert you Prolific ID *

15. What is your age range? *

- 18- 24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+
- Prefer Not To Say

16. How do you self identify? *

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

17. Highest educational attainment? *

- GCSE or lower (or equivalent)
- A-Level (or equivalent)
- Foundation Degree (or equivalent)
- Undergraduate Degree (or equivalent)
- Master's Degree (or equivalent)
- PhD (or equivalent)
- Higher than PhD level
- Prefer not to say

18. Have you worked (current or previous) in the manufacturing sector or have experience of the manufacturing sector? *

- Yes
- No

Appendix H – Assistive Automation Experiment One Demographic Questions

The demographic questionnaire was held within the PsychoPy experiment (Peirce et al., 2019). Participants were able to select options, with each ',' separating the choices.

1. How do you identify?

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

2. What is your age range?

    18-24, 25-34, 35-44, 45-54, 55-64, 65+, Prefer not to say

3. Please describe your job title?

    job title or leave blank

4. How many hours a day would you say you use a smartphone device?

    0-1, 2-3, 4-5, 5-6, 6+, Prefer not to say

5. How many hours a day would you say you use a personal computer?

    0-1, 2-3, 4-5, 5-6, 6+, Prefer not to say

6. How many hours a day would you say you use automation for work (this could be anything from a word processor or excel sheet to assisted driving)?

    0-1, 2-3, 4-5, 5-6, 6+, Prefer not to say

7. How many hours a day would you say you use automation for non-work (this could be anything from a word processor or excel sheet to assisted driving)?

    0-1, 2-3, 4-5, 5-6, 6+, Prefer not to say

Appendix I – Trust Scale

A trust scale by Jian et al. (2000) was included as part of the Quality Control Experiments. These were on a seven-point Likert scale. This was included in PsychoPy (Peirce et al., 2019) within the first experiment and within MS Forms in the second (the numbering was removed and a single number was present for the total scale).

1. The system is deceptive
2. The system behaves in an underhanded manner

3. I am suspicious of the system's intent action, or outputs
4. I am wary of the system
5. The system's actions will have a harmful or injurious outcome
6. I am confident in the system
7. The system provides security
8. The system has integrity
9. The system is dependable
10. The system is reliable
11. I can trust the system
12. I am familiar with the system

Appendix J - Assistive Automation Experiment Two Demographic Questions

The second assistive automation experiment included a demographic questionnaire that varied from the first experiment.

3. Prolific ID *
4. What is your age range? *
    - 18-24
    - 25-34
    - 35-44
    - 45-54
    - 55-64
    - 65+
    - Prefer not to say
5. Highest Educational Attainment *
    - GCSE or lower (or equivalent)
    - A-Level (or equivalent)
    - Foundation Degree (or equivalent)
    - Undergraduate Degree (or equivalent)
    - Masters Degree (or equivalent)
    - PhD or higher (or equivalent)
    - Prefer not to say
6. How do you self indentify? *

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