PHD

Investigating Spatial Augmented Reality for Collaborative Design

Giunta, Lorenzo

Award date:
2022

Awarding institution:
University of Bath

Link to publication

Investigating Spatial Augmented Reality for Collaborative Design

Lorenzo Giunta

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Mechanical Engineering

January 2022

– Acta est fabula plaudite –

ACKNOWLEDGEMENTS

I want to thank my supervisors: Elies Dekoninck and James Gopsill whose patience, knowledge, dedication, and kindness has helped me through the toughest parts of this doctorate. It has been an absolute pleasure working with them and the insights they provided have been invaluable. I would never have been able to achieve so much without their help.

A thank you to all the members of the SPARK Consortium: it was an honour to pursue this project with you making your acquaintance has left me a better person for it.

I also want to thank my family for their continued support during this PhD journey, your support has been invaluable to the success of this work. For those that have been with me throughout and for those who sadly did not see me complete this path. This thesis is dedicated to you, from the bottom of my heart: Thank you for helping me achieve all I have.

Lastly, a thank you to all my friends, colleagues, and other staff at the University of Bath. A PhD may be a road one must travel alone, but knowing you were all there following your own paths always gave me heart.

In memoriam: Vincenzo Giunta, Domenica Ragusa, Vito Ardita, Trofimena Di Lieto

ABSTRACT

This thesis presents an analysis of the implementation of Spatial Augmented Reality (SAR) to support collaborative design. SAR is a class of Augmented Reality (AR), where the AR effect is independent from the viewer's perspective. This property should make SAR uniquely well suited to support collaborative design sessions as it allows multiple participants to simultaneously view the same AR enhanced models concurrently. However, the adoption of SAR in the field of design remains limited. Thus, there is a knowledge gap that surrounds SAR, due to the novelty of the technology, and how its implementation can be configured to better support design. The aim of this research is therefore to:

"Investigate SAR Systems' Characteristics and Features and how these Affect the Efficacy of Co-Design Sessions ".

A review of the literature surrounding collaborative design is provided to highlight evidence that enhanced collaboration during design sessions results in improved outcomes. Thereafter AR, and SAR in particular, are analysed to provide a basis for the continued investigation of SAR as a valid tool to support co-design sessions. In addition to the review, the state-of-the-art in implementing SAR in design and collaborative design is presented and discussed to frame the subsequent research conducted as part of this PhD. Moreover, literature is provided covering the metrics that will be used to analyse SAR supported co-design sessions.

The methodology implemented throughout this thesis was one of empirical research conducted through a number of experimental studies aimed at evaluating the impact of SAR on co-design sessions. The experimental nature of this research thus required the development of an SAR platform to be used during experimental testing. The platform was developed iteratively and improved on the basis of feedback received from experiments. These experiments focused on three main areas:

1. The impact of SAR on co-design sessions as a whole
    - These experiments aid in understanding the impact that SAR has on the overall design session, focusing primarily on the design outcomes
2. The analysis of specific characteristics and features of SAR and their impact on design outcomes
    - Experiments used to analyse the impact of specific elements that are innate to SAR to evaluate their individual impact on SAR supported design sessions
3. Experiments and data collection aimed at gathering industry feedback
    - Used to understand the industry perspective to allow for the optimization of SAR to cater to the needs of industry and ultimately allow for its wider adoption

The results highlighted the viability of SAR as a tool to support collaborative design sessions. However, the studies also identified a number of areas for improvement in the implementation of SAR. The thesis concludes with a summary of thirty-two findings, condensing them into two lists of recommendations. One of these lists for technical improvements and one for features and affordances, aimed at assisting future researchers wishing to develop improved SAR platforms for co-design.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Full Form	Abbreviation	Additional Information
Antwerp Management School	AMS	University of Antwerp's autonomous business school
Artefice Group		Milan based design consultancy with a focus on packaging design
Augmented Reality	AR	
Augmented Virtuality	AV	
Characteristics of SAR		A core, defining element of SAR
Client		Individual(s) who have commissioned the design process
Collaborative Creation	Co-Creation	
Collaborative Design	Co-Design	
Collaborative Design Review	CDR	
Creativity Support Index	CSI	A psychometric survey designed for evaluating the ability of a creativity support tool to assist a user engaged in creative work (Cherry and Latulipe, 2014)
Design Research Methodology	DRM	
Digital Prototype		A prototype that exists only in digital space and is therefore not physically tangible
End User		Final consumer of the product being designed
Finite Element Analysis	FEA	
Features of SAR		Elements whose presence is not necessary for a platform to be considered SAR
Grenoble INP	GINP	Grenoble technological university
Graphical User Interface	GUI	
Hand-Held Display	HHD	
Head Mounted Display	HMD	
Keystone Effect		Image warping that gives projected images a trapezoidal appearance
Mixed Reality	MR	The combination of real and virtual environments
Physical Prototype		A prototype that exists in real space and is physically tangible
Politecnico di Milano	POLIMI	Technical university of Milan
Real Environment		The physical environment
Reality Virtuality Continuum		Spectrum ranging from fully real environments to fully virtual environments
SPARK Consortium	SPARK C.	Collaborating members of the SPARK project
SPARK Platform		The platform designed to achieve the goals of the SPARK project
SPARK Project		EU Horizon 2020 funded project to develop an SAR tool to support collaborative design
Spatial Augmented Reality	SAR	
Spatial Augmented Reality for Co-creativity	SPARK	
Stimulo		Barcelona based design consultancy with a focus on product design

System Usability Scale	SUS	A usability scale that can be used for global assessments of systems usability (Brooke, 1996)
Throw Ratio		Ratio between distance of projector lens to projection screen and screen width
University of Bath	UBATH	University of Bath
User Interface	UI	
UV Map		A map used to overlay a 2D drawing onto a 3D object
Virtual Reality	VR	Virtual environment
Viseo		Global IT a digital consulting firm

1 INTRODUCTION

Chapter One introduces the work that will be conducted as part of this PhD. Firstly, the motivation behind the PhD is laid out. Secondly, the background on the SPARK project, which was one of the funders, is discussed. Thirdly, the chapter provides a summary of Collaborative Design (co-design) as well as the role that Augmented Reality (AR) has played in design thus far, is presented. Chapter One then concludes by outlining the aim of the research and subsequent thesis structure.

1.1 DRIVERS FOR THE RESEARCH

Spatial Augmented Reality (SAR) is a form of Augmented Reality that achieves the mixing of real and digital environments independently from the users' perspectives. For example, when SAR is used to augment an object, the SAR effect is visible independently from the user position relative to said object. This enables SAR to function without the user needing to wear a headset, hold a tablet, or require any other individual means of viewing the AR effect. Further, SAR enables multiple users to view the same augmented scene in real time, as they can all view the same object, rather than having individual AR representations that must be synchronised together. In recent years an increasing amount of interest has been drawn to AR technologies, and SAR in particular, as a tool to support design (Park *et al.*, 2015; O'Hare *et al.*, 2018b). The viability of AR technologies, and of SAR in particular, as a tool to support design are further discussed in sections 2.4 and 2.5 respectively.

Of particular interest is the application of SAR in co-design as where it provides an intermediary object that can improve and support communication between participants (Boujut and Blanco, 2003; O'Hare *et al.*, 2018b). Collaborative Design (Co-Design) refers to a design activity where the inputs for the development of a design are not taken from a single individual but rather from a group. Often the use of a group as a source of input will be taken as an opportunity to merge expertise from multiple disciplines and/or backgrounds. Furthermore, participants in a co-design session need not all be designers. Some may be end-users or clients who had commissioned the work. Combined, they form the stakeholders in a design project. Co-design sessions pertaining to product and packaging design could particularly benefit from an SAR tool as it would facilitate the sharing and communication of concepts and ideas that are otherwise hard to express verbally. The etymology and implications of co-design are further explained in section 2.1.

The impact of SAR on co-design has attracted some interest from the scientific community. In the field of architecture, for example, the impact of SAR on co-design sessions has already solicited some interest (Ben Rajeb and Leclercq, 2013; Calixte and Leclercq, 2017). In their paper, Ben Rajeb and Leclercq (2013) note how the use of SAR enables a more immediate link between the words spoken by collaborators and the changes that are being made to a design. Calixte and Leclercq (2017) highlight the utility of SAR as a tool to support collaborative design. Calixte and Leclercq (2017) also highlight some challenges faced by those who wish to implement SAR systems. In particular, they note the difficulties that surround the calibration and the potential for cognitive overload. However, the suitability of SAR systems in supporting co-design has not been investigated in depth, especially in the context of product and packaging design.

As of yet, the impact of the inherent features and characteristics of SAR on co-design in general, and in product and packaging in particular, is not well understood. Understanding how the inherent SAR characteristics can impact a co-design session is desirable to allow for a better understanding of the value that SAR can have. This knowledge would aid in understanding the role SAR could play in future co-design sessions as a tool to support designers, clients, and end-users. Furthermore, the features of SAR have not been fully explored or catalogued to understand their value during co-design sessions. In order to support the development of future SAR systems for use in design these knowledge gaps must be bridged; the research presented within this thesis aids in the pursuit of this goal.

1.2 TIMELINESS OF THE RESEARCH

AR technologies have also attracted the interest of businesses and the commercial sector. This interest has spurred some market analysis into the long-term feasibility and commercial viability of the technology by Gartner, a market consultancy. They have developed a tool, known as the "hype-cycle", to aid investors with understanding the maturity and market readiness of emerging technologies. This tool and their specific mapping of AR technology help to show the timeliness of the research presented in this thesis.

At its core, the hype-cycle is a graph that shows the development of various technologies by mapping the public's expectations against time. An illustrative example of this is shown in Figure 1. The graph is subdivided into five sections that describe the public's expectations and behaviour towards the technology (Fenn and Blosch, 2018).

Figure 1 explains the progression of technologies along the hype-cycle. The first phase, labelled the "Innovation Trigger", is the period where the technology first emerges and goes from a completely unknown or niche technology to gaining wider-spread appeal and interest. Once early adopters begin to embrace the technology it enters the "Peak of Inflated Expectations" phase. Here, the technology begins to be embraced beyond the first early adopters and expectations continue to grow. As more adopters promote the technology the risk of overselling its capabilities rises causing a bubble. Once the technology fails to meet all the expectations, no matter how overinflated or unrealistic they might have become, negative backlash begins to lower the expectations of the general public. The technology then enters the "Trough of Disillusionment" phase where the public's expectations have considerably lowered towards the technology, but development is still ongoing. From here the technology rebuilds confidence in it as new methodologies are adopted, past mistakes are learned from, and support is built up in a more gradual and sustainable manner the technology enters the "Slope of Enlightenment" phase. Ultimately the technology continues to develop and gain support until it reaches the "Plateau of Productivity". Here the technology is considered to have matured beyond the point of being considered emerging. It should be noted that not all technologies fully develop throughout the entire cycle and different technologies will progress at different rates.

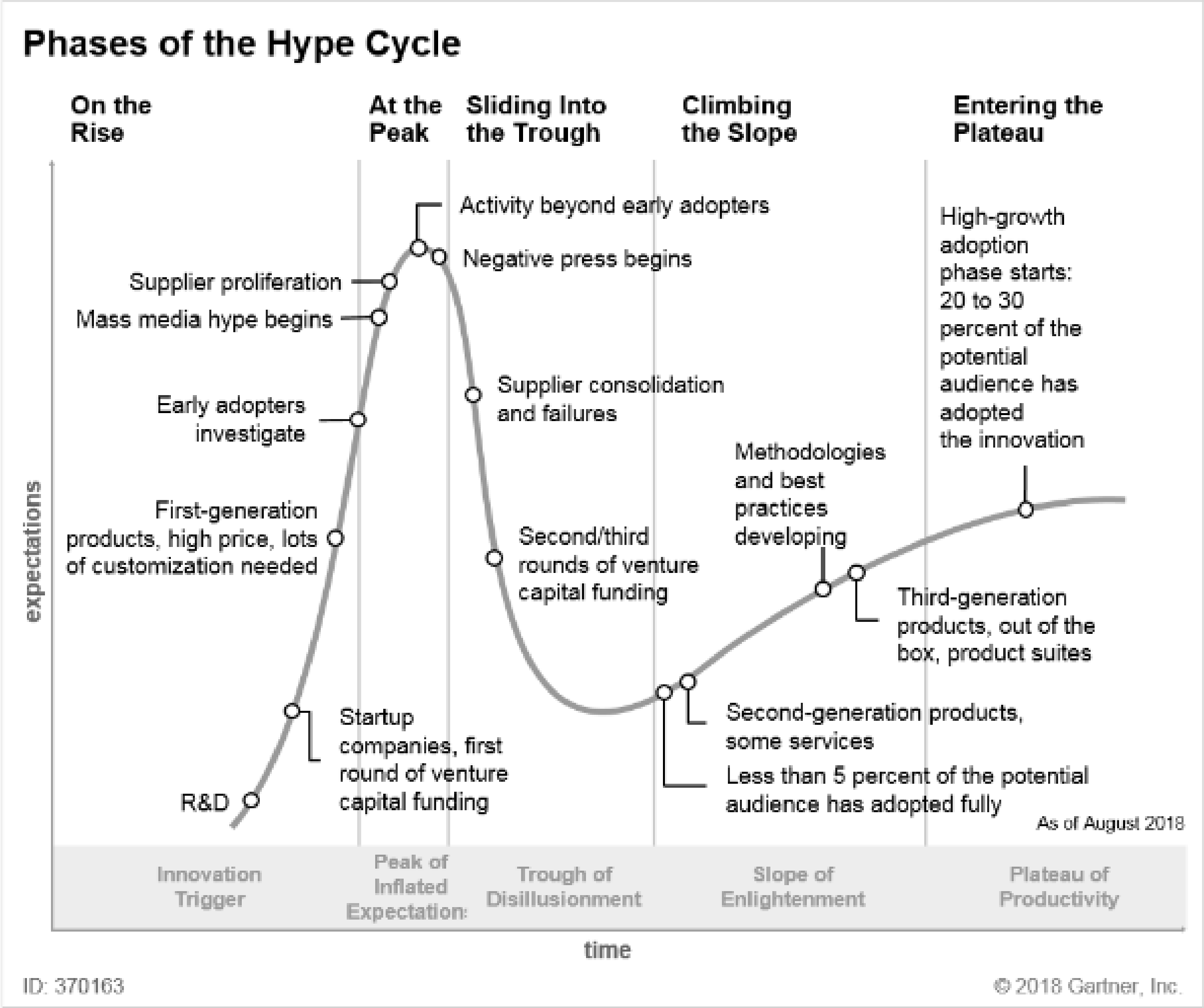


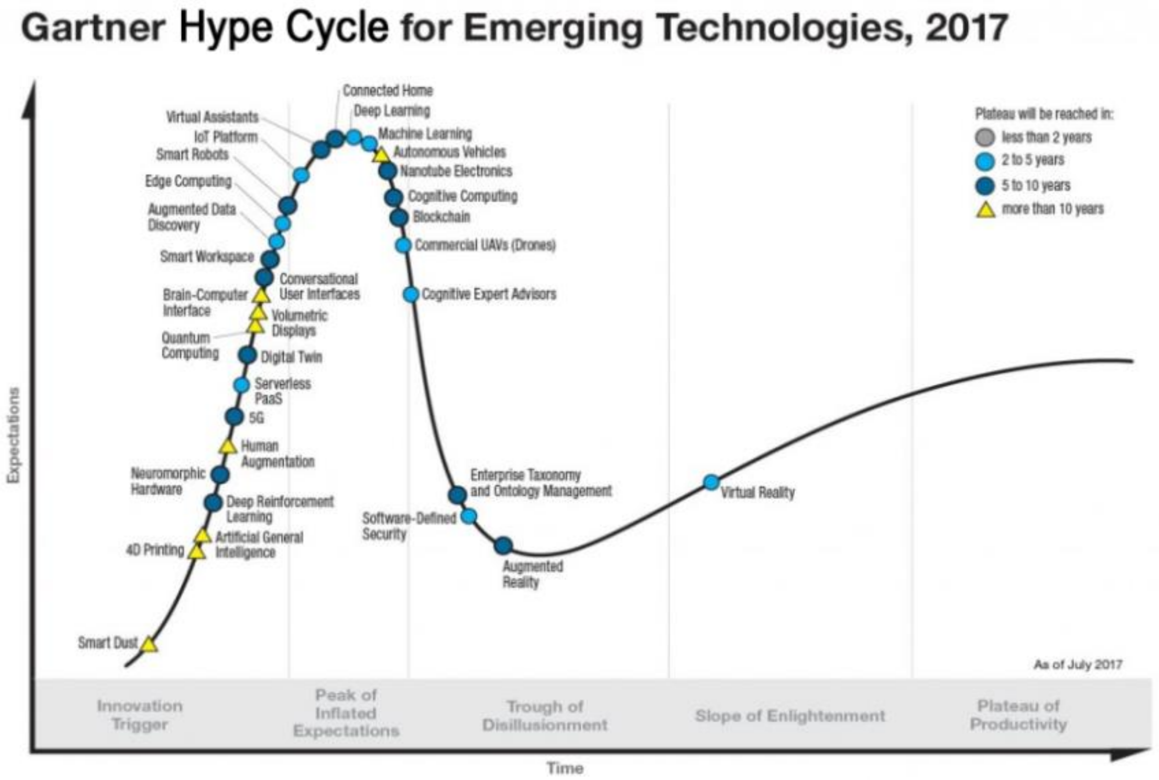
Figure 1. Progression of a Technology along Gartner's Hype-Cycle (Fenn and Blosch, 2018)

Figure 2 shows the placement of AR technologies along the Gartner hype-cycle over the last 5 years. In 2017, Gartner evaluated AR technologies as being in the "Trough of Disillusionment" (Panetta, 2017) and expected that the technology would require between 5 and 10 years to reach the plateau of productivity. It was unsurprising that, only one year later, in 2018, AR technologies were placed in the same position. However, one noteworthy addition in 2018 was "Mixed Reality" (MR). The technology, which had not been tracked the previous year, appeared directly in the Trough of Disillusionment, albeit at an earlier stage when compared to AR. This is a somewhat confusing choice as, in academia, Mixed Reality acts as something of a catch all term for all the technologies that lie between the real environment and the virtual environment on the Reality-Virtuality Continuum (Milgram *et al.*, 1995). Additional details to understand what the relationship between MR and AR are provided in section 2.3.

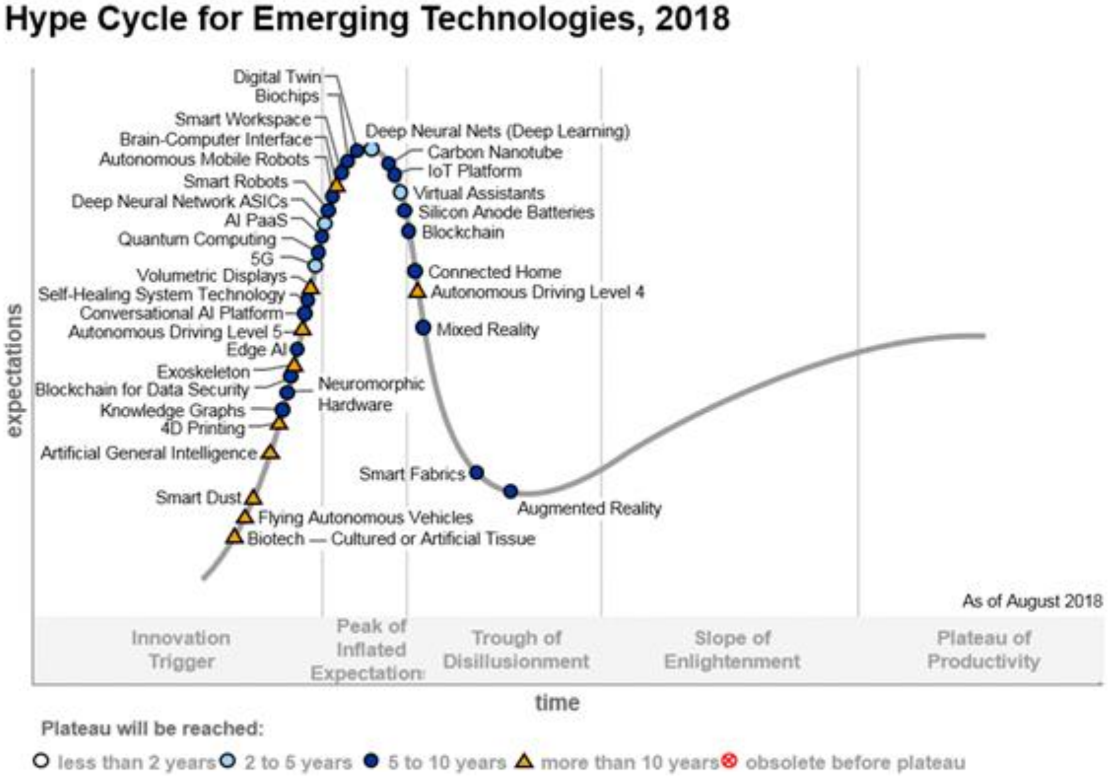
The 2019 edition of the hype-cycle manages to simplify matters. Both AR and MR are absent from the list of emerging technologies. The Gartner report for 2019 on emerging technologies (Burke and Smith, 2019) discusses why both AR and MR were removed from the hype-cycle. In the case of MR, the rationale given is that, while MR remains an emerging technology, it was removed to highlight other trends. In contrast, AR was considered to be rapidly approaching maturity, and as such no longer considered an emerging technology to be included within the report.

To explain this discrepancy it becomes relevant to note that Gartner has made no distinction between the various sub-categories of AR (Panetta, 2017). Since Hand-Held Display (HHD) AR systems (such as Pokémon Go or Minecraft Earth) were amongst the first types of AR to gain mainstream acceptance, it is likely that the report predominately focuses on these over SAR or Head Mounted Displays (HMD), which may have been be categorised as MR. Supporting this evaluation is the work conducted by Rauschnabel, Brem, and Ivens (2015) who undertook two preliminary studies, finding that there is potential for acceptance of AR HMD technology, by analysing the correlation between personality traits of potential adopters and likelihood of technological adoption. Their findings, combined with Burke and Smith's (2019) opinion that AR is reaching maturity indicate that AR is gaining wider acceptance and adoption. This is, however, not indicative that all AR technologies are approaching maturation, but that AR adoption is occurring and that there is potential for novel or less explored forms of AR to gain attention and acceptance.

Additionally, while AR might have been omitted from the 2019 and later editions of the hype-cycle a new category has emerged. Augmented Intelligence, in the Innovation Trigger phase, seems to have become a more generic term for more advanced forms of AR and Virtual Reality (VR) that incorporate greater amounts of artificial intelligence as well as increased responsiveness (Columbus, 2019). From this, we can infer that the foundations of AR have become ubiquitous enough to no longer be considered emerging technologies with newer forms of AR emerging.

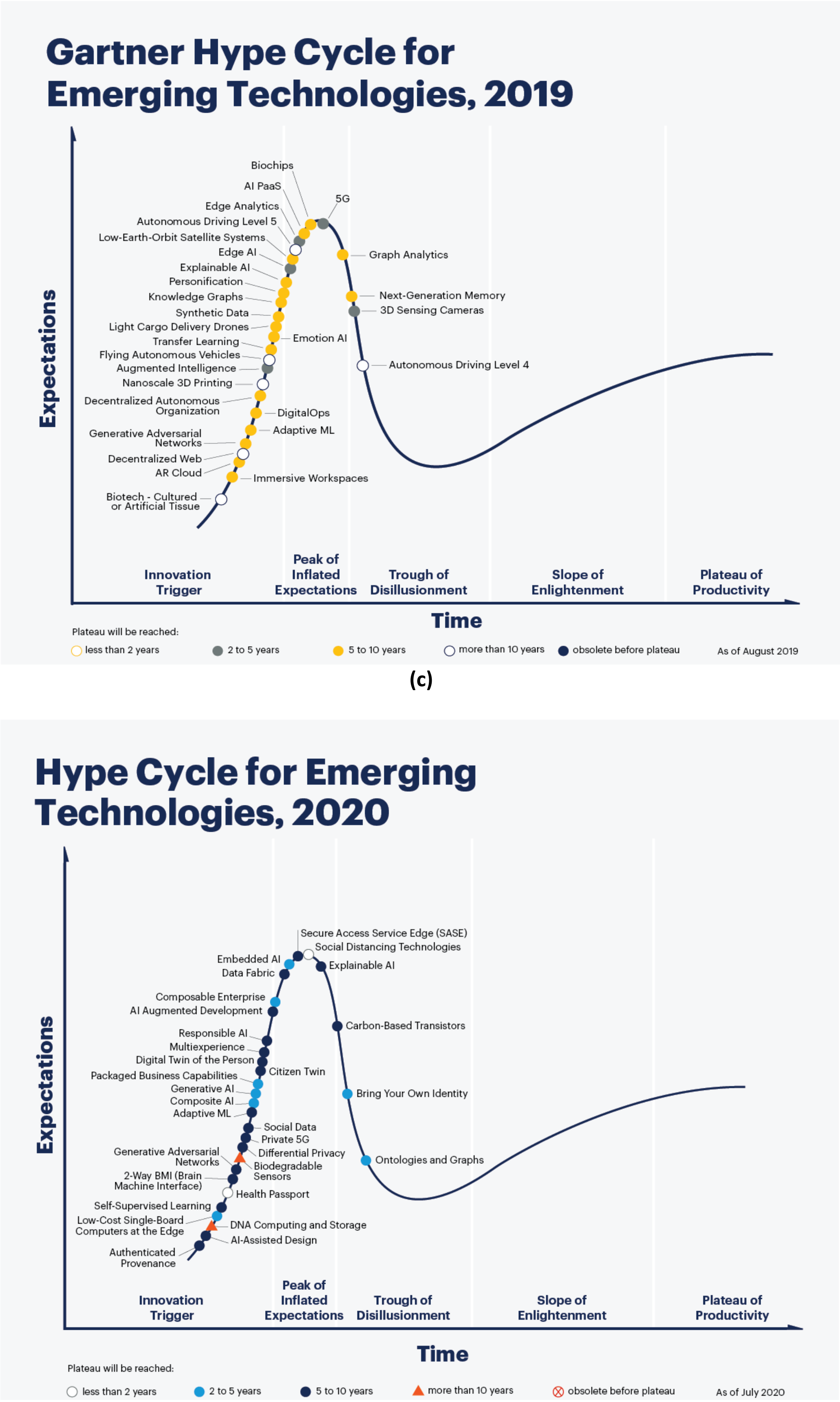


(a)



(b)

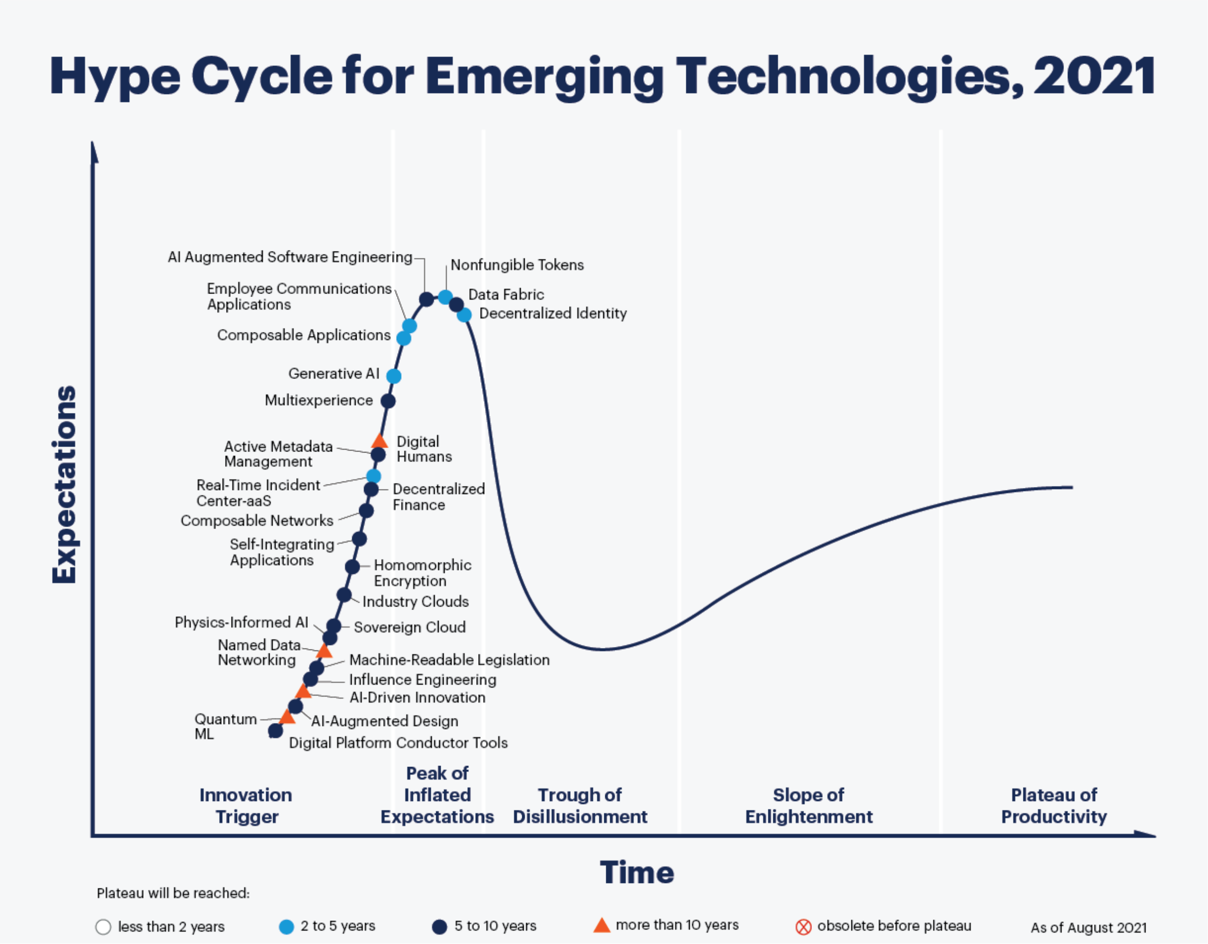
Figure 2. Gartner's Hype-Cycle analysis for Emerging Technologies in (a) 2017 (Panetta, 2017), (b) 2018 (Fenn and Blosch, 2018), (c) 2019 (Panetta, 2019), (d) 2020 (Panetta, 2021b), and (e) (Panetta, 2021a)



(c)

(d)

Figure 2. (cont.) Gartner's Hype-Cycle analysis for Emerging Technologies in (a) 2017 (Panetta, 2017), (b) 2018 (Fenn and Blosch, 2018), (c) 2019 (Panetta, 2019), (d) 2020 (Panetta, 2021b), and (e) (Panetta, 2021a)



(e)

Figure 2. (cont.) Gartner's Hype-Cycle analysis for Emerging Technologies in (a) 2017 (Panetta, 2017), (b) 2018 (Fenn and Blosch, 2018), (c) 2019 (Panetta, 2019), (d) 2020 (Panetta, 2021b), and (e) (Panetta, 2021a)

1.3 SPARK PROJECT AND PLATFORM EXECUTIVE SUMMARY

1.3.1 SPARK Project Background

Spatial Augmented Reality for Co-Creativity (SPARK) was a Horizon 2020 (a European Union funding programme) project which ran between January 2016 and the 1st of January 2019. The goal of the SPARK Consortium, who contributed to the SPARK project was to develop a: **"Responsive ICT platform that exploits the potential of Spatial Augmented Reality for supporting and fostering collaborative creative thinking"**(SPARK Consortium, 2015). The project co-funded this thesis thereby providing the opportunity to investigate the research gaps identified in this thesis as well as providing a range of SAR systems with which to experiment, in addition to connections with industry to support experiments.

The SPARK project was administrated by the SPARK Consortium which is composed of: The University of Bath (UBATH), Politecnico di Milano (POLIMI), Grenoble INP (GINP), Antwerp Management School (AMS), Stimulo, Artefice Group, and Viseo. The PhD started in 2017 and the SPARK consortium had already made considerable progress having already developed an initial SAR platform prototype as well as metrics to evaluate SAR in relation to design activities. Figure 3 shows the overlap of work and interests of the SPARK Project and this PhD. As shown, there is a large overlap; especially where data generation and collection, by means of studies, is concerned. It should be noted that, while a considerable portion of the work evidenced in Figure 3 is shared between this PhD and the SPARK Project, this does not mean that the work was led by different researchers or that it was not conducted with the overall PhD in mind. The figure serves merely to demonstrate the symbiotic relationship between PhD and SPARK Project. Appendix section J provides a breakdown of the author's contributions to the SPARK Project deliverables. The work performed for these deliverables was incorporated within the PhD.

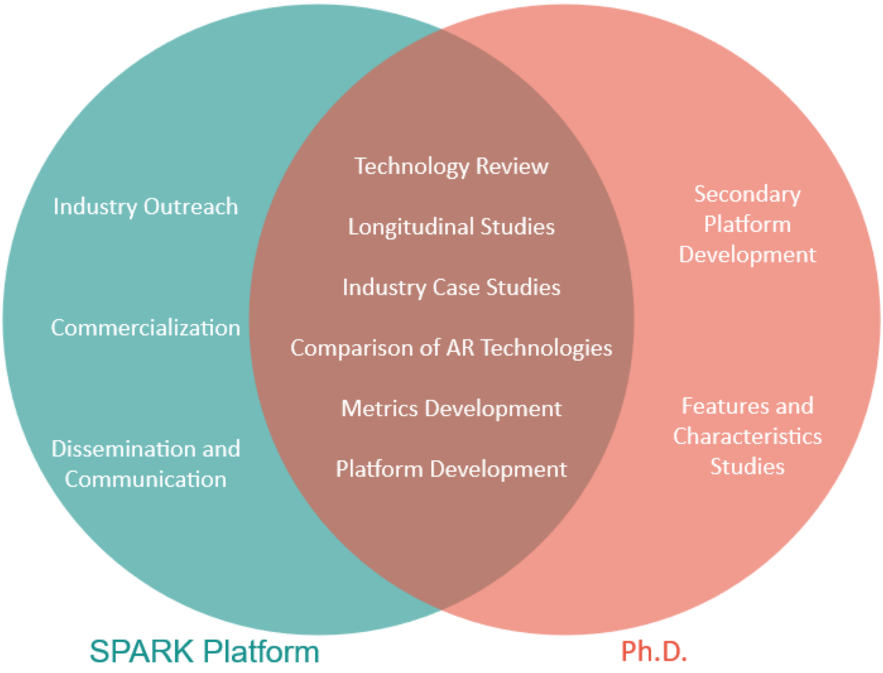


Figure 3. Graph showing overlap and separation of tasks and labour between PhD and SPARK Project

The SPARK project was motivated by a report by the European Union that identified the need for increased co-creativity and co-design in small to medium enterprises (SPARK Consortium, 2015). The report found that the majority of SMEs in Europe consist of micro and craft enterprises. These types of enterprises are on the smaller side of the Small-Medium Enterprises. The report continues by making a number of recommendations for the future of SME development in Europe advocating the investment in the development of skills for the workforce focusing primarily on co-collaborative and inter-disciplinary skills (Buschfeld *et al.*, 2011). The SPARK Consortium aimed to meet these needs by providing an SAR-based co-creation platform, the development of which is described in Chapter 4.

This PhD was envisioned as a continuation of work already performed as part of a prior MSc dissertation (Giunta, 2017), to offer blue-skies research capability and an opportunity to build on the achievements of the SPARK Project. As part of the MSc dissertation, metrics were developed that could be used to effectively and efficiently analyse design sessions involving AR. The initial motivation and driving force behind the PhD was a desire to utilise these metrics to further investigate the potential of AR as a support tool for design. It is important to note that, while the PhD was indeed partially funded by the SPARK project, and as such attempted to provide outputs relevant to it, its goals are considerably more research oriented than the goals of the SPARK project. Indeed, the SPARK Consortium's final goal is to develop a commercially viable platform, whereas this PhD is aimed at researching and investigating the applicability and usability of SAR tools as a method to support co-design sessions. It is, however, only natural that there is considerable overlap between the goals of the SPARK Consortium and this PhD.

The relationships are laid out in Figure 4 with the main goals for each of the three activities displayed. The activities are also laid out chronologically. As can be seen, the SPARK Project began prior to either the MSc or this PhD. The MSc was fully a part of the SPARK Project however, building upon the knowledge and research that had already been gathered as part of the SPARK Project. The PhD inherited some of the findings from the MSc and was partially conducted within the scope of the SPARK Project. However, the PhD's intention was always to explore beyond the exact scope of the SPARK Project and to continue with research aimed not only at commercialization (as previously noted in Figure 3).

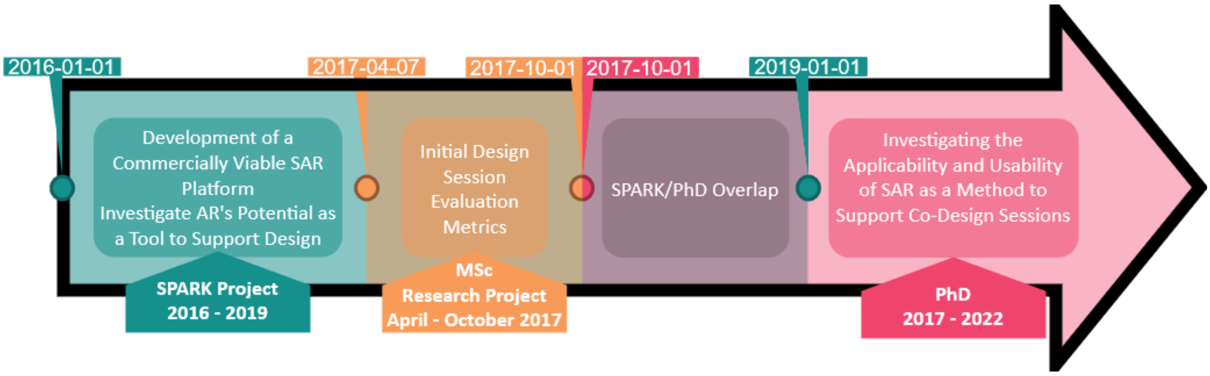


Figure 4. Timeline illustrating the relationships between SPARK Project, MSc, and PhD.

1.3.2 SPARK Platform Background

The major output of the SPARK Project was the SPARK Platform. This platform was designed to achieve the aforementioned goal of realising an ICT platform that utilises SAR to support collaborative design. Figure 5 shows the SPARK Platform as it was initially envisioned during the preliminary stages of the SPARK Project, prior to any technical development being undertaken. The figure illustrates how the platform is envisaged as a device to be used during co-design meetings, allowing multiple individuals to comment and review the prototype being worked on. Of note is the presence of a physical prototype, and the ability to interact with it in real time while the SAR projection is maintained. Figure 5 furthermore illustrates a preliminary concept for the interface, allowing for simple interactions with the graphics placed on the physical prototype to modify the SAR projection.

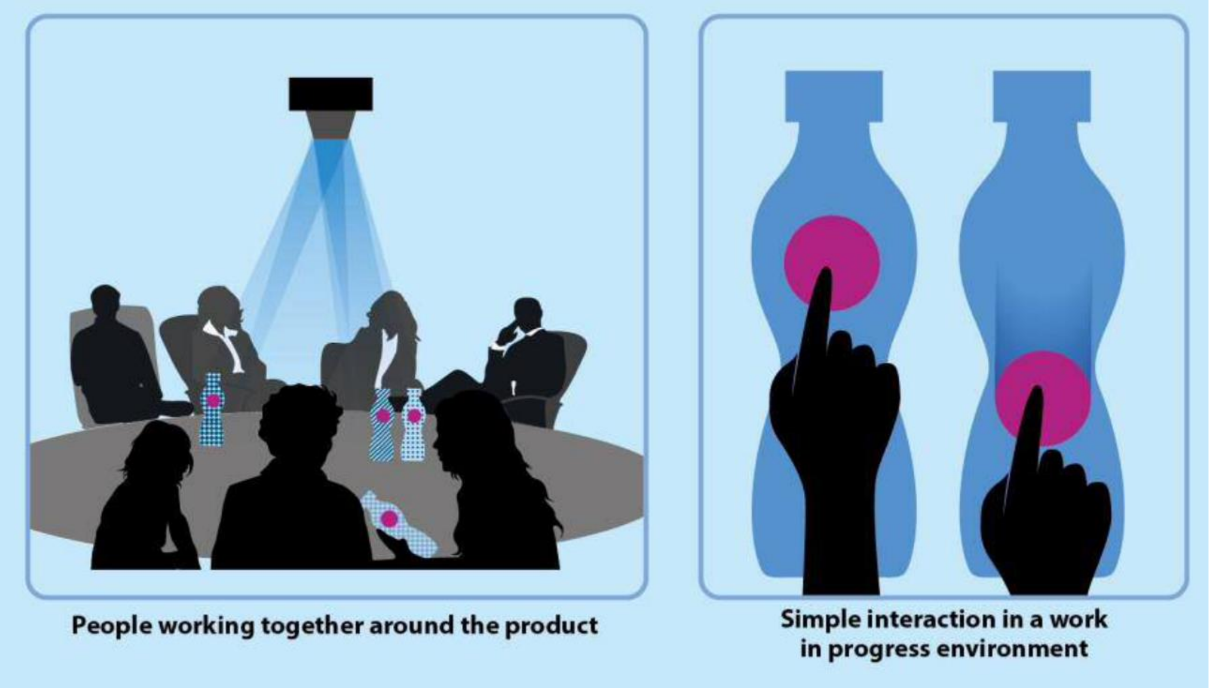


Figure 5. Illustration showing a preliminary concept for the SPARK Platform, showing the intended use for the platform (Left) and the interaction method (right) (SPARK Consortium, 2015)

Figure 6 shows the first iteration of the SPARK Platform. While later iterations of the SPARK Platform increased in complexity, the basic setup is laid out here: a Graphical User Interface (GUI) serves as a user input to control the SAR projection. This projection is made possible by a projector and projected onto a physical prototype. Subsequent iterations of the SPARK Platform (discussed in further detail in section 4) added tracking of the physical prototype, permitting users to interact with the physical prototype while maintaining the SAR effect. Furthermore, additional projectors were used to cover more angles of the physical prototype.

The SPARK Platform development was iterative. Feedback from users as well as members of industry led to the development of new features and capabilities, both for the SPARK Platform hardware as well as software. Installation and implementation of the platform was improved to better support SAR sessions. Once other such development was the creation of a portable SPARK Platform, termed "miniSPARK". Additional details regarding the process of development of this platform can be found in section 4.3.

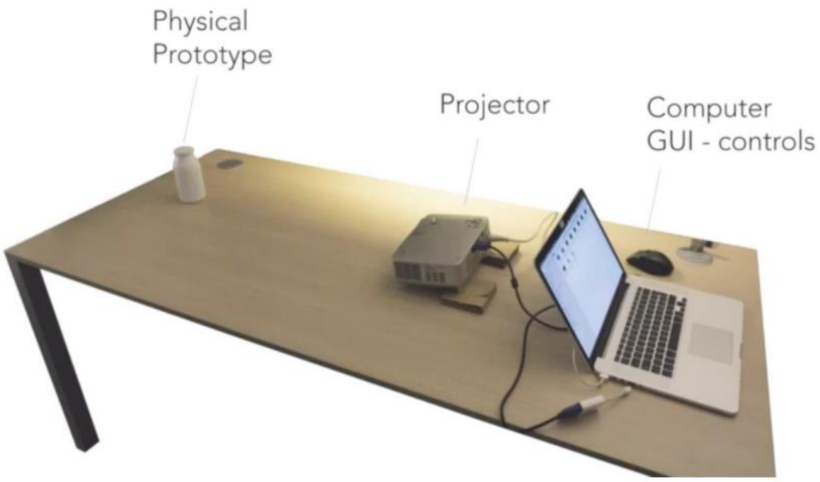


Figure 6. First SPARK Platform Iteration (SPARK Consortium, 2017)

1.3.3 Executive Summary of the SPARK Project

In summary, the SPARK Project aimed at developing a "Responsive ICT platform that exploits the potential of Spatial Augmented Reality for supporting and fostering collaborative creative thinking"(SPARK Consortium, 2015). The SPARK Project was administered by the SPARK Consortium. This was comprised of seven members (UBATH, POLIMI, GINP, AMS, Stimulo, Artefice Group, and Viseo). The project was initiated on the 1st of January 2016 and ran until the 1st of January 2019.

The main output of the SPARK Project was the SPARK Platform. The SPARK Platform was developed as a tool to utilise SAR to assist in collaborative design. The most basic components of the SPARK Platform, without which the SPARK platform cannot function, were: the GUI, the projector(s), and the physical prototype. The GUI permits users to interface with the SAR projection allowing them to modify it to suit their needs. The projector projects this SAR overlay onto the physical prototype which acts as a canvas. Additional functionalities were added to the SPARK Platform in order to allow it to track the physical prototype as well as permit projection from multiple projectors. These changes to the platform were due to feedback from members of industry as well as experiments undertaken to understand user interaction with the SPARK Platform. One such improvement was the development of "miniSPARK" a portable SPARK Platform.

1.4 RESEARCH FOCUS

As first discussed in section 1.1, The opportunity to further support co-design and the maturity of AR formed the drivers for this research. This is discussed in further detail in sections 2.1.2, 2.4, and 2.5. Therefore, the focus of this research is to investigate how SAR systems' characteristics and features affect collaborative product and packaging design. Equipped with this knowledge, one could seek to optimally configure SAR for co-design activities.

The research aims, questions, and objectives will be presented and discussed in full in sections 3.4.1, 3.4.2, and 3.4.3 respectively. The purpose of this section is to lay the foundations for the work that will be conducted in this thesis by providing a brief overview of the key terms and concepts.

In order to unpack the research focus, it is important to note the distinction between the characteristics and the features of an SAR system. The definitions used in this thesis are:

- **Characteristics –** The inherent properties that define the system as a SAR. Without these an SAR system could not be considered such.
- **Features –** Additional capabilities that are often accompany SAR characteristics but are not required.

For example, a characteristic of SAR would be how an SAR system is always independent of the users' point of view. Another example would be the presence of a physical projection model, i.e., the object onto which the projection occurs. For projective SAR to function a model must be present for the projector to project onto. While this model could take any number of forms or be made of any (sufficiently reflective) material it must be present for it to be a projective SAR system. An additional example would be the presence of projection. While the make, model, size, or type of projector does not matter in the slightest from a characteristics point of view the presence of a projector is an inherent characteristic of projective SAR without which the system could not be called an SAR system. Thus, in summary, characteristics of SAR are taken to mean any element of an SAR system that, if removed, would make it impossible to continue calling the system a form of SAR. A more comprehensive set of SAR characteristics is presented in section 4.1.1

Conversely the features of an SAR system are not inherent properties but rather elements or capabilities that often accompany a SAR system. A more detailed discussion of the typical features found in an SAR platform is provided in section 4.1.2. Their presence enables novel, customised and tuned experiences for given activities and opens new pathways to exploiting the SAR technology without changing the fact that it remains a form of SAR.

The simplest example of this would be: if an SAR system were to implement a colour correction system, to ensure high colour accuracy between a screen and the projection model, this could be considered a feature of that specific SAR system. This is because colour accuracy is not an inherent element of SAR. After all, the SAR effect can be achieved in shades of grey, or even with a terrible discrepancy between the interface and the model. A colour correction system is therefore a feature of an SAR system.

Another example would be the interface. An AR system must, by definition, be responsive and interactive. However, the type of interface to achieve this interactivity is not prescribed. As such, there is no specific limitation on how this interaction should occur. Theoretically using a single button as an interface that allows the users to interact with the digital projection would suffice. However, a screen where users can select a specific part of the model and easily change its colour to a new one would open the door to a much wider range of use cases. Consequently, the type of interface is a feature of an SAR system.

The purpose behind the investigation of the characteristics of SAR is to highlight whether the adoption of SAR brings with it any inherent issues or advantages. For example: projective SAR systems almost invariably use a physical model for projecting onto to create the Augmented Reality effect. If the presence of a physical model is taken as a given, for SAR to become more widely accepted in design it then becomes relevant to ask whether the presence of this model introduces new biases, or how it affects pre-existing biases in designers, and how this might impact the design session overall. A similar point can be made about the features of SAR, this time focusing on the new affordances that SAR provides that previous approaches did not; or did differently. Thus, looking at both the characteristics and the features, this thesis is able to investigate Spatial Augmented Reality for collaborative design.

1.5 THESIS STRUCTURE

Having identified the Research Focus, Figure 7 presents the structure of the thesis and is split into three phases. The first phase, from the start of the thesis up to and including the methodology, focuses on laying the groundwork for the remainder of the thesis. Here the overall scope of the thesis is formed (Chapter 1).

This is followed by the literature review (Chapter 2). Therein, the relevant literature is analysed to identify any knowledge gaps, as well as lay the foundations of the research. The main topics explored are:

- Collaborative Design (Section 2.1)
  Examines the definition of co-design and provides an explanation of the taxonomy thereof. Additionally, some common co-design techniques are provided as examples and analysed.
- Augmented Reality (Section 2.3)
  Discusses the taxonomy surrounding AR, and utilises this to analyse a number of different AR technologies.
- Augmented Reality in Design (Section 2.4)
  Explores implementations of AR as a tool for design collating and briefly discussing a number of papers to analyse the type of AR technology used and their potential impact for design.
- SAR in Co-Design (Section 2.5)
  Focuses specifically on literature relevant for the analysis and evaluation of SAR as applied to co-design sessions.
- Developing Design Research Metrics (Section 2.6)
  Analyses and discusses the development of the design research metrics utilised as part of the SPARK Project, describing their development and implementation.

Lastly, phase one concludes with the methodology for the thesis as a whole (Chapter 3). The chapter discusses the research aim, questions, and objectives and how these will be addressed by the subsequent body of work.

Thereafter, in phase two, the thesis focuses on the development of the SAR research platform (Chapter 4). All the different studies performed over the course of the PhD are described, their methodology explained, and their results laid out and discussed (Chapters 5, 6, and 7). Chapter 5 analyses the overall impact of SAR on design sessions. Chapter 6 analyses the impact of specific characteristics and features of SAR on the design process by means of more controlled studies. Lastly, chapter 7 discusses the industry feedback collected to understand the benefits of SAR, the potential barriers to implementation, as well as the necessary improvements required of the technology to overcome these barriers.

Finally, in phase three, the thesis concludes with a discussion and conclusion chapter (Chapter 8). The chapter provides a succinct summary of any important results as well as shows any interactions that have been noted across the various studies (section 8.1). The subsequent section focuses on summarising the overall findings of the study, tying them back to the research questions, aim, and objectives, evaluating whether these have been answered as well as highlighting additional knowledge that has been gained and any new knowledge gaps that have been identified (sections 8.1.1 and 0). Finally, the thesis concludes with an overall summary of the work presented within it, its limitations, the recommendations for the development of future SAR platforms, as well as presenting future work that could be conducted to expand on the thesis and the work presented therein (sections 8.2, 8.3, 8.4, and 8.5 respectively).

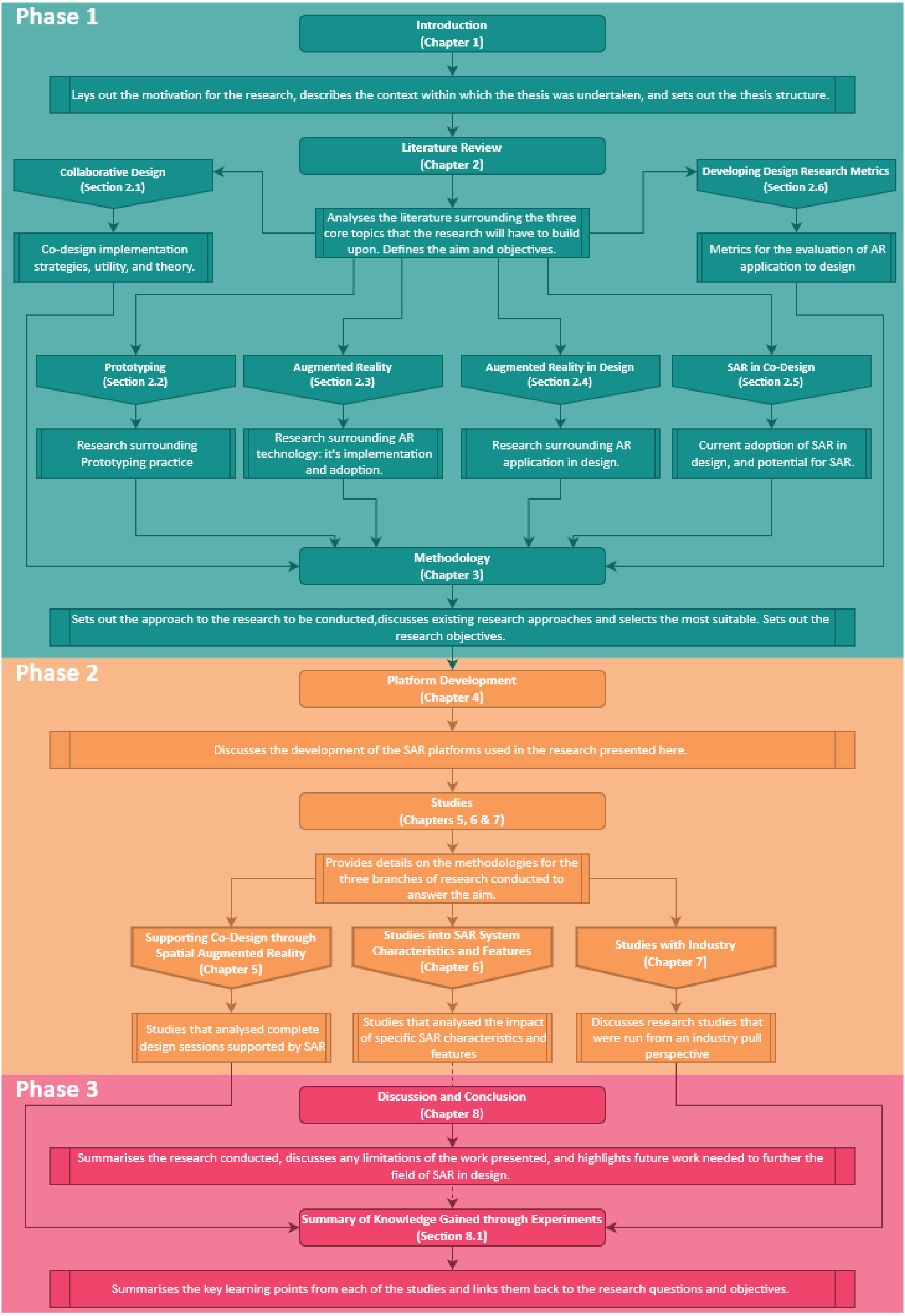


Figure 7. Thesis Structure

2 LITERATURE REVIEW

Before one can understand how SAR can support co-design, it is important to first review the state-of-the-art to arrive at the underlying research questions. In order to do so, the co-design literature is reviewed to:

- provide the definition for this activity that will be used throughout this thesis (Section 2.1.1);
- elicit co-design's benefits (Section 2.1.2); and,
- categorise existing methods of co-design (Section 2.1.2.1).

Secondly, section 2.2 discusses the literature surrounding prototyping and prototyping practice. This is done in order to contextualise the advantages of SAR as a tool to support prototyping and co-design.

Thirdly, the literature concerning AR has been evaluated to understand how these are categorised (sections 2.3.1, 2.3.2, and 2.3.3), and what the advantages and disadvantages of each AR technology are (Section 2.3.4).

Fourthly, an analysis of AR technologies currently used to support design has been performed to aid in further narrowing the scope of the research presented in this thesis and ensure that the research presented here is sufficiently novel (Section 2.4)

Finally, it is important to understand how others have studied the specific application of SAR in co-design (sections 2.5 and 2.6). It is important to note that these two sections draw heavily on the work performed by the SPARK Consortium during the development of the SPARK Project prior to the start of this PhD as first outlined in section 1.3.1 and illustrated by Figure 4. The work performed as part of the SPARK Project prior to the start of the PhD was foundational to much of the work performed during the PhD but was, however, extraneous to the work presented within this thesis. As such it has been located here, within the literature review, to allow the reader to understand the prior work performed without implying that it construed part of the PhD.

2.1 COLLABORATIVE DESIGN

The following section will examine the definition of co-design and provide an explanation of the taxonomy thereof. Thereafter some common co-design techniques are provided as examples. Relevant papers, which are discussed and analysed in sections 2.1.2.1 and 2.1.2.2, were found through Google Scholar, searching for terms such as "co-design", "descriptive", "prescriptive", "methodology", and "methods". In addition, the references used in the SPARK Literature review (SPARK Consortium, 2015) were analysed and used as a basis to find additional papers. Papers cited in this literature review that seemed promising were read and, on the basis of the citations made within each paper, new papers were found to be analysed.

2.1.1 Definition

The term co-design seems, at first, intuitive and has already been repeatedly used in this thesis. However, it must be noted that the simplicity of the term hides more convoluted origins. This is, in part, due to co-design's relatively recent origins. The first conference aimed specifically at the discussion of co-design, aptly titled CoDesigning, was held in 2000 (Scrivener *et al.*, 2000). The first journal aimed purely at the discussion of co-design, separately from the general field of design, was not published till 2005 (Scrivener, 2005). However, some of the techniques and approaches to design, that today would be labelled as co-design, have existed and been discussed in academia prior to this; though often without a specific framework to identify them as such (Valtonen, 2005; Szebeko and Tan, 2010). Indeed co-design, as well as Collaborative Creation (co-creation), were considered, until recently, synonymous to the umbrella term “participatory design” (Sanders and Stappers, 2008). However, these terms have since diverged to become separate entities, driven in part by the expansion of the field of human centred design.

As a result, Collaborative Design, or co-design, is not a term with a perfectly agreed upon definition (Ulrich *et al.*, 2003). Attempts have been made to clarify the term: Sanders and Stappers (2008) propose the definition of co-design as the application of collaborative creation during the entire design process. They state that co-design is “… collective creativity as it is applied across the whole span of a design process” and then further specify that “[they] use co-design in a broader sense to refer to the creativity of designers and people not trained in design working together in the design development process”. This definition by Sanders and Stappers (2008) has become widely, but not universally, accepted.

Figure 8 shows an overview of how Sanders and Stappers (2008) viewed the state of human centred design research. As can be seen from the figure, they highlight and group a number of different strategies for design involving users; participatory design is highlighted as a form of human centred design where the user is seen as a partner. However, in their graph, Sanders and Stappers (2008) do not directly describe collaborative design or collaborative creation. A prior publication (Sanders, 2006), where an initial version of this graph was presented, claimed this was a deliberate choice as the author stated that the current framework is not fully complete and was to serve as a foundation for further research and discussion. It is perhaps for this reason that, in their subsequent paper, Sanders and Stappers (2008) immediately follow up the presentation of the graph by providing their previously discussed definitions for co-creation and co-design. This all serves to highlight, however, the difficulty of defining these terms, and evidences the need for a coherent definition to be used throughout this thesis.

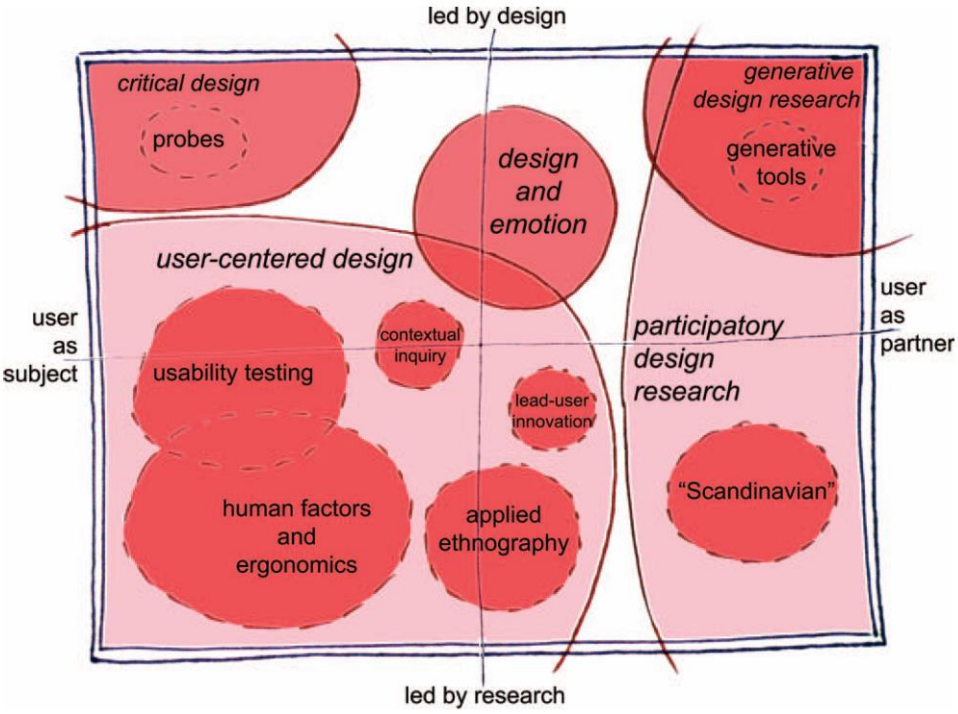


Figure 8. Sanders and Stappers's overview of the landscape of human centred design research (2008)

As such, for the purpose of this thesis, when the term co-design is used, it will be in accordance with the definition provided by Sanders and Stappers (2008) which refers to the collaboration of designers and/or non-designers during the design development process. It should also be noted that co-design is related to co-creation. Again Sanders and Stappers (2008) provide a definition by stating that co-creation can be considered "…any act of collective creativity, i.e. creativity that is shared by two or more people".

In summary: the umbrella term "participatory design" refers to any design process where the end user participates, in some way, in the design development process. Co-creation and co-design can be seen as distinct subsets of participatory design. Co-creation involves two or more participants collaborating in the creative aspect of the design. Co-design can be seen as a step further, as the application of co-creation throughout the entire design development process. These relationships are illustrated in Figure 9. In any one of these three scenarios, be it participatory design, co-creation, or co-design, the collaborative element can be between any two or more individuals. This is to say that there is no design or engineering knowledge requirement for a participant to be able to engage in participatory design, co-creation, or co-design: cross-disciplinary participants are perfectly acceptable. It should, furthermore, be noted that all the co-design techniques reviewed do not discriminate between types of design (product, packaging, UX, UI, etc.) and seem equally applicable in all these cases.

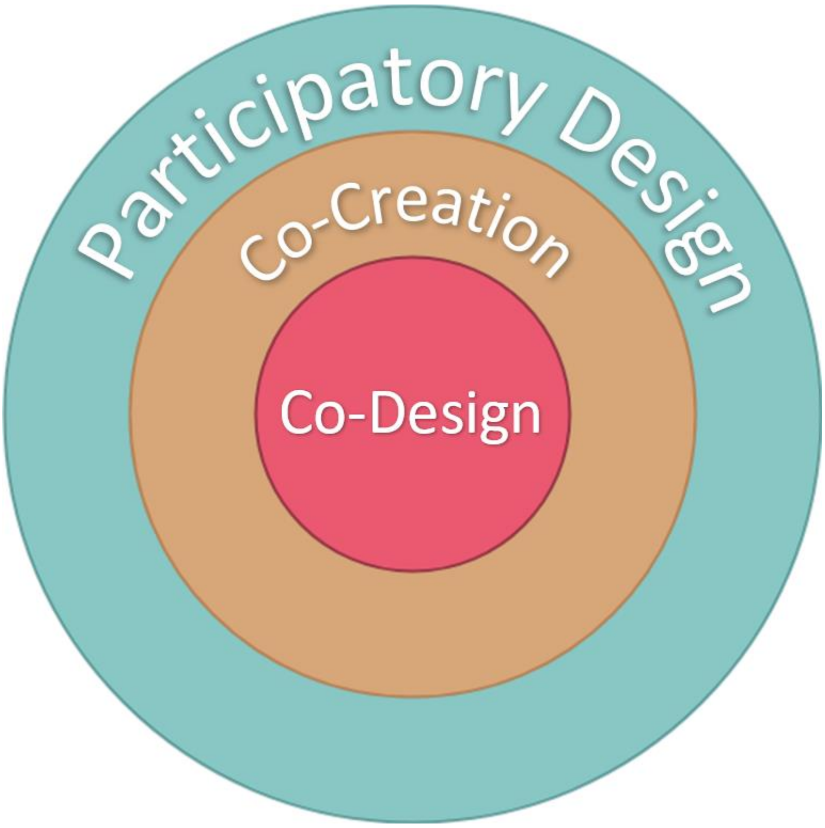


Figure 9. Diagram showing the relationships between Participatory Design, Co-Creation, and Co-Design

2.1.2 The Benefits of Co-Design

It was important to define these terms due to their, at times, confounding use in the literature. The most important point to address however is that, while there is evidence that substantiates the importance of collaboration to support value creation for customers and users (Prahalad and Ramaswamy, 2004; Payne *et al.*, 2008) the exact mechanism of action is not yet fully understood.

Research relating to co-design can be broadly categorised as either prescriptive or descriptive research. Prescriptive research focuses on analysing the impact of novel tools or methods aimed at improving co-design sessions both in terms of the activity and its outputs. Descriptive research analyses both the design development process, as well as the design sessions themselves. In doing so descriptive research attempts to expand the knowledge space and enables a deeper understanding of how co-design influences design sessions and their outputs. This provides a baseline that permits a deeper understanding of the co-design phenomena. Exploring some of the more relevant results of both approaches creates a more thorough understanding of what the impact of co-design is as well as providing insight into which areas are best suited, or most in need of, support from an SAR system.

The work of Shah et al. (2001) and Perttula, Krause and Sipilä (2006) are examples of the former category, prescriptive research aimed at testing novel tools methods that support co-design. Examples of prescriptive research are presented and discussed in section 2.1.2.1. Examples of the latter category, descriptive research, which refers to looking to map out and expand the knowledge space on co-design, are represented by the work of Franke and Piller (2004) and Gultekin-atasoy et al. (2014). They have focused on investigating the design sessions structure as well as its format. In addition, Mugge, Schoormans, and Schifferstein (2009) have analysed how the interaction between expert designers and end-users can impact the output of design sessions. Examples of descriptive research are presented and discussed in section 2.1.2.2.

2.1.2.1 Prescriptive studies in Co-Design

The number of possible design tools and methodologies that have, throughout the years, been proposed to support co-design is large enough that a detailed analysis is unfeasible; investigating and exploring all these would be beyond the scope of this thesis. As such the tools and methodologies discussed here will be based on the literature identified by the SPARK project (O'Hare *et al.*, 2016a) during its initial stages. Using this as a starting point, allows for a narrower scope that provides a link to some of the literature adopted as the underpinning of the SPARK project, thus having a direct impact on this thesis.

One of the most widely used methods is that of Osborn, who originally developed and popularised the use of "Brainstorming" (Pahl *et al.*, 2007). Brainstorming sessions rely on having the participants share ideas amongst themselves while being free from criticism. To ensure a criticism free environment, free even from implied criticism, it is necessary for the hierarchy within the group to be flat. During the session the ideas of the participants are recorded as they are produced and participants are encouraged to build upon each other's ideas (Pahl *et al.*, 2007).

A similar technique is known as the "Gallery Method" where participants sketch ideas silently and independently from one another. Once the sketches are complete, they are shared within the group with any missing or incomplete details being expounded upon. Constructive criticism and improvements are provided and once the discussion phase is complete the participants return to their individual sketches to make improvements based on the feedback received (Pahl *et al.*, 2007).

One other technique is known as the Collaborative Sketching (C-Sketching) method (Shah *et al.*, 2001). It consists of making the participants, independently from one another, create a sketch of their individual design solutions. Once all the participants have completed their sketch, they are passed around the table to the next person. Upon being passed to the next person they will be able to modify the sketch by drawing onto it, making additional changes and building on the ideas already detailed in the sketch (Shah *et al.*, 2001). Simply put, the C-Sketching can be seen as playing a game of "telephone" using sketches, only here the goal is not to convey the original message back to the start of the chain but to make incremental improvements upon concepts towards a concrete design.

Closely related to C-Sketching is the 6-3-5 method (Linsey *et al.*, 2005). The name of this method refers to the presence of 6 participants, who will each draw 3 solutions to the design problem, and then the proposed solutions will be passed around the group 5 times so that all participants can see all the proposed designs (see Figure 10). After each exchange the participants are invited to make modifications to the 3 concepts they have before them (Pahl *et al.*, 2007). While there are considerable similarities between the 6-3-5 method and C-Sketching, some important differences should be noted: Firstly the 6-3-5 method only allows for 6 participants. C-Sketching is more flexible in that any number of participants can be used. Secondly, C-Sketching relies only on sketches for communication between participants while the 6-3-5 method allows for participants to use words, or other methods, to describe their concepts.

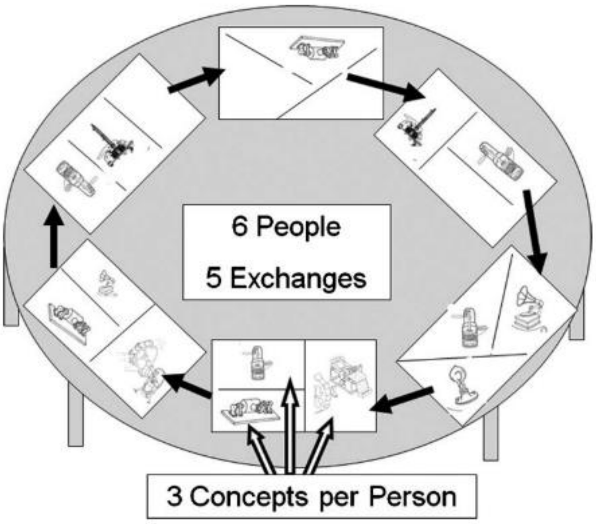


Figure 10. The 6-3-5 method illustrated (Linsey *et al.*, 2011); six participants each generate three concepts which are exchanged five times so that all participants can view and adapt the concepts.

All the approaches listed above, Brainstorming, the Gallery Method, C-Sketching, and the 6-3-5 method can be categorised as being either a type of Gallery or Rotational viewing (Linsey *et al.*, 2011). Gallery type viewing are defined by the fact that the participants have simultaneous access to all the concepts, sketches, or ideas that have been generated during the session and this simultaneous access is maintained during the discussion or critique phase. Conversely Rotational type viewing relies on the participants only having access to one set of concepts at any given time.

Linsey *et al.* (2005) found that techniques using Rotational type viewing, such as C-Sketching or the 6-3-5 method, resulted in greater design outputs when compared to Gallery viewing approaches. Further studies however revealed that, whilst the number of ideas generated is higher when using Rotational type viewing techniques, the quality of the output was higher when using Gallery viewing methods (Linsey *et al.*, 2011). The study also found that neither the novelty nor the variety of the ideas generated seemed to be influenced by the use of either a Rotational or Gallery type viewing. For these reasons, the study proposed using a hybrid system of both Rotational and Gallery type viewing techniques. These results have been corroborated by Perttula *et al.* (2006) who, in their study to analyse the impact of idea exchange on productivity, found similar results also recommending a hybrid method. In their paper, Perttula *et al.* (2006) hypothesise that:

> *"Individuals who are able to share ideas momentarily with other persons produce more ideas than those who work in solitude; however, idea exchange would not increase the variety of ideas produced."* (Perttula *et al.*, 2006)

In order to test their hypothesis, the experiment tested thirty-two mechanical engineering students, who were currently in the process of obtaining their master's degree. The authors (Perttula *et al.*, 2006) note that the participants can be considered to be novice designers based on their level of expertise. The experiment consisted of asking participants to generate ideas for an automated food package collection and sorting device. Three conditions were tested. In the first condition participants independently had to develop ideas. In the other two conditions participants were made to develop ideas individually, then share these in a group (either of two or three people according to the condition being tested) and then continue generating ideas individually. For each condition, the experiment analysed the quantity of ideas generated, the variety of ideas, and the non-redundant variety (i.e. ideas that were shared by members of the same group). The result of the study, as mentioned, was that collaboration led to a larger number of ideas being generated but did not necessarily result in a greater variety of ideas.

2.1.2.2 Descriptive studies in Co-Design

Franke and Piller (2004) looked to understand whether there was added value to participatorily designed products and, if so, to quantify said added value. To do so they looked at the study participants' 'willingness to pay' for a product as a measure of whether there was an improved value. The study was run by having participants design their own watch in an online system. The study found that participants in the study were willing to pay almost twice as much for a watch they had designed themselves when compared to a standard watch they had not designed. The paper argues that this behaviour is caused by a participant's personal investment in the design process even though the watch that was being evaluated was not their own design. The study also notes that there was considerable heterogeneity in the designs produced by the participants. This finding highlights the applicability of collaborative design techniques as a tool to improve the exploration of the design space when provided with the right opportunities.

The Business Innovation Observatory of the European Commission has also investigated the value that co-design and co-creation can have, as well as its potential impact on business-client relationships (Dervojeda *et al.*, 2014; Probst *et al.*, 2014). They found a shift in the paradigm that controls this relationship between businesses and customers. Designers, including professionals, are no longer seen as the sole creators and drivers of product innovation but are increasingly collaborating with end-users and clients to develop, design, and manufacture new and innovative solutions for current design challenges. This approach has allowed designers to better tackle the needs and meet the expectations of clients and end-users alike.

Others such as Mugge, Schoormans and Schifferstein, (2009) have, however, noted that the increased collaboration between designers, clients, and end-users can come at a cost. By including external sources within the design development process, the designers themselves lose control of this process and thus have a reduced impact on the final outcome. This lack of control can lead to untrained or inexperienced individuals falling prey to common pitfalls that more experienced design practitioners know to avoid resulting in extended product development times and non-optimal design outcomes. Gultekin-atasoy *et al.* (2014) highlight the balance between the foreign input and the more professional knowhow of the designers is one of the main challenges currently limiting outcomes for design development processes that rely on co-design. The technical knowledge, and wider perspective, of the professional designers helps reduce scope creep where the design is often favoured by the design session participants but does not meet the larger market needs. (Gultekin-atasoy *et al.*, 2014)

2.1.3 Summary

In summary, this section has highlighted the presence of discord within the design community with regards to an exact definition for the terms “co-design”, “co-creation”, and “participatory design. Subsection 2.1.1 highlights the definitions devised by Sanders and Stappers (2008) that were selected in order to harmonise the terminology used throughout this thesis.

In addition to defining co-design for use within this thesis, the literature review provided a basis for the value of co-design. In doing so it is possible to pre-empt any debate about the effective utility of using SAR to support co-design. Subsection 2.1.2 addresses this issue by providing examples of how applying co-design can improve design outcomes. Subsection 2.1.2 continues by identifying two main types of categories for analysing co-design studies, grouping them into prescriptive and descriptive studies in co-design. The former studies being more aimed at testing the validity of different techniques used to support co-design sessions and the latter at analysing the outcomes of sessions run using co-design tools.

Prescriptive co-design studies, discussed in subsection 2.1.2.1, analysed some methods for supporting co-design to understand how the sharing of ideas can affect design outcomes. Methods discussed in this section, such as “C-Sketching” and the “6-3-5 Method”, resulted in increased design output. However, whilst the quantity of ideas generated was affected, the variety of these ideas often remained unchanged. Perttula *et al.* (2006) corroborated these findings and concluded that a hybrid approach, not relying too heavily on either the Gallery or the Rotational viewing approach, was likeliest to yield the most desirable outcome of an increased number of ideas whilst also supporting a greater variety in the ideas generated.

Lastly, the analysis of descriptive co-design studies, discussed in subsection 2.1.2.2, attempted to provide a more detailed understanding of the costs and benefits associated with using co-design. Some of the drawbacks associated with co-design, in particular the need for experienced designers to relinquish control of the design process to inexperienced participants prone to pitfalls avoided by more seasoned individuals was highlighted (Mugge *et al.*, 2009; Gultekin-atasoy *et al.*, 2014). However, despite the drawbacks, there are clear advantages to co-design. Participants more involved in the design process are likelier to attach an added value to the product they have participated in designing, increasing their willingness to buy (Franke and Piller, 2004). Furthermore, there has been a clear shift away from designer-centric design and towards co-design (Dervojeda *et al.*, 2014; Probst *et al.*, 2014). This shift has allowed the designer to better target the needs of clients and end-users, resulting in products that more closely match the needs and wishes of the community.

2.2 PROTOTYPING

The following section discusses the literature surrounding prototypes and prototyping. This is done in order to establish the significance of prototypes and prototyping to the design process and thus the impact SAR could have thereon. The section first discusses the general literature and definitions of prototypes and prototyping (section 2.2.1). With the background established additional literature is highlighted pertaining to intermediary objects (section 2.2.2) and purposeful prototyping (section 2.2.3). Finally, a summary of the section is provided to highlight the major learnings of the section and their relevance to this thesis (section 2.2.4).

2.2.1 Prototyping Background

The terms “prototype” or “prototyping” lack precise definitions that are accepted by the research community at large (Jensen *et al.*, 2016). To add additional complication, it appears that the definitions vary across research domains (Beaudouin-Lafon and Mackay, 2012). In their analysis of common prototyping techniques, Mathias et al. (2018) attempt to provide a characterisation that can be used in lieu of a precise definition. In their paper they discuss the existing schools of thought and, citing Ullman (2003), Mathias et al. (2018)arrive at a categorisation that can serve to distinguish types of prototypes by their purpose:

- **Proof-of-Concept:** […] is used to identify what approach to take in the initial stages when designing a new product.
- **Proof-of-Product:** […] helps develop the physical embodiment and manufacturing viability.
- **Proof-of-Process:** […] demonstrates that the chosen materials and production methods meet the product requirements.
- **Proof-of-Production:** […] shows that the complete production process can achieve the required results.

Camere and Bordegoni (2016), however, describe the prototyping process not just by the output but as the "activity of engaging with the product-to-be, instantiating the design process." (Camere and Bordegoni, 2016). In response to this view, Mathias et al. (2018) attempt to define a prototype by expanding on the six characteristics described by Jensen et al. (2015) (material, interactivity, visual detail, purpose, surroundings, and technology).. This is achieved by the addition of a new dimension: technique. Here technique is defined as either the method by which the prototype is created or the tools used to accomplish the creation of the prototype, as discussed by Blomkvist and Holmlid (2011) and Hallgrimsson (2012).

Nonetheless, while it is perhaps possible to categorise prototypes, the prototyping process itself is even less uniform in methodology and implementation. This is exacerbated by the lack of data surrounding the prototyping process, as once the prototype has served its purpose, it is often discarded as there is often little value in communicating with parties external to the design team (McAlpine *et al.*, 2006).

2.2.2 Intermediary Objects

Intermediary objects were first described by Vinck and Jeantet (1994) in an attempt to address what they saw as a “disastrous opposition” between technical and social aspects of design. In their paper, Vinck and Jeantet (1994) argued that the gulf between social and engineering sciences in design caused problems when needing to reconcile the different aims, research approaches, and value judgements. These differences cause friction during the design process, reducing the likelihood of good output. The intermediary object thus acts as a tool for the facilitation of knowledge transmission and communication between participants.

Papadimitriou and Pellegrin (2007) connect the concept of intermediary objects to that of boundary objects first introduced by Star and Griesemer (1989). Papadimitriou and Pellegrin (2007) argue that the intermediary objects, unlike boundary objects, are less prone to neglect as they may be less charged with meaningful action.

Boujut and Blanco (2003) expand on this by analysing how computer aided design can be used as a intermediary object to support cooperation during the design of parts. While the intermediary object encompasses more than just prototypes, prototypes fall within the category of intermediary objects as they support and facilitate the exchange of ideas and information during the design process (Lauff *et al.*, 2018).

As such, the literature surrounding intermediary objects highlights the importance of prototyping as an invaluable tool for communication between design session participants (Lauff *et al.*, 2017). By being able to influence and modify the prototype using SAR it may be possible to improve the value of the prototype as an intermediary object, allowing for more efficient and effective communication of ideas and intent. This would, in part, be driven by the non-static nature of SAR prototypes which can be modified on the fly, unlike other more traditional prototypes.

2.2.3 Purposeful Prototyping

Petrakis *et al.* (2019) in their discussion of the development of a new taxonomy for prototypes introduce the concept of the "purposeful prototype". They describe this as "[…] the establishment of the prototype's purpose, prior to initiating the prototyping process". In doing so, Petrakis *et al.* (2019) argue that the approach requires additional reflective thinking. Thus, aiding to decouple the process whereby the prototype is created, and the knowledge gained thereby, from the actual intended use of the prototype.

This more thoughtful approach to considering the purpose of prototypes is echoed by Lauff *et al.* (2019) who caution against the sunk costs of prototyping merely for the sake of prototyping. Lauff *et al.* (2019) suggest that when prototyping becomes the goal or focus of a design activity, this can lead to inefficiencies in the design process. They argue that, by paring prototypes down to a "minimum viable prototype" focus can be maintained on the specific outcomes desired from the prototyping activity. Hansen *et al.* (2020) built on these findings by running an empirical study comprised of 125 novice designers using a tool aimed at supporting purposeful prototyping. Their study found that participants reported that the tool developed aided them in purposeful prototyping and that, as a result, communication between team members was improved.

2.2.4 Summary

In summary: there is presently no universally accepted definition for the characterisation of prototypes and prototyping. It is, nonetheless, possible to focus on certain specific characteristics of prototypes to categorise them (Mathias *et al.*, 2018).

However, from the research surrounding intermediary objects it is possible to ascertain that prototypes play a vital role in the communication of ideas and intent between participants in design sessions. The role of prototypes is also an important aspect to consider. The simple generation of prototypes does not necessarily result in superior design outcomes. Purposeful prototypes allow participants to maintain focus on the expected outcomes of their work. This is valuable for the development of an SAR platform as it highlights the value of clear and effective prototypes and their positive impact on communication.

2.3 AUGMENTED REALITY

Before beginning to discuss Augmented Reality, it is important to understand what Augmented Reality (AR) is and the terminology used. Figure 11 shows the Reality-Virtuality Continuum, created by Milgram *et al.* (1995), to help understand the relationships between the Real and Virtual Environments. The Real Environment is the world, as all humans perceive it with no foreign augmentation of any kind. The Virtual Environment is an environment that has been computer generated and completely immerses a user when interacting with it. Figure 11 shows these as the two extremes of the continuum. Between these lies all of “Mixed Reality” (MR). This term is used to refer to technologies that are a combination of both virtual and physical. It is within the MR segment of the continuum that AR can be found. AR lies closer to the Real Environment than the Virtual Environment along this Continuum. This is because AR is the augmentation of the Real Environment with Virtual elements. This can be achieved by any means but the core element which must be present to consider a technology a form of AR is the presence of the Real Environment and the use of Virtual elements to enhance this (Milgram *et al.*, 1995).

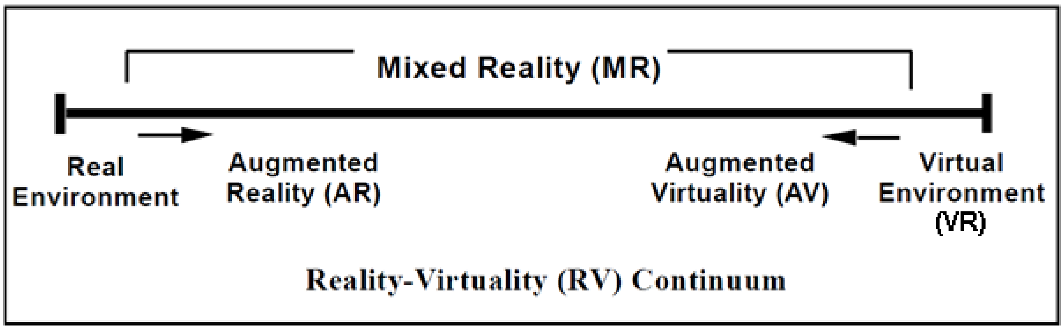


Figure 11. The Reality - Virtuality Continuum (Milgram *et al.*, 1995)

With the terminology defined, it is possible to begin to review technologies that produce AR environments. Bimber and Raskar (2006) provide a categorisation for AR technologies (Table 1). Their taxonomy focuses on the application of the technology relative to the user. The three main categories are: Hand-Held Displays (HHD), Head-Mounted Displays (HMD) (also referred to as head-attached) and Spatial Augmented Reality (SAR). The categories are then further divided into sub-categories based on the type of technology used to achieve the AR effect.

Table 1. Taxonomy Proposed by Bimber and Raskar (2006)

HEAD-ATTACHED DISPLAY			HANDHELD		SPATIAL	
Retinal	Head Mounted	Projective	Hand-Held Display	Hand-Held projector	See-Through Display	Projective

Figure 12 provides an illustration of this taxonomy and helps contextualise it better by illustrating the user’s perspective and how each technology would be placed relative to that. A more in-depth discussion of how each technology functions can be found in sections 2.3.1, 2.3.2, and 2.3.3. Those sections go into greater detail, and provide examples of, HMDs, HHDs, and SAR systems respectively.

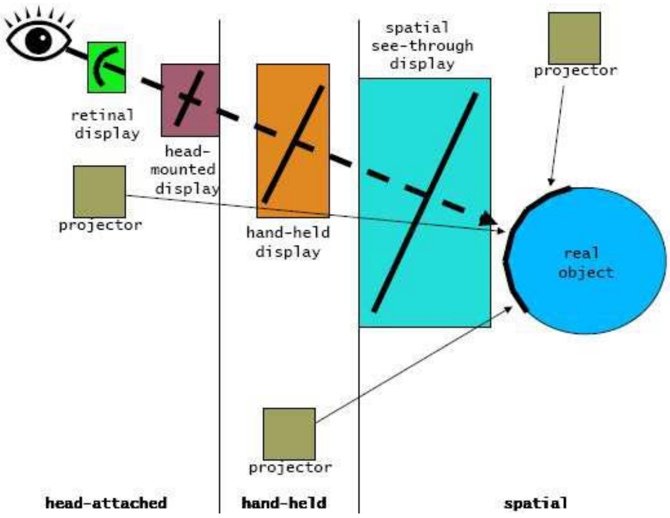


Figure 12. AR technology taxonomy proposed by Bimber and Raskar (2006)

Van Krevelen and Poelman (2010) have further expanded Bimber and Raskar's (2006) work by refining the definitions to provide descriptors that are more reflective of the technologies that were in development at the time. Table 2 shows this updated taxonomy; additions to the taxonomy include:

- "Video" element added to head worn technologies
- Hand-Held Display and hand-held projector have been merged into a single "All" category
- "Spatial" section has the see-through display relabelled as "Optical"
- A new element, "Video", was added to the "Spatial" category

Table 2. Refined Taxonomy Proposed by Van Krevelen and Poelman (2010) based on the original definitions of Bimber and Raskar (2006)

Head-Worn				Handheld	Spatial		
Retinal	Optical	Video	Projective	All	Video	Optical	Projective

It is in this further subcategorization that Van Krevelen and Poelman (2010) differs from Bimber and Raskar (2006). In their original work Bimber and Raskar made a further distinction between types of head mounted displays (this further subcategorization is not shown in Figure 12) by attempting to subdivide between LCD based systems and video-based systems. Van Krevelen and Poelman (2010) did not make this distinction, due to the very close similarity between LCD systems and retinal displays making it hard to draw a line between the two systems.

In summary:

- The model shown in Figure 12 originates from Bimber and Raskar (2006).
- Van Krevelen and Poelman (2010) provided an update of this model, summarised in Table 2.
- While the underlying model has changed due to some refinement in the terminology used, the graphic shown in Figure 12 remains unchanged since, when the graphic was originally

made Bimber and Raskar (2006) it did not originally make the distinction between LCD vs retinal in their graphic.

- This seemingly small change is nonetheless relevant, the smallness of the change suggesting that the model is mostly complete and worth using. As such, the model shown in Figure 12, is valid for both the taxonomy proposed by Bimber and Raskar (2006) and Van Krevelen and Poelman (2010).
- Only the underlying definitions, as highlighted in Table 2 change between the two.
- Thus the model proposed by Van Krevelen and Poelman (2010) was chosen over the one by Bimber and Raskar (2006) due to its simpler and more refined nature.

Peddie (2017), citing Padzensky, similarly discusses an extended taxonomy for augmented reality devices primarily focusing on whether the technology is worn and the means of interacting with the technology by sensing it and controlling it (Figure 13, showing the use of the type and sensing categories to group AR technologies). This taxonomy shares similarities with that of Van Krevelen and Poelman (2010) in their top-level ordering: Van Krevelen and Poelman suggesting Head-Worn, Handheld, and Spatial as top-level orderings; and Padzensky proposing Head Mounted, Mobile Non-wearable, and Projected. However, the two begin to differ in the subsequent sub-categorization. Beyond the type category, one observes the two taxonomies diverge with Padsnzky categorising the technologies by Design, Form Factor, Class, Sensing, and Control.

These additional sub-categories provide additional nuance when it comes to the classification of AR technologies, especially Head Mounted systems, but add complexity in identifying the correct sub-category. Furthermore, some of the distinctions made seem to rely on somewhat arbitrary decisions. For example, for the Mobile Non-wearable Type, the Form Factor category discriminates between tablet devices and smartphones. However, labelling a device as a smartphone or tablet is subjective, especially when one considers that most tablets can act as smartphones and that many smartphones are large enough that some might consider them tablets.

Furthermore, Padzenky's taxonomy seems to focus much more on categorizing Head Mounted systems rather than Projected systems. Whilst, van Krevelen and Poelman's (2010) categorization makes a distinction between optical and video systems, which Padzensky's taxonomy does not. Furthermore, Padzensky's (cited in Peddie, 2017) taxonomy does not apply the Design category to Projective AR. Additionally, the Class category is not applied. As such the only meaningful way of distinguishing between types of Projective AR using Padzensky's taxonomy is by assessing the Form Factor, Sensing and Control used. Padzensky divides the Form Factor category for Projected AR into three sub-categories: Surface, Vehicular Windscreen, and Retro-reflective. Respectively these sub-categories refer to:

#1 Surface: An AR system that projects onto a flat surface. E.g.: a keyboard that is projected onto a flat surface and detects user input as the virtual keys are pressed.
#2 Vehicular Windscreen: A system where the AR projection is placed onto a vehicle's windscreen. E.g.: Continental's AR Heads Up Display (HUD)
#3 Retro-reflective: Any kind of system where projections are reflected back from an existing object or shape (barring a flat surface). E.g. a head-mounted system with integrated projectors to augment the objects the wearer is looking at.

This approach to categorizing Projected AR systems becomes somewhat restrictive, the first two subcategories, Surface and Vehicular Windscreen, seem to be very narrow, seemingly focused on the implementation of the system rather than its underlying technology. For example, any Retro-Reflective system would immediately be considered a Vehicular Windscreen system if it were installed in a car, without any underlying change to the technology. The distinction made between Surface and Retro-reflective is also not very clear: any 3D object will be made of surfaces. As such, projecting onto only one surface would make that system fall under the surface subcategory whereas projecting onto multiple surfaces of the object would make it fall within the Retro-reflective subcategory. Furthermore, having an entire category dedicated only to the use of AR in vehicles, and specifically the vehicles windscreen, seems overly narrow and specialised in particular when the context of this thesis is taken into account.

Lastly, Padzensky’s method of further subcategorizing Projected AR systems by looking at the type of Sensing and Control used, appears - from the examples provided - to add restrictions beyond that of the other classes. For example, according to the taxonomy, a Retro-reflective system cannot make use of Geo-positioning to sense its environment. This seems to bring in an unnecessary restriction that has no technical case for the mutual exclusion. Multiple use cases could be imagined for a portable Projected AR system that tracks the user’s location within a building to adapt the projection made in real time. This would, for example, allow the Projected AR system to change its functionality depending on the location it is in, as might be required in a factory along an assembly line.

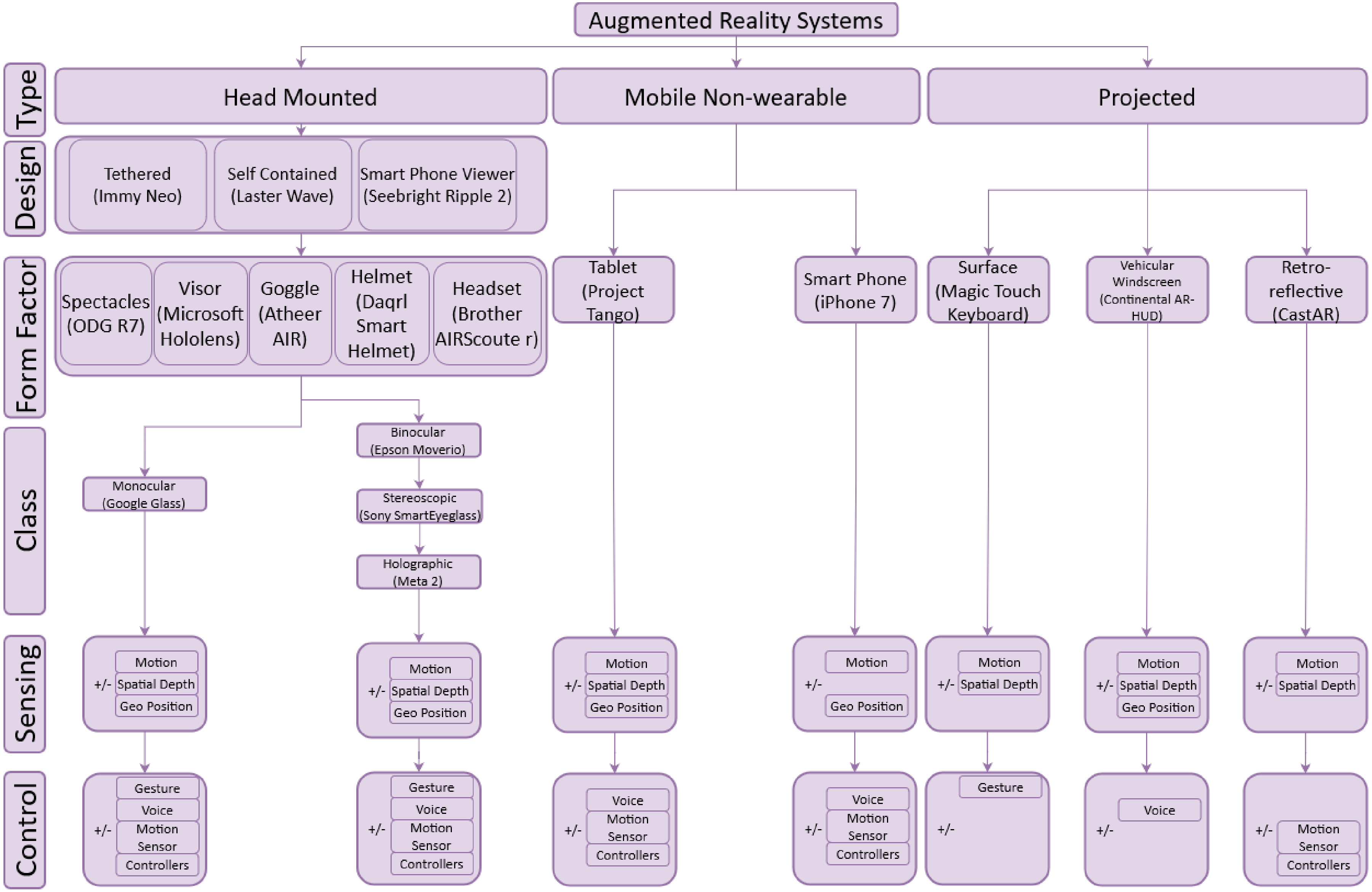


Figure 13. AR technology taxonomy proposed by Padzensky (cited in Peddie, 2017) names between brackets are exemplars of each subcategory

Following the review of AR taxonomies, the categorization developed by van Krevelen and Poelman (2010) with adaptations from Padzensky's work was deemed most appropriate for this research. The clear subdivision, based on the technology used to display information, provides an objective measure that can be mapped to reliably.

Indeed, the use of Padzensky's taxonomy helps overcome a failing of Krevelen and Poelman's taxonomy. Krevelen and Poelman's work, conducted more than a decade ago, could not take into account the modern adoption of smartphones. Their ubiquitous presence in everyday life has opened new portals to the exploration and exploitation of AR technologies. Grouping all handheld systems into the same category, while certainly sensible at the time, now seems less reflective of the current state of AR technology.

For this reason, the author's adaptation is to split the handheld category into two sub-categories: Scene Augmentation and Object Augmentation. The former refers to scenarios where virtual objects are placed within a real scene, whereas the latter refers to the augmentation of real objects, within a real scene, with digital elements. Table 3 summarises the taxonomy and terminology that will be used throughout the remainder of this thesis. Sections 2.3.1, 2.3.2, and 2.3.3 below further explain and provide examples for each of the sub-categories.

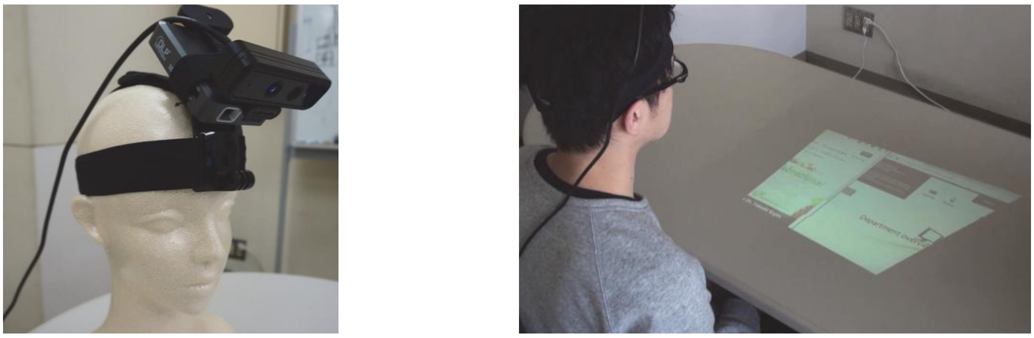
Table 3. Taxonomy used in this thesis

HEAD-MOUNTED DISPLAY				HANDHELD		SPATIAL		
Retinal	Optical	Video	Projective	Scene Augmentation	Object Augmentation	Video	Optical	Projective

2.3.1 Head-Mounted Displays

There are a number of types of Head Mounted Displays (HMDs). In this section, the following types will be discussed: Projection, Retinal, Optical, and Video.

Projection HMD AR is achieved by wearing one or more projectors (Figure 14a). These projectors then project a digital overlay onto the physical world that is being viewed by the user (Figure 14b). Some form of tracking is usually implemented to allow the images to adapt to their surroundings, although exceptions do exist (Cortes *et al.*, 2018).

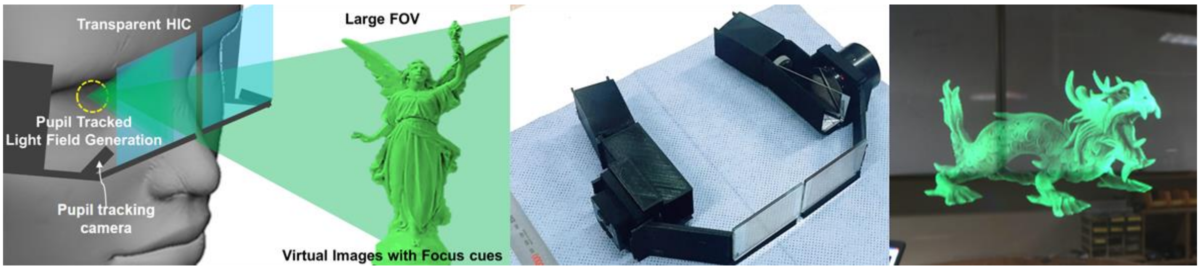


(a) Prototype Setup **(b) Prototype in use**

Figure 14. Prototype Projection HMD developed by Kemmoku and Komuro (2016)

Retinal HMDs project directly into the user's eyes. To aid in visualizing this it helps to contrast it with how a television, or monitor, works. There, an image is created on a screen and the light then travels to the human eye. The focal point of the image is the reflective screen itself. Retinal systems create the image not by first putting it in focus on a screen but by projecting it directly into the human eye (de Wit, 1999) in such a way that the focal point of the image is exactly on the retina. In doing so, they are able to mix incoming light from the physical world with the light from the projectors. This gives the illusion that the digital world has overlaid or replaced parts of the real world as shown in Figure 15c. It is important to note the distinction between optical systems, and retinal systems. While both may make use of a semi-transparent film to blend the real and virtual worlds, retinal systems will always ensure that the virtual image is focused directly in the eye.

Various technologies exist that can achieve this effect (Cakmakci and Rolland, 2006). Figure 15 shows one example where this is achieved by having a set of semi-transparent screens placed before the user's eyes, at times called Holographic Image Combiners (HIC), as is shown in Figure 15a. Digital images can be projected onto these screens which, due to their semi-transparent nature, will blend them with the incoming light from the user's surroundings.



(a) Overview of how a Retinal HMD works **(b) Retinal 3D prototype** **(c) User view with Retinal 3D**

Figure 15. Overview of a Retinal HMD developed by Jang et al. (2017) called Retinal 3D

Figure 16 shows the functioning of the Retinal HMD used in the Retinal 3D headset (Jang *et al.*, 2017). The Laser Beam Projector (LSP) mounted on the headset contains three Laser Diodes (LDs). For each of the LDs the emitted light is passed through a set of beam shaping lenses (BLs) and then combined by a set of dichroic mirrors (DMs). Once the light exits the LSP by means of a mirror (M1) it is passed through a set of lenses, L1 and L2, that shape the beam. The light is then passed through an Attenuation Filter (AF) and a Colour balancing Filter (CF). The light is then reflected off a Half Mirror

(HM) and onto a fast moving electronically controlled steering mirror (M2) it is this mirror that allows the light to always be shone directly into the pupil even as the eye moves. The light then passes through the HM and onto the Holographic Image Combiner (HIC) where the light is mixed with the light from the surrounding environment (Jang *et al.*, 2017).

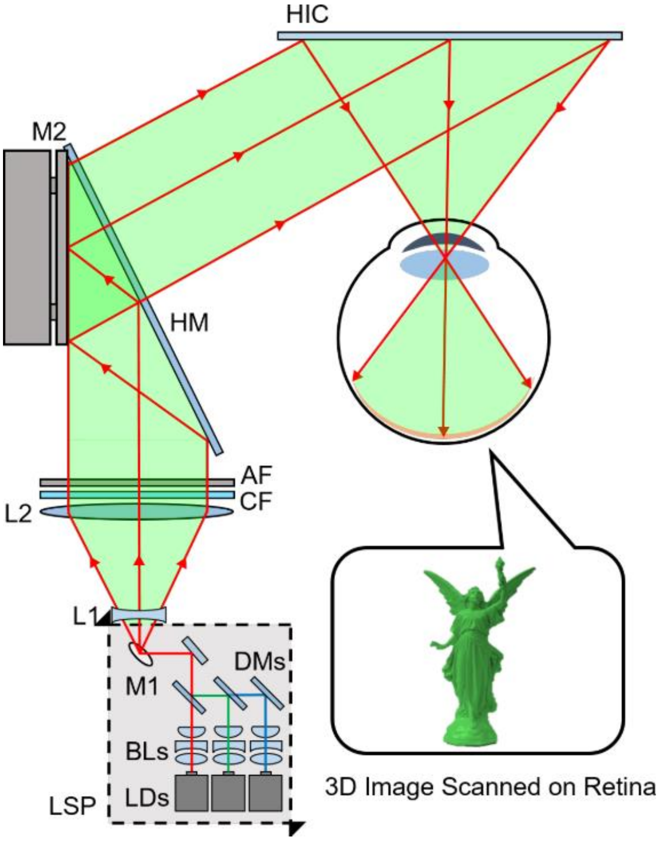
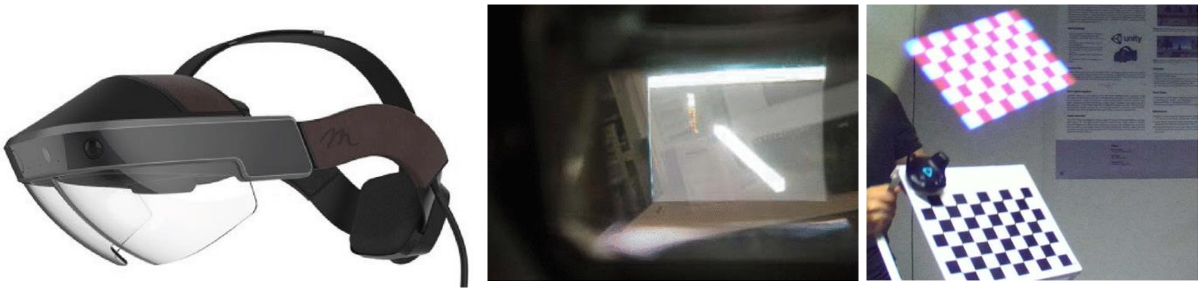


Figure 16. Retinal projection system used in the Retinal 3D HMD (Jang *et al.*, 2017)

The method described in Figure 16 serves as an example of how a Retinal HMD can function, different devices will use slightly different approaches. The figure serves to highlight the challenge in achieving the focus of the image on the retina itself rather than on the semi-transparent screen used to combine the light from the surrounding environment and the projected light.

Optical HMDs function by using a semi-transparent screen onto which they create a virtual image that allows the real and virtual scenes to blend and be seen as one by the viewer as shown in Figure 17b (Cakmakci and Rolland, 2006). Figure 17a shows one such device; the goggles at the front act as the semi-transparent layer. The headband contains all the necessary infrastructure for the AR effect to be obtained. Figure 17c shows how the AR effect appears to the user and is comparable to that seen in Figure 15c. Figure 17c also shows the AR effect during the calibration phase. The black and white checkerboard represents a physical object currently being tracked by the Optical HMD. The red and white checkerboard is the virtual image being created by the HMD to augment the physical object. As the calibration phase is not yet complete, the two checkerboards do not match up. Once the process has been completed the augmented checkerboard will follow the real object as it is moved giving the impression that the physical object is actually red and white instead of black and white.

Commercial examples of the technology include Microsoft's Hololens, LusoVU's Eyespeak, and Google's Google Glass (Peddie, 2017). It must be noted that Google Glass has been relaunched (Savov, 2017). This relaunching however demonstrates that companies remain convinced of the potential of the technology. The continued development of new devices shows that the technology is still relatively immature and that a dominant design has not yet been established in the market.


(a) An example of an optical HMD (Syberfeldt *et al.*, 2017)

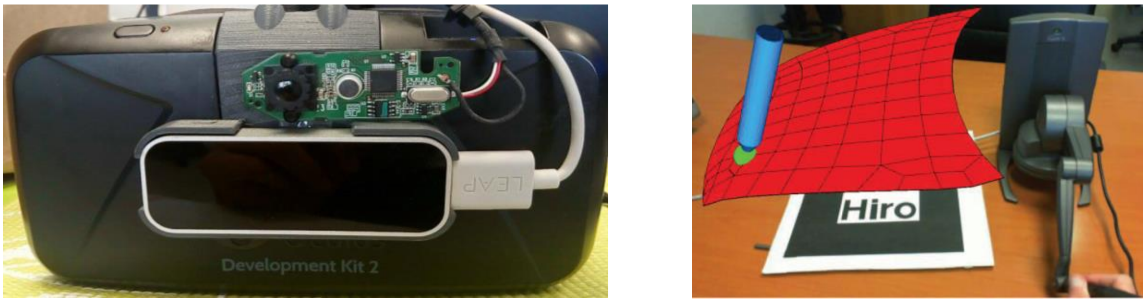
(b) Semi-transparent screen with virtual scene projected as seen by user (Zhen *et al.*, 2017)

(c) User's perspective of the AR effect here shown during calibration (Ballestin *et al.*, 2019)

Figure 17. General overview of optical HMD systems

Video HMDs are similar in function to VR headsets (Figure 18a). Like VR headsets, they only transmit digital images by way of a screen. However, unlike VR, users still see their surroundings due to a live video feed. AR comes into play when parts of the visuals shown to the user are modified to intersect a digital reality with the physical one as shown in Figure 18b.

Many VR headsets, such as the Oculus Rift (Figure 18a) the HTC Vive, and the Steam Index, can act as a Video HMDs. This is as many of these headsets are natively equipped with cameras which can "pass through" video of the user's surroundings.



(a) Video HMD Prototype. An Oculus Rift with a camera attached. (Wai and Abd Manap, 2018)

(b) User perspective when using Video HMD (Valentini and Biancolini, 2018)

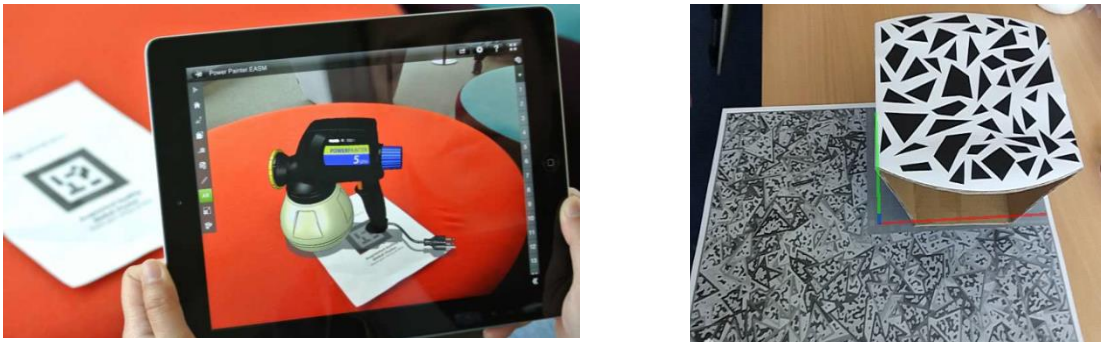
Figure 18. Video HMDs often function in the same way as VR platforms but make use of a camera feed showing the real world with augmentations rather than an entirely virtual world.

2.3.2 Handheld

Hand-Held Devices (HHD) are one of the more common methods for implementing AR (Dey *et al.*, 2018) due to the ubiquitous nature of smartphones and tablets. They provide a simple, low-cost platform for AR. In a survey of U.S. VR and AR gamers, 77% of a 1000 respondents stated that they made use of smartphone based VR and AR systems (Statista, 2018). HHDs can be divided into two types: Object Augmentation and Scene Augmentation systems. Object Augmentation functions by overlaying information onto physical objects (e.g.: texture and information tooltips), whilst Scene Augmentation places digital objects within a physical scene. The nature of HHD's means that the AR effect is only visible through the device itself and thus requires users to hold up the device within their field of view to maintain the illusion.

A common method for achieving Object Augmentation is through visual markers, for example QR codes, placed on the objects to be augmented. Figure 19a shows the application of a QR code on a sheet of paper. When the QR code is identified by the tablet's camera, a digital overlay of the object associated with the QR code is placed into the scene. This method requires an unbroken line of sight between the QR code and camera to maintain the Scene Augmentation effect. Another approach is to cover an existing object with an irregular pattern to break-up the otherwise regular surface of the object. This "covering", shown in Figure 19b, allows the entire physical object to be tracked and potentially be fully overlaid.

Both situations require calibration, so the AR software knows how to position the digital overlay relative to the object. Figure 19b shows the initial stage of this process where the model is placed on a reference plane and is then scanned so that the software can recognise the pattern and relates this to the reference place.



(a) Combining a HHD and a QR code to create AR (Chin, 2013)

(b) Calibration of an irregularly patterned object

Figure 19. Tracking objects using different techniques

An approach that avoids the use of visual markers is through image recognition. While this eliminates the reliance on visual markers it adds to the difficulty of implementing computer vision. Notwithstanding, it has found use in assembly and maintenance (e.g.: Patent 9,448,758 (2012)). In this situation, the camera will recognise objects and track them to overlay the digital information. A commercially available example is the IKEA Place app, shown in Figure 20, which uses the camera to detect the room size and angle relative to the camera. In doing so the app can calculate the position of the digital objects to overlay, placing them into perspective and scaling them to the correct size relative to the other furniture in the room. In contrast to the object detection, which recognises specific objects and tracks them, the IKEA Place app overlays a digital object on the scene, using the camera only to adjust the perspective of the object so that it appears to be properly aligned with the scene.

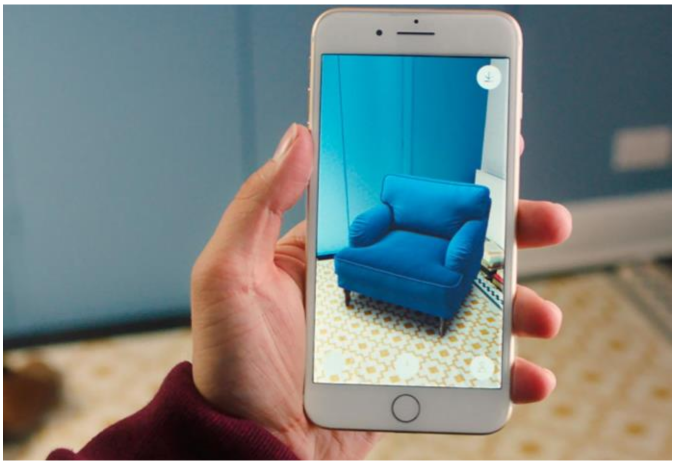


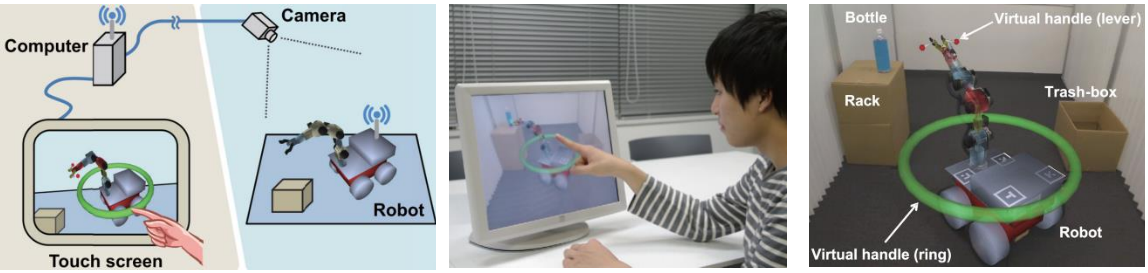
Figure 20. IKEA Place App showing a virtual couch being placed over a rug (Dasey, 2017)

Location-based tracking has also been used to enhance the AR experience. Most commonly used in outdoor settings, this approach to AR relies on mapping the location of users and then presenting them with images based on their current position. Examples of this would be Pokémon Go or Minecraft Earth. These apps allow multiple users in the same location to view the same digital object (a Pokémon or Minecraft blocks). They can interact with this object and the interactions will be transmitted in near real-time to other users, so that they too will see the updated scene. This form of implementation requires a constant connection, either between the AR devices directly or by means of a central server, to coordinate the digital objects displayed and update these as necessary.

2.3.3 Spatial

Where HHDs and HMDs are characterised by the technology positioned in relation to the user, Spatial Augmented Reality (SAR) is characterised by the overlay of digital images over a specific physical area as well as the independence of the system from the user. Indeed, unlike the previous two categories, SAR systems are not worn and tend to be more static solutions. SAR can be achieved either through video, projection, or holographic/optical displays.

Video SAR is the superposition of digital images onto a video feed of the real world. This implementation is one of the most ubiquitous forms of AR currently in use. For example, sporting events often overlay valuable information, such as score, time remaining, and names onto the live feed of a match. Conversely, Cave Automatic Virtual Environment (more commonly known by the recursive acronym CAVE) is a platform that allows for more immersive and interactive setups, such as those found in simulators (Peddie, 2017). Figure 21 shows the former type of Video SAR. Unlike HHD, the scene being recorded is separate from the user and the augmentation takes place on the live video feed before being presented to the user (Figure 21a). This means that from the perspective of the user it appears as if they are interacting with a regular computer screen (Figure 21b), however the view the user sees is a live feed with additional augmentation added to it in real time which is responsive to the changing environment (Figure 21c).



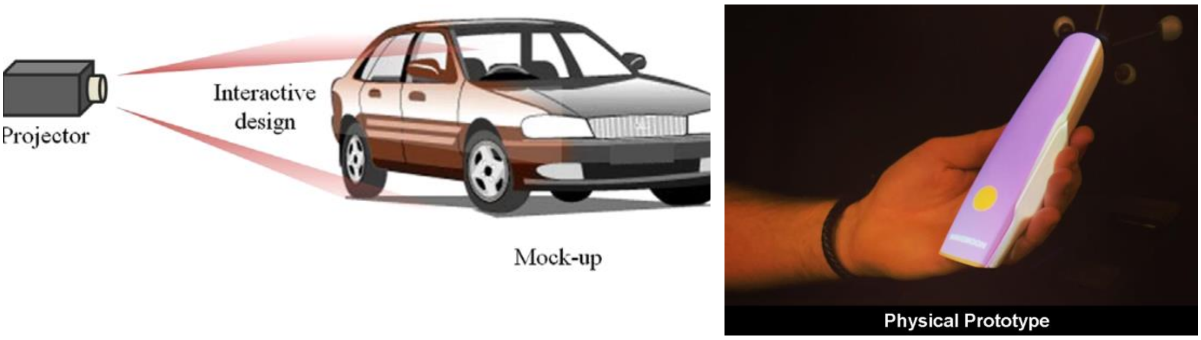
(a) Process of video capture, augmentation, and presentation on screen for user

(b) View of user interacting with the AR visualization of a separate room

(c) User view of the separate room presented on screen

Figure 21. A Video SAR system used to interact and control a robot (Hashimoto *et al.*, 2013)

Projective SAR uses projectors to render images directly onto physical prototypes (Figure 22a). These prototypes are usually matte white, to allow for better colour rendition, and are occasionally tracked to enable users to interact with them. If tracked, the projection will follow the physical prototype around as it is moved. Otherwise, the physical prototype must remain static and only the digital overlay can be interacted with. Figure 22b shows an interactive projective SAR model created by using the SPARK platform. This SAR system enables the projection of additional features onto physical surfaces to support the development of marketing and packaging material (Caruso *et al.*, 2016b).



(a) Standard Projective SAR setup. A projector projects a digital interactive design onto a blank physical mock-up (Park *et al.*, 2015)

(b) Projective SAR in use. The model shown is actually completely white but the projection gives the illusion of colour (SPARK Consortium, 2018a)

Figure 22. Projective SAR

Optical SAR (also known as Pepper's Ghost) relies on projecting digital images onto semi-transparent materials. This partially reflective layer allows users to see objects beyond it as well as any digital images that have been projected. This type of technology has been used in head-up displays and, more recently, been implemented in music performances to allow dead musicians to "play" in live performances next to live artists as shown in Figure 23 (Peddie, 2017).

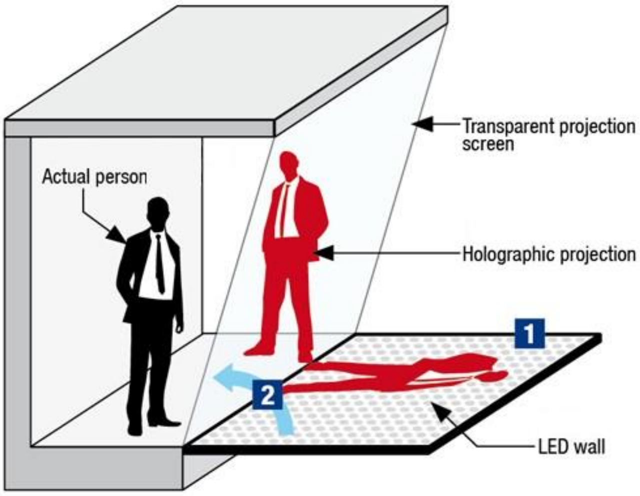


Figure 23. Optical SAR. (1. An image is generated on a LED wall, 2. The image is reflected off a semi-transparent screen to create the illusion that the real and digital person are standing next to each other (Gerriets, 2018)

The development of SAR technology has, to date, led to few commercial applications despite having been the subject of various research projects and studies (Bimber and Raskar, 2005; Bottani and Vignali, 2019). Unlike HMDs, custom hardware is rarely necessary and off-the-shelf components are often used to 'put together' an SAR system. Although no custom hardware is required, SAR typically requires extensive preparation and setup to ensure that it works effectively. This is may be a block in the commercial development of SAR and makes it less attractive to potential consumers (Park et al., 2015).

The flexibility in the type of hardware used, and the ease with which commercial off the shelf systems can be integrated, complicates the definition of SAR. Where other AR technologies can often be grouped together by their features, this is much harder to do with SAR. As such, it becomes much harder to properly perform a systematic analysis of SAR technologies to assess their impact and utility. SAR systems will often be ad hoc and developed for a specific purpose with specific requirements and limitations. At present, this makes it much harder to understand which elements of SAR provide benefits to specific tasks.

2.3.4 Summary

This section has sought to briefly introduce the most prevalent AR technologies and provide a brief overview of each in accordance with the AR taxonomy outlined in Table 3. It is now possible to better understand the reasoning that has led to the selection of projective SAR as the technology to support collaborative design. A more detailed analysis, investigating how AR has been applied in the context of design specifically, is available in section 2.4.

While SAR may be less prevalent than either HHDs or HMDs (Bottani and Vignali, 2019), SAR presents some unique advantages that make it well suited for supporting collaborative design sessions. Table 4 shows each technology's ability to support specific elements required for co-design. Whilst the elements used for this comparison are not exhaustive nor fully representative of all those necessary for co-design to occur, they are each individually necessary and have been discussed as such in the literature.

Firstly, unlike HHDs and HMDs, SAR is less cumbersome for users to operate; participants are not required to continuously hold up a tablet or wear a heavy headset, two issues which have prevented industry from adopting AR technologies due to health and ergonomic concerns (Uva *et al.*, 2018). Secondly, SAR allows all the participants to view the same shared representation at the same time. With the exception of projective HMDs, all other non-Spatial forms of AR require each user to have their own device to view the AR effect; this increases complexity, as the devices must be synchronised, and further isolates the designers from one another as they must focus on their individual viewing screen (Morosi *et al.*, 2018a). Furthermore, the weight of constantly having to hold a screen or headset, as in the case of HMDs or HHDs, can contribute to fatigue, limiting the maximum length of a co-design session and contributing to the exhaustion of the participants (Uva *et al.*, 2018). Projective HMDs can also achieve similar results to SAR systems but force the user wearing the projective HMD to act as a camera stand as well restricting their ability to communicate and collaborate with others. These points are summarised in Table 4.

Table 4. Comparison of AR technologies to Co-Design Support elements. The matrix rates each technology's ability to support elements of Co-Design as either poor (--), middling (+-), or good (++)

			Co-Design Support	
		Cumbersomeness	**Shared Representations**	**Communication**
Technology	**HMD** Retinal	--	--	+-
	Optical	--	--	+-
	Video	--	--	+-
	Projective	--	+-	+-
	HHD Scene Augmentation	--	+-	--
	Object Augmentation	--	+-	--
	SAR Video	+-	+-	--
	Optical	+-	++	+-
	Projective	++	++	++

For these reasons, out of all the available AR technologies it would appear that SAR is most suited to supporting co-design sessions and in particular, Projective SAR (Table 4). Projective SAR does not restrict the viewing angle of the object being augmented or limit interaction. Video SAR and Optical SAR, by the very nature of how the SAR effect is achieved, prevent participants from interacting with the object(s) being augmented, limiting the effectiveness with which participants in co-design sessions can communicate.

Out of the three types of SAR, only projective SAR has seen adoption in the field of design with various prototype systems identified during the course of this review (Porter *et al.*, 2010; Park *et al.*, 2015; Calixte and Leclercq, 2017). Porter et al. (2010) mentioned the ability of Projective SAR to allow improved interaction with physical prototypes being designed when compared to other forms of AR. Park et al. (2015) describe how the decision to pursue an SAR system was based on SAR's low cost, intuitiveness, and flexibility. Lastly, Calixte and Leclercq (2017) discuss how SAR is well suited to supporting collaboration and communication between design session participants, in particular in co-design sessions. For these reasons, projective SAR was selected as the focus for the SPARK project and this research. To aid with clarity, future references to SAR are intended to refer to projective SAR unless otherwise specified. A more detailed analysis, investigating how SAR has been applied in the context of co-design specifically, is available in Section 2.5.

2.4 AUGMENTED REALITY IN DESIGN

This section reviews the application of AR in design. It begins by describing the literature review methodology followed by the presentation and discussion of the results.

2.4.1 Literature Collection Methodology

Firstly, the methodology for classifying the AR technologies is discussed. Thereafter the search criteria and approach for finding the AR technology literature is shown in addition to the keywords used in the search. Lastly, this section concludes by explaining how the AR technologies are to be tabulated and displayed.

AR technologies are a relatively new development: not all are commercially ready. As such, some extrapolation was needed. This is especially true where prototypes or patents are concerned. Particularly, those prototypes developed only for research purposes rather than commercial exploitation. As with any literature review, in particular one dealing with the re-organization of literature to explore new dimensions, decisions needed to be made in how to organise and display the data.

The review identified relevant research papers through the use of Google Scholar. Google Scholar was chosen due to being independent from major publishing houses. In addition, it has the ability to link to commercial material which could enable the identification of technologies not published in academic papers. In addition, Google Scholar provides a feature, Related Articles, that automatically attempts to find similar and related articles to the one selected. These were also queried to collect an initial set of papers.

Search term keywords used were:

- Mixed Reality
- Augmented Reality
- Design
- Assembly
- Creativity

The keywords were selected as follows: Mixed Reality and Augmented Reality were used as keywords to guarantee that the search would contain results pertaining to Augmented Reality. The term Mixed Reality was also used as some papers, older work in particular, use the two terms interchangeably. Design was used as a keyword to attempt to find all works that linked to design and design practice. Assembly was used as a key term to capture those papers that used augmented reality to evaluate ergonomics or assess the impact of AR on basic prototype construction tasks. Creativity was used as a keyword to attempt to capture papers aimed at assessing the impact that AR had on the idea generation, concept development, and concept evolution elements of the design process.

These keywords were searched in the following way: Mixed Reality and Augmented Reality were searched for first independently. Thereafter any search was a combination of Mixed Reality or Augmented reality as well as the additional terms (Design, Assembly, and Creativity). For example, searches were “Mixed Reality Design” or “Augmented Reality Assembly”. Once all combinations had been searched for, and the relevant related articles had been explored, the search was considered complete.

Only the first ten results listed (the full contents of the first page of results) were taken into consideration. Where empty cells were still present a second search was conducted in an attempt to identify any additional papers that may have been overlooked in the original search. This second search looked at the first twenty listed results. Search terms used for this second search revolved around combining the AR technology with the design process phase. The keywords for the second level search were:

- Head-worn
    - Video
    - Retinal
    - Projective
- Handheld
- Spatial
    - Video
    - Optical
    - Projective
- Task
- Design Specification
- Concept
- Preliminary Layout
- Definitive Layout
- Product Documentation

In both cases not all papers that were listed as results by Google Scholar were included in the matrix. This is because some papers might only mention either design or AR but not go into depth or not have it as the focus of the paper. When considering which papers to include, the abstract for each was read. A clear link between design and augmented reality had to be present within the abstract to allow for the paper to be included.

The references collected were then coded in terms of the technology used and the stage in the design process. Understanding where the reference could be used to support the design process was the more complex of the two tasks. In the event that the reference explicitly stated how the technology could be used to support the design process, the categorization provided by the authors was used. However, many of the references identified tended to showcase the technology more and focus less on potential applications. Furthermore, simply because a reference failed to mention a potential design application, choosing instead to focus on other fields, did not mean that it could not be used for design purposes as well. In the interest of clarity and transparency each reference used has been provided with a brief summary to better understand the logic behind the categorization into a specific design process stage in addition to describing the original intended purpose should it not have been originally intended for design. It is important to note that some technologies can be used in multiple stages of the design process.

2.4.2 Results

The AR technologies reviewed are listed in Table 5. The review included commercial applications and showed that they required additional development before they could be considered mature enough to gain wider acceptance. Many of the technologies reviewed were still in a prototype stage or were not designed in such a way that they could be easily used off the shelf by a consumer. Indeed, many technologies required considerable knowledge and skill to assemble, maintain, and deploy in such a way that only someone who was intimately familiar with the design could make the system work in the desired way.

Table 5. Summary of reviewed AR Technology Papers

REFERENCE	AR TECHNOLOGY	SUMMARY
ONG ET AL.(2011)	**HEAD-WORN: RETINAL**	Review of different AR technologies. Analyses technologies for design review, manufacturing, and robot programming. Of particular interest is the design review section which matches the Definitive Layout stage through the use of Retinal HMD.
TAWARA (2011)	**HEAD-WORN: VIDEO**	Development of a system for visualization of CT/MRI scan data allowing users to manipulate 3D models and see cross sections. Also enables manipulation of 3D models in real time. Such a system could be used for either Concept or Preliminary Layout stages.
KAUFMANN ET AL. (2011)	**HEAD-WORN: VIDEO & RETINAL**	Development of a 3D modelling environment for use in teaching. Two technologies are investigated: both a retinal and a video HMD. Use of primitive shapes and basic implementation of geometry as well as simple mechanics suits the Concept stage as it enables a simplification and subsequent analysis of the problem.
JANG ET AL. (2017)	**HEAD-WORN: RETINAL**	Head worn retinal AR system that enables image visualization by means of eye position tracking. Light is projected directly into the eye to reduce ambient light interference. Low image quality and colour resolution limits applications to early concept stage.
WAI ET AL. (2018)	**HEAD-WORN: VIDEO**	Adaptation of an Oculus Rift and a Leap Motion Controller to allow a user to use hand gestures to interact with the digital representations. Video setup means that the real world is captured by a camera and augmented on-the-fly before being streamed to user. Appropriate for Definitive Layout stage.
VALENTINI ET AL. (2018)	**HEAD-WORN: VIDEO**	HMD video system that combines a force feedback device to allow the user to modify the scene being viewed. The example provided by the authors is the modification of a surface representing a car bonnet. This makes the technology ideal for the Concept and Preliminary Layout stages

Kemmoku et al. (2016)	**Head-worn: Projection**	An attempt at creating an interactive desktop interface by means of a head mounted projector. Hand tracking allows the interface to be interactive. The technology lends itself well to the Product Documentation stage. The example provided in the paper shows the system in use to review pictures and text.
Ma et al.(2011)	**Head-worn: Video & Retinal**	Overview of prototypes of AR technologies currently in use or under development in industry. Use scenarios include AR as an aid in assembly of components and as a prototyping tool to explore layouts.
Poh et al. (2005)	**Head-worn: Video**	Discussion of a potential architecture for the development of an AR based CAD system. Use of markers for drawing and measuring in digital space. Preliminary Layout stage.
Valentini (2009)	**Head-worn: Video**	Manipulation and interactive assembly of virtual objects through the use of AR headset and gloves to detect hand movement. Virtual assembly of basic parts and simple manipulation of objects indicate function at the Preliminary Layout Stage.
Park (2008)	**Head-worn: Video**	CAD visualization system for the setup and assembly of models. Smaller models can be manipulated and modified as well as reassembled. Preliminary Layout Stage.
Radkowski et al. (2009)	**Head-worn: Video**	A setup for realistic lighting and colour rendering for CAD models in AR. Due to the fine detail finishing this can be categorised under Definitive Layout.
(Reyes et al., 2020)	**Head-worn: Video**	A device to support the assembly is presented in the paper. The paper utilises the AR technology to support the assembly of motherboards. The study analysed how the AR system supported participants in their assembly task, finding their ability to properly perform it improved. The system discussed in the paper is best suited to a Product documentation stage.
Hua et al. (2011)	**Head-worn: Projective**	A review of projective HMDs showcasing different technologies available for development as well as potential applications. Majority of applications suggested are for visualization or basic manipulation of objects suggesting suitability for a Preliminary Layout stage.
(Potts *et al.*, 2019)	**Head-worn: Optical**	The paper presents a tool, ZenG, aimed at fostering creativity by combining AR with an electroencephalography tool to allow the system to modify the visual output on the AR display not just based on user inputs through controllers but also brain activity. The system is best suited to a Concept generation stage due to the creativity support the tool is designed to provide.
(Eder *et al.*, 2021)	**Head-worn: Optical**	Head worn device to assist assembly tasks on production lines. The system is designed to provide the user with usable and intelligible feedback and instructions on how to perform assembly operations. This makes the system best suited to a Product documentation stage.
(Carrasco and Chen, 2021)	**Head-worn: Optical**	A study to analyse the impact of AR on supporting design review of an architectural design. The study indicates that the AR setup provides additional support when compared to the more traditional 2D media. The nature of the AR setup, aimed at design review, makes the system most suited for a Preliminary or Definitive Layout stage.

(Chalhoub and Ayer, 2018)	**HEAD-WORN: OPTICAL**	The paper proposes and discusses the value of AR as a tool to support building information modelling. The AR system presented is contrasted to the existing use of 2D drawings to convey information about electrical layouts within buildings finding that the use of AR simplified the task. The tool presented in the paper appears thus best suited for a Product documentation Stage.
XIN ET AL. (2008)	**HANDHELD**	3D sketching interface for portable tablet PCs. Markers are used to define a sketch space which is then drawn in through the tablet interface by the user. The basic designs being generated, and the creation of working principles suggests that this application would be best suited for the Concept or Preliminary Layout stages depending on the specific use made by the designer.
LIESTØL (2011)	**HANDHELD**	Analysis of situated simulations as a method for displaying architectural buildings. Used to simulate and visualise ancient structures in their historical locations where otherwise only ruins can be seen. Low modifiability with high degree of accuracy in representation makes this application best suited to Definitive Layout stage.
STUTZMAN ET AL. (2009)	**HANDHELD**	Development of MARTI (Mobile Augmented Reality Tool for Industry) platform developed specifically for mass market use in industrial settings. MARTI is designed to aid in assembly and setup of machinery as well as other fine-tuning aspects as the user input is less required and the system is more geared towards giving instructions. This makes the system best suited for Definitive Layout as well as Product Documentation.
ZHANG ET AL. (2010)	**HANDHELD**	Virtual panel for the visualization of CNC machining pathways. The system aims to assist users to follow specific pathways for machining. This makes it most suited for the Definitive Layout stage.
(BRUNO ET AL., 2019)	**HANDHELD**	The paper explores the use of AR within the context of industry 4.0. The technology presented therein is aimed at supporting information exchange between professionals involved in the design and production processes for the oil and gas sector, making the system adequate for both the Concept and Preliminary Layout stages.
(KERR AND LAWSON, 2020)	**HANDHELD**	Discusses the development of an AR prototype to act as an art installation that would aid members of the public, such as students and non-designers, to better understand the foundations of landscape architecture. The tool is used as a teaching and digital storytelling tool and is thus adequate to support both the Definitive Layout and Product documentation stages.
SMPAROUNIS ET AL. (2007)	**SPATIAL: VIDEO**	Virtual and Augmented Reality tool for collaborative design. The AR part consists of both an online and offline interface that permits the visualization of a room wherein designers can make changes collaboratively to the layout by moving set pieces around. Preliminary Layout stage.

Marner et al. (2011)	**Spatial: Projection**	Discussion of how SAR can be applied to Design Specification, Concept, Preliminary Layout and Definitive Layout stages of the design process. Various techniques for SAR and potential applications, both present and future are described here including how SAR can be implemented without having to rely on a pre-existing physical model.
Marner et al. (2009)	**Spatial: Projection**	SAR applied to user interface and finishing design. Due to the nature of SAR a semi-final physical model for projection is required. As such this is best suited for Preliminary and Definitive Layout stages.
Löchtefeld et al. (2011)	**Head-worn: Projection** **Spatial: Projection**	Investigation of how pico-projectors may be used to develop new applications for SAR and head-worn projection systems. A number of potential applications, from games to map design and augmentation are suggested as potential avenues for exploitation of the technology. Due to the relative simplistic nature of the augmentation and the ability to present sketches and render them interactively this seems best suited for the Concept stage.
Marner et al. (2010)	**Spatial: Projection**	Presentation of a new technique for real-time simultaneous modelling of both physical and digital worlds. Use of foam physical prototypes with projection that can be cut and modified as guided by the projection for additional flexibility. This supports the Concept stage in the design process.
Israel et al. (2009)	**Spatial: Projection**	An investigation of how CAVE can be used to combine 2D and 3D sketching for designers in the Concept stage.
Calife et al. (2009)	**Spatial: Projection**	Discussion on the creation of Robot Arena, an AR platform to aid in the development of games in combination with existing physical robot models. Useful for Preliminary stage design as it allows setup and organization of initial concepts.
Irlitti et al. (2013)	**Spatial: Projection**	Discussion of how new techniques can be used for constraint driven design using SAR resulting in the creation of a prototype, named SARventor. This approach to SAR best suits the Definitive Layout stage.
Hashimoto et al. (2013)	**Spatial: Video**	Development of an AR system to remotely control a robot. The user views a video feed of a room with a robot that has a digital overlay that allows interaction with the robot. Of interest to support the Preliminary and Definitive Layout stages
(Dalinger et al., 2020)	**Spatial: Video**	Presents a tool, termed Murison, which is designed to aid in the training of teachers. The system simulates a virtual classroom to allow for a realistic teaching environment to be simulated. The study discusses the feedback received from users of the system when compared to in person observations of classrooms, finding that the tool has a positive impact on the teachers in training. The system appears to be best suited to a Product Documentation stage

(YAMAGUCHI ET AL., 2020)	**SPATIAL: VIDEO**	The paper discusses a tool designed to aid users visualise and create tutorials for the assembly of objects. The system relies on a set of cameras to track user input during the assembly process and visualises the next step in the process, highlighting and explaining any complex procedures. This makes the tool most to the Product documentation stage.
CARUSO ET AL. (2016B)	**SPATIAL: PROJECTION**	Projective system designed to support product and packaging design in collaborative environments. Aimed predominantly at the Preliminary and Definitive Layout stages due to the need for a physical object that needs to be agreed upon for projection.

It should be noted that the results listed in Table 5 are based on the review by Giunta et al. (2018) but has been expanded from the original 21 papers to now analyse 37. The methodology followed was unchanged during this expansion and is the methodology reported on in section 2.4.1. The review by Giunta et al. (2018) highlights a number of potential avenues that AR technologies, and SAR technologies in particular, could pursue in order to support design activities. The analysis also highlights that SAR is best suited for supporting design activities that are further along the design process and that earlier stage design activities are better supported by other forms of AR.

2.4.3 Analysis

As previously mentioned in section 2.1.2, there is reason to believe that co-design can improve design session output and thus value for end-users and clients (Prahalad and Ramaswamy, 2004; Payne *et al.*, 2008). The question now becomes: what role can AR, and more specifically SAR, play in supporting co-design sessions? SAR certainly seems to have the potential to support co-design sessions, and this has attracted increasing interest from both the scientific community as well as from industry (Verlinden *et al.*, 2010). One of the current hypotheses purports that SAR, when applied to design, enables participants to communicate more efficiently, thereby having their explanations, ideas, and opinions more readily understood by their interlocutor, thus resulting in improved outputs (Bordegoni *et al.*, 2009; Calixte and Leclercq, 2017). The design output from SAR sessions has, in part, been investigated by O'Hare et al. (2018a) as part of the SPARK project.

Their study focused on the impact that SAR and AR technologies have on the novelty and quality of ideas generated in co-design sessions by comparing them to traditional design sessions. Traditional design sessions are sessions that are conducted without the use of AR technology to support them, and often rely on paper models, sketches, and computer screens. Their study showed that the use of SAR increased idea novelty and quality but also pointed out challenges in using and implementing SAR consistently. This is in line with the findings of Akaoka et al. (2010) as well as Calixte and Leclercq (2017). Both these papers noted the usefulness of SAR during the design session. Akaoka et al. (2010) noting participants enjoyed using SAR. Calixte and Leclercq (2017) found the platform "support[ed] encouraging the understanding of complex shapes" but found, like O'Hare et al. (2018a), that the setup and calibration was particularly complex.

Other aspects of SAR have also been investigated, such as its ability to supplement physical feedback. Porter et al. (2010) analysed the potential of SAR systems as a User Interface design tool. The study tracked the finger movement of participants to simulate button presses. The study showed that mean "button-press time" (i.e., the time taken to press all the buttons indicated in the task) was 1.2

times higher when using the SAR prototype when compared to a physical prototype. Again, participants in the study mentioned the limitations of the technology, in particular the lack of tactile feedback. Nonetheless, the participants felt confident in using SAR as a design tool. This finding is supported by Park and Moon (2013) who investigated the impact of AR on design evaluation studies. Their research found that the main difficulties faced during the implementation of AR relate primarily to: hand occlusions, difficulty in interacting with the prototypes, and discomfort of headsets due to their weight.

2.5 SAR IN CO-DESIGN

Section 2.1 analysed the benefits of co-design to establish whether supporting co-design was a worthwhile goal. Section 2.4 looked at how AR can be used to support design in general. Now, from looking at the potential of AR in design, this section focuses more specifically on the potential of SAR in co-design and in eliciting the requirements for its adoption. It should be reiterated that the work presented within this section was part of the work performed by the SPARK Consortium for the SPARK Project prior to the start of this PhD, as noted in section 1.3.1.

SAR differentiates itself as an augmenting technology by applying environmental projection onto physical scenes thus not requiring any attachment to the user. This has shown to be advantageous when compared to Virtual Reality, which requires headsets or screens, as it has been proven to be more comfortable for long design sessions, provides a greater field of view, and supports collaboration to a greater extent (Caruso *et al.*, 2015). This is of particular interest for work taking place later in the design process, where modifications are less centred around physical shape (Ong *et al.*, 2011). The technology is of particular interest for the packaging and advertising industry as well as any design whose colour, material, and finish is being evaluated (Caruso *et al.*, 2016b). It should be noted that SAR has been used in multiple studies with a wide range of applications in design, such as the prototyping of both interfaces and products (Porter *et al.*, 2010; Park *et al.*, 2015; Dey *et al.*, 2018; Morosi *et al.*, 2018a).

2.5.1 SAR Format for Co-Design

Billinghurst and Kato (2002) and Ben Rajeb and Leclercq (2013) suggest that SAR would improve collaboration from a theoretical perspective. Billinghurst and Kato (2002) state that the real world (which lies at one of the extremes of the reality-virtuality continuum (Milgram *et al.*, 1995)) plays a large role in communication and collaboration, in particular with design and spatial collaboration. They posit that AR systems capable of combining tangible interaction methods with the richness of digital interfaces and overlays would greatly improve users' shared understanding. Ben Rajeb and Leclercq (2013) support this claim by stating that the use of SAR allowed participants in collaborative design sessions to communicate more effectively as the causal link between what is being said and the actions undertaken is maintained. Thus, the SAR enhanced model acts as a boundary object that all the participants can relate to and use to express their thoughts, with changes made by each participant visible in real time, enabling easier communication of intent.

One of the hypotheses surrounding the application of SAR in design is that it will enable participants of a design session to communicate more effectively resulting in improvements in the number and quality of design outputs. The number, novelty, and quality of design outputs was investigated by O'Hare et al. (2018a) who studied how SAR and AR technologies influence ideas in collaborative design. It was revealed that SAR increased the novelty and quality of ideas but extended session setup time due to the need to calibrate the system. This is corroborated by Akaoka, Ginn and Vertegaal (2010) whose study found that participants enjoyed the interaction environment it provided. The participants highlighted that SAR requires configuration to support the intended exercise and a single off-the-shelf implementation is not suitable for fully supporting design activities. The interaction afforded by SAR has also been investigated by Porter et al. (2010). They evaluated the potential of SAR for User Interface design where finger tracking could be employed to indicate button presses. Their study revealed the mean "button-press time" increased by a factor of 1.2 for the participants using the SAR prototype when compared to the participants that made use of a traditional physical prototype.

However, the speed of design iteration was greatly reduced thereby enabling more design iteration in a given timeframe. Participants also reported technical limitations and, in particular, the lack of tactile feedback in the SAR models. Participants nonetheless felt confident that SAR would be useful as a design tool. This is corroborated by Park and Moon (2013) who applied AR to design evaluation studies, which found that the primary hurdles faced in implementing AR related to: hand occlusions, interaction difficulties with the prototypes, and the weight and inherent discomfort of headsets. Giunta et al.'s (2018) review of AR in design research highlighted that it has centred around the Concept, Preliminary Layout, and Definitive Layout stages, with gaps in the application of AR in early and late stages of the design process. Overall, the paper highlighted that the technology shows potential and that some commercial applications have been tried but that it still requires additional development before it can be considered mature enough. Supporting the development of AR by understanding how it affects design can lead to a better understanding of how it can be made ready for more widespread adoption.

What is less well understood is the form and format that SAR could take to support co-design sessions. Some research has been conducted to analyse how SAR could support Industry 4.0 (Uva *et al.*, 2018; Butt, 2020; Masood and Egger, 2020). These studies have found that AR in general and SAR in particular can support activities such as assembly and manufacturing. However, less research has been conducted into how SAR can be used to support design itself. One explanation for this may be that, due to the technology still being in its infancy, more research has focused on the development of tools rather than analysing the impact of the tools themselves: in particular in more comprehensive scenarios. One interesting finding however comes from Masood and Egger (2020) who note that while the adoption of AR may be limited by the technological aspects, much less focus has been given to organizational issues. These appear to play a larger role in determining whether AR technologies are ultimately adopted, but these organizational issues have, till now, been more neglected.

However, assuming that organizations are willing and able to adopt SAR into their work process, there is still an insufficient amount of information and knowledge regarding the format that SAR should take to support co-design sessions. As SAR is an emerging technology, most of its implementations have been prototypes, often tailored to a specific need. This has led to difficulties both in analysing the impact of SAR on design sessions and in understanding what format SAR must take to be able to support these sessions (Morosi *et al.*, 2018a).

2.5.2 Establishing Requirements for an SAR Co-Design System (SPARK)

To address the issue of the form and format that SAR should take to better support co-design, the SPARK project benchmarked conventional co-design practice. In addition, interviews were conducted to establish the requirements for a co-design SAR system. This took place prior to this PhD (as highlighted in Figure 4) and is therefore treated as part of the literature review.

The SPARK project's aim was to develop a "Responsive ICT platform that exploits the potential of Spatial Augmented Reality for supporting and fostering collaborative creative thinking"(SPARK Consortium, 2015) therefore the requirements for the development of the platform needed to be laid out. Two industry partners, Stimulo and Artefice Group, helped provide additional insight into the industry requirements that might otherwise not have been explored. While both partners are design consultancies, Stimulo specialises in product and interface design whereas Artefice Group focuses more on packaging design and branding.

Part of this groundwork was an investigation into the needs of designers in order to identify the specific areas where the use of AR could provide added value, particularly within the context of co-design sessions. In doing so it was possible to develop an understanding of the requirements for an SAR platform aimed at supporting co-design. Furthermore, the investigation was able to identify additional opportunities for the application of SAR in co-design.

The investigation was split into three studies. The first two studies, one for each industry partner, focused on the analysis of conventional co-design sessions. The methodology for the analysis between the two design consultancies was unvaried. The third study consisted of interviews with the participants of the first two studies as well as interviews with external organizations identified as potentially interested in the technology. These were predominantly other design consultancies and consumer goods manufacturers (O'Hare *et al.*, 2016b).

2.5.2.1 Gesture and Verbal Analysis

In the first two studies the co-design sessions were allowed to proceed as they normally would have when run by the respective design agencies and no changes to their structure, layout, or timing was made other than to have a number of high-resolution cameras in the room recording the proceedings. The recordings were then coded based on the gestures and interactions of the participants (O'Hare *et al.*, 2016b). This data was collected in order to explore co-design sessions to aid in the development of a SAR platform.

The coding scheme is presented in Figure 24 and further elaborated on in Table 6. Figure 24 shows how the various gestures were coded between the client, designer, and artefact. These are colour coded green, blue, and orange, respectively. Each type of interaction is given a number that can be looked up in Table 6 to understand the types of interaction. Table 6 shows who the initiator of the interaction is in column "Level 1.1". Column "Level 1.2" relates the definitions of the interactions to the interactions shown in Figure 24.The gesture interaction framework discussed in Figure 24 and Table 6 is consistent with the framework first developed by O'Hare *et al.* (2016a). This framework enabled the analysis of the interactions between the design session participants, namely: the designers leading the collaborative design sessions; the clients participating in the sessions; and any artefacts present in the session used as tools to support design creativity (O'Hare *et al.*, 2016b).

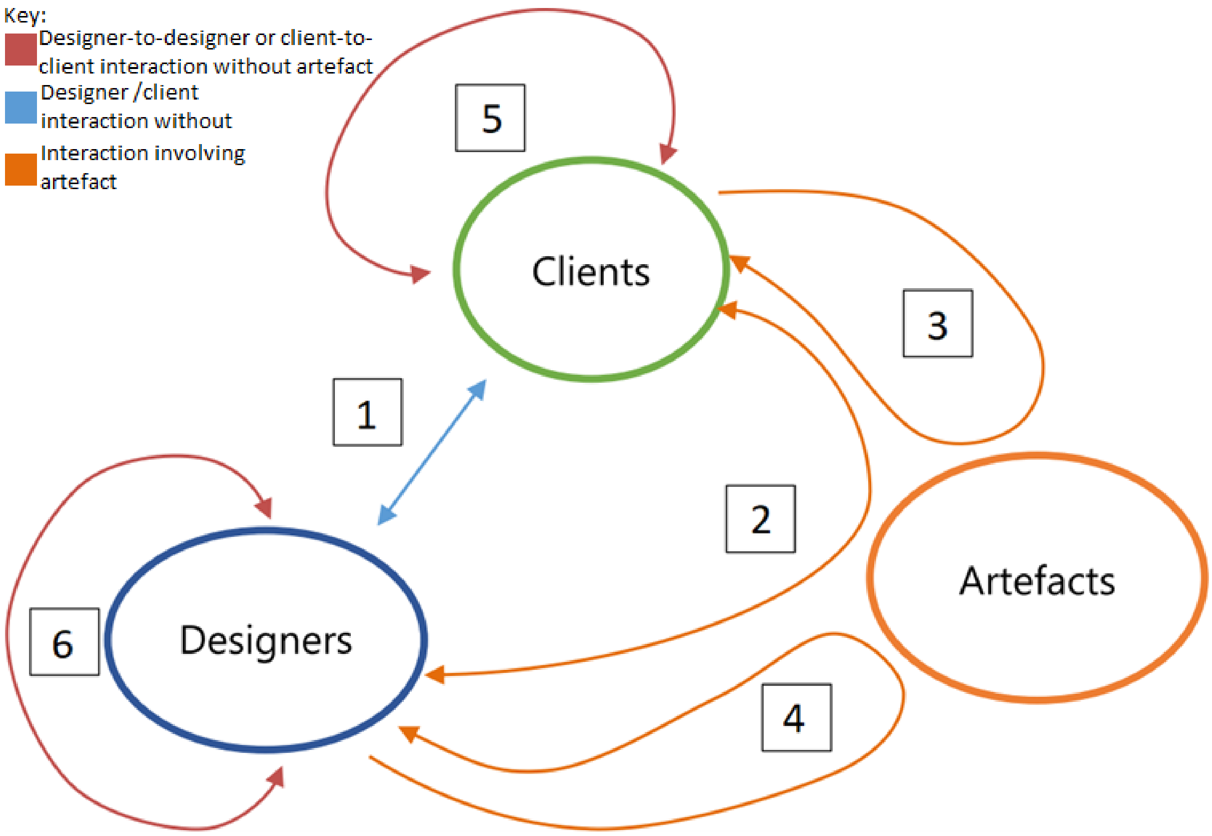


Figure 24. Structure of the Gesture Analysis Framework (O'Hare et al., 2016b). The interaction behaviour between designers and clients is described in Table 6.

Table 6. Gesture Analysis Framework Interactions and Definitions (O'Hare et al., 2016b)

LEVEL 1.1	LEVEL 1.2	INTERACTION	DEFINITION
C	1a	Interaction from the client to the designers, without artefact	The client will explain/show something to the designer without using an artefact
D	1b	Interaction from the designers to the client, without artefact	The designer will explain/show something to the client without using an artefact
C	2a	Interaction from the client to the designers, through an artefact	The client will explain/show something to the designer by using the artefact
D	2b	Interaction from the designers to the client, through an artefact	The designer will explain/show something to the client by using the artefact
C	3	Interaction of the client with an artefact	The client will use/manipulate the artefact for himself
D	4	Interaction of the designers with an artefact	The designer will use/manipulate the artefact for himself
C	5	Interaction between the clients, without artefact	The clients will explain/show something/talk together without using an artefact
D	6	Interaction between the designers, without artefact	The designers will explain/show something/talk together without using an artefact

Verbal interactions were also coded (Table 7). The coding scheme was divided into three layers of analysis. Table 7 shows and provides additional details on each layer of analysis performed. Additional information concerning the exact categories, statements, and key phrases searched for as part of the verbal analysis can be found in “Results from the Experimental Activities and Presentation of the Research Metrics Framework” (O’Hare *et al.*, 2016b) and “Case Studies and Evaluation Criteria”(O’Hare *et al.*, 2016a) where these metrics were first identified and discussed.

Table 7. Three Layers of Analysis Performed to Analyse the Verbal Interaction (O’Hare et al., 2016b)

LAYER 1	Relevance of the content and interaction for the development of the SPARK platform (5 mutually exclusive categories on the nature of design items with reference to their applicability to the mixed prototype “to-be”)
LAYER 2	SAR-related topic emerging from the discussion (8 different mutually exclusive categories describing what the items are)
LAYER 3	Distinctive features from the topics considered and coded at layer 2, with reference to the characteristics the designer would like to change or keep (9 different features describing the features to be changed or kept in the proposed design)

The first study analysed two co-design sessions provided by Artefice, a design consultancy that focuses predominately on packaging design. The first co-design session focused on the packaging design for a range of biscuits. The second co-design session concerned itself with the brand identity of an ice-cream brand (O’Hare *et al.*, 2016b).

The second study looked at the co-design activities of Stimulo, who focus primarily on product design. Again, two co-design sessions were analysed, one focusing on the redesign of a personal locator beacon and the second on the design of a gas barbecue (O’Hare *et al.*, 2016b).

The findings from these two studies highlighted the importance of boundary objects in communication between co-design session participants as circa 90% of interaction time involved using some form of design representation or artefact (O’Hare *et al.*, 2016b). Additionally, the study found that clients tended to use tangible models more often than digital representations, though both were used during sessions. Furthermore, participants relied on gesturing using their hands when a suitable representation was not available to express their views. Lastly, product design sessions mostly discussed size changes, changes to number of instances of items (e.g. number of logos, buttons, or other features) and changes in shape. Conversely packaging design sessions focused more on changes to colour, look, position, and number of instances of items (O’Hare *et al.*, 2016b).

2.5.2.2 Structured Interviews

The third study consisted of interviews with experienced design practitioners well versed in co-design sessions. The study was split into two: interviews with participants from the previous two studies and interviews with organizations external to the SPARK project (O’Hare *et al.*, 2016b). In both cases the interview took the form of structured interviews, with directed questions by the interviewer aimed at answering the underlying questions posed by the study. All interviews were recorded.

In the case of the interviews conducted with those who had participated in the previous two studies, the aim of the interviews was to answer the research aims, questions, and objectives stated in Table 8. A complete list including the exact interview protocol, as well as the exact responses from the participants can be found in "Results from the Experimental Activities and Presentation of the Research Metrics Framework" (O'Hare *et al.*, 2016b).

Table 8. Research Aims, Questions, and Objectives for the Interviews with Participants of the Previous Co-Design Studies (O'Hare et al., 2016b)

RESEARCH AIM	Explore the impact of design representations in the observation sessions as perceived by designers in order to understand their best practices
RESEARCH QUESTION	What was the impact of design representations in the observation sessions?
RESEARCH OBJECTIVES	RO1: Establish which design representations were reported as used during the observation sessions RO2: Establish how design representations were reported to be used during the observation sessions RO3: Establish what affordances were perceived to be associated with design representations during the observation sessions RO4: Establish what challenges were perceived to be associated with design representations during the observation sessions

The findings of the interviews conducted with the participants of the first two studies indicate that the design representations used during the design process tend to be accurate representations of design concepts and that this enables the participants to explore design elements. Furthermore the interviews confirmed that design representations, when acting as boundary objects, facilitate collaboration between participants (O'Hare *et al.*, 2016b). This is in line with the literature on boundary objects and their significance in the design process (Star and Griesemer, 1989; Carlile, 2002; Bergman *et al.*, 2007).

In a similar fashion to the interviews conducted with those who participated in the first two studies, the interviews conducted with those who were external to the SPARK project were aimed at attaining a set of specific aims. These are summarised in Table 9. Again, a fuller list including the exact interview protocol, as well as the exact responses from the participants can be found in "Results from the Experimental Activities and Presentation of the Research Metrics Framework" (O'Hare *et al.*, 2016b) as well as "Spatial Augmented Reality for Design Representations in Product Design and Development" (Sola, 2016).

Table 9. Aims for the Interviews with Participants from Design Agencies External to the SPARK Project (O'Hare et al., 2016b)

1	To understand the types of co-creative session that currently take place across the New Product Development process.
2	To understand the types of design representation currently used within these co- creative sessions.
3	To understand the challenges that practitioners face with their current use of design representations
4	To gather opinions from practitioners as to how they might use the SPARK platform and what their requirements would be.

The interviews conducted with the external organizations resulted in a better understanding of the basic characteristics of co-design sessions. Sola (2016) discusses some of the results extracted from the interviews, which are summarised in Figure 25. The interviews conducted indicate that the main barriers to effective implementation of co-design sessions are:

#4 Misinterpretation of design representations by participants who fail to grasp the intended meaning/purpose of the design representation
#5 Lack of realistic design representations to evaluate end-user/client interaction due to the approximate nature of mock-ups and prototypes (Sola, 2016).

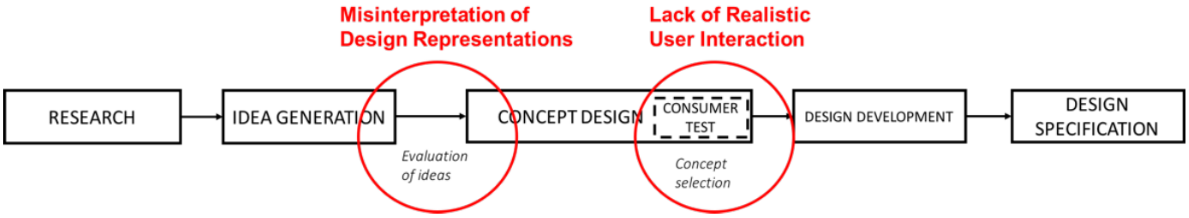


Figure 25. Limitations encountered by designers (Sola, 2016)

The interviews reveal visualization is pivotal to design companies with the ability to accurately portray colours, materials and textures playing a large role in their willingness to adopt SAR (Sola, 2016). However the importance of accurate image rendering decreases at the initial stages of the design process where the design is usually rougher and less detail oriented (O'Hare *et al.*, 2016b). Additionally, the viewing angle is an important element influencing their interest in SAR. The ability to track the movement of the physical prototype, thereby allowing projected images to follow as it is moved, was less requested (O'Hare *et al.*, 2016b). One request that was made by the interviewees related to the portability of the system and its ease of deployment and use (Sola, 2016). Table 10 summarises the list of requirements made by the interview participants ordered by level of importance.

The remaining technical requirements, in order of ranking, focused on:

- The ease of setup and use of an SAR platform
- The resolution of the projection itself
- The requirements of the room where the SAR system was to be setup (e.g.: size, external light, etc.)

- The cost of the SAR system
- The noise emitted by the system whilst in operation.

Table 10. Ranking of Importance of SAR Technical Requirements (Sola, 2016)

RANKING	TECHNICAL REQUIREMENT
1	Accurate rendering of materials, colours, and finishes
2	Visibility of model from various vantage points
3	Projection response speed in relation to movement
4	Ease of set-up and use
5	Projection resolution
6	Room requirements
7	System cost
8	System noise

2.6 DEVELOPING DESIGN RESEARCH METRICS

In addition to the analysis of the literature surrounding AR, SAR, and Co-Design it is important to review and understand the metrics used to evaluate co-design sessions. Some metrics have already been mentioned in the context of limited studies, in section 2.5.2.1. However, a set of robust metrics to be utilised throughout the remainder of this thesis needs to be identified, or developed, in order to provide the opportunity to compare the results of SAR and non-SAR co-design sessions. It should be reiterated that the work presented within this section was part of the work performed by the SPARK Consortium for the SPARK Project prior to the start of this PhD, as noted in section 1.3.1.

2.6.1 Introduction

Successfully evaluating design session outcomes requires a robust set of metrics that can be applied in a repeatable manner thereby enabling design session comparisons. It was for this reason that the SPARK project set out to create a set of co-design metrics building on the work of Shah *et al.* (2000, 2003) and Mombeshora *et al.* (2017), and have been subsequently used in this thesis. The motivation behind their development as well as the relevant theory that underpins the metrics is now discussed.

Prior to the beginning of this PhD, Mombeshora et al. (2017) and O'Hare et al. (2016a, 2016b) had begun to develop a set of metrics to be used as part of the SPARK project that used the work of Shah, Vargas-Hernandez and Smith (2003) as a basis. The metrics developed by them, while perfectly functional, were complex to implement due to their reliance on transcripts of the co-design sessions.

Thus, the objective was to develop a set of metrics that could be implemented more efficiently. The development of the metrics took and iterative approach with four versions being generated over the course of the SPARK project.

Figure 26 shows the breakdown of the development of the metrics over time, linking them to the relevant works that inspired the improvements to the metrics. Each version of the metrics was developed in response to studies conducted that highlighted potential flaws in the methodology. The papers highlighted in the figure analysed the previous version of the metrics and suggested improvements.

As can be seen from Figure 26, the main development of the metrics consisted of five iterations. The initial metrics developed as part of the SPARK Project bid, the two subsequent refinements, and the two versions of the rapid implementation metrics. These metrics evolved chronologically as a result of learnings through the implementation of the metrics in studies. Highlighted in red are the published works that aided in the development of the metrics. The arrows in black denote changes to the metrics that arose through learnings as the metrics were used (Metrics v1 to Metrics v2) or as result of a deliberate change to achieve faster and simpler metric implementation (Metrics v2.1 to Metrics v3).

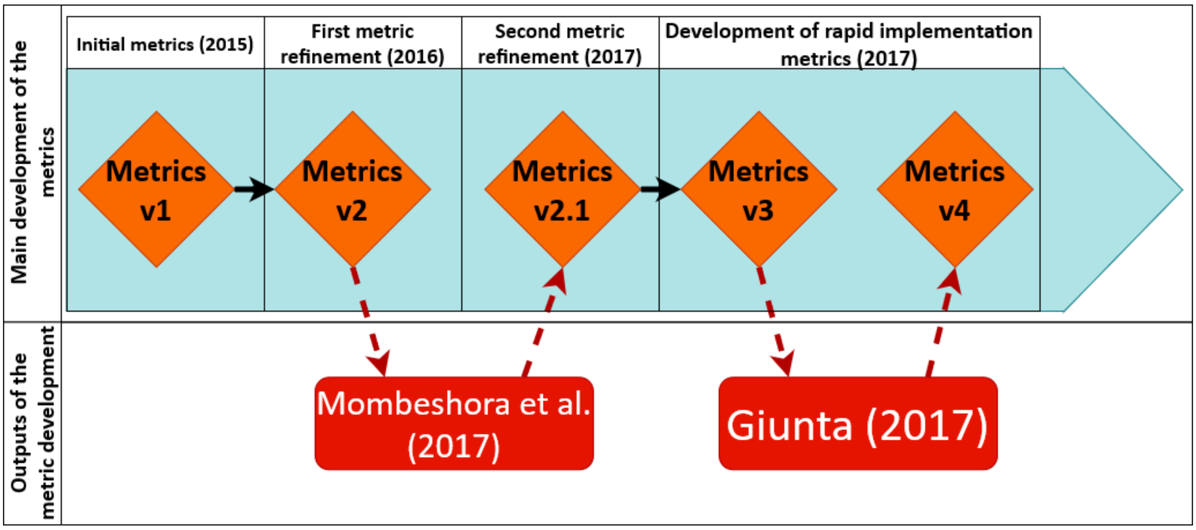


Figure 26. Chronology showing the development of the metrics. Adapted from Ben-Guefreche et al. (2017). The arrow indicates the progressive development of the metrics with the boxes in red highlighting the published work that drove some of the changes

2.6.2 Design Research: In-vivo vs In-vitro studies

In vitro studies are studies conducted “within glass” that is to say in petri-dishes or test tubes in (Oxford University Press, 2018a). This can be contrasted to the term in vivo, which are studies conducted on living beings (Oxford University Press, 2018b). While these definitions are perfectly fine for describing laboratory experiments and clinical trials, when speaking of design studies it is important to note that in vivo and in vitro take on slightly different meanings. In vitro is usually used to refer to a study that is “artificial” i.e. that the design problem being worked on is removed from reality in some way. This can be because, amongst other potential reasons, the end-users participating are actually students or because the design problem is created artificially and with little real-world context. Conversely, in vivo studies focus on analysing design sessions “in the wild” that is to say by observing the participants without removing them from their normal environment where they typically conduct their work and without providing them with cases. Rather letting participants work on the design problems they normally face as part of their regular duties.

Protocol analysis is a strategy, originally developed within the psychology community, commonly used in design session analysis to evaluate design sessions and their outputs (Akin and Lin, 1995). It relies on recording the work of participants and then categorizing it into separate sections, allowing conclusions to be drawn on the length of time, or number of times, that specific elements are brought up (Gero and Mc Neill, 1998). While there are methodologies that have attempted to codify the approach to take in applying protocol analysis, there is still a considerable discrepancy in how this is applied (Shah *et al.*, 2000). For example, one very common method for data collection is the "think aloud" method. Here participants are asked to speak their thoughts out aloud so that audio recordings and transcriptions of their thought process can be made (and eventually transitions and discussion points identified). However, some have disagreed with this method as it may impede the designers' ability to think properly and can become distracting (Akin and Lin, 1995).

Chai and Xiao (2012), in an analysis of papers published from 1996 to 2010, identify protocol analysis as the most published approach to design studies. Their findings do, however, also confirm those found by Shah, Kulkarni and Vargas-Hernandez (2000) - that a more robust approach to design research is necessary. They provide a set of four metrics for the purpose of more empirical analysis of design sessions: Quantity, Quality, Novelty and Variety (Shah *et al.*, 2000). These, unlike results obtained through protocol analysis, focus on design session outputs to help eliminate the qualitative aspect of categorization. The metrics measure the following properties:

- Quantity: The number of ideas generated
- Quality: How well the ideas meet any design constraints or requirements
- Novelty: How unexpected the ideas generated are
- Variety: The breadth and depth of the solution space explored

2.6.3 Design Study Metrics

Expanding on their previous findings Shah, Smith, and Vargas-Hernandez (2003) provide a more thorough methodology for implementing the metrics previously proposed by them. In their paper they explain the implementation of each of the metrics mentioned previously. Furthermore, the paper expands on how each metric can be calculated and therefore how to effectively assess a design session. Their work is, however, not without criticism. López-Mesa and Vidal (2006) suggest improvements to the way the Novelty metric is implemented by attempting to suggest a new approach to reduce bias. Nelson et al. (2009) suggest combining the Variety and Novelty metrics to improve the functionality of the Variety metric. Furthermore, they attempt to change the implementation of the methodology. Contrary to Shah, Smith, and Vargas-Hernandez (2003), Nelson et al. (2009) propose that the metrics should only be implemented at the idea level rather than focusing on the features, which the ideas are made of. Srivathsavai et al. (2010) have argued the exact opposite, suggesting that the metrics only be applied at the feature level rather than the ideas as a whole.

Snider, Dekoninck and Culley (2012) looked specifically at developing methodologies and coding schemes for small scale design studies. They promote the use of a deductive approach to the building of a coding scheme. Their argument is that using pre-existing theory to underpin the coding scheme is likely to result in a high chance of producing useful results.

The set of metrics developed by Shah, Smith, and Vargas-Hernandez (2003) explicitly calls for an external observer to analyse the design session. It is to support their quantitative analysis that their method calls for the use of transcripts. Golafshani (2003) citing Denzin and Lincoln (1998) states that: "[researchers must] emphasise the measurement and analysis of causal relationships between variables”. It is in pursuit of identifying these causal relationships that the use of transcripts is recommended.

Nonetheless, it is also true that not all research can be quantified. As such it is also important to consider qualitative approaches. Golafshani (2003) quoting Patton (2001) defines qualitative research as “research [that] uses a naturalistic approach that seeks to understand the phenomena in context-specific settings such as ‘real world settings [where] the researcher does not attempt to manipulate the phenomenon at interest.’”. This approach seems more appropriate for the research being conducted in the SPARK project. This is because, while quantification is important for comparisons, knowing how the introduction and/or changes to a SAR-enabled co-design session is having a positive or negative impact is also beneficial. The use of an observer aids in bridging the gap between quantitative and qualitative analysis. Thus, a framework is needed that can provide structure, thereby improving the quantitative element by aiding categorization. Nonetheless, the use of an observer to code the sessions based on their interpretation of events, rather than on specific predefined cues (e.g. hand gestures or key phrases) enables a modicum of qualitative analysis to be injected into the process. This approach proves beneficial by allowing a balance to be struck between qualitative and quantitative analysis of co-design sessions.

The metrics developed by Shah, Smith, and Vargas-Hernandez (2003) for analysing design sessions will be used as the basis for the metrics to be applied in this PhD. Their methodology was chosen because, despite the previously mentioned disagreements about how the metrics should be implemented, the overall method itself remains rather robust. Furthermore, the metrics and methodology developed by Shah, Smith, and Vargas-Hernandez (2003) is considerably less resource intensive in implementation when compared to other methods, such as protocol analysis. Lastly, the metrics and methodology show potential for adaptation to the specific needs of SAR supported co-design sessions, as the methodology for their implementation can be adapted.

2.6.4 Version 2.1

Amongst the first set of metrics used to evaluate design sessions in the SPARK project were the version 2.1 metrics (O’Hare *et al.*, 2016b). These metrics were an adaptation of the metrics defined by Shah, Smith, and Vargas-Hernandez (2003) to suit the purposes of the SPARK project. The un-adapted metrics, as originally defined by Shah, Smith, and Vargas-Hernandez (2003) constitute version 1. The reasoning behind the naming is that these metrics were an update of the original metrics proposed in the initial project proposal and literature review (SPARK Consortium, 2015). Small improvements were then made to the implementation process to obtain version 2.1 (Mombeshora *et al.*, 2017). The timeline for the implementation of the metrics to obtain results from analysing a co-design session is outlined in Figure 27. As can be seen the process typically takes approximately two months within the SPARK project, due in large part to the need to transcribe, and occasionally translate, the design sessions.

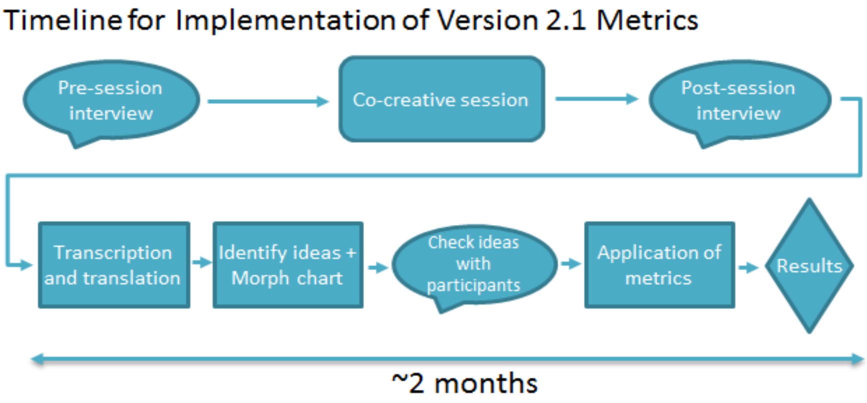


Figure 27 - Process for implementing v2.1 metrics. Typically two months are needed for final results to be made available (Boujut *et al.*, 2017)

Version 2.1 consists of six metrics: Quantity, Quality, Variety, Novelty, Task Progress and Filtering Effectiveness. The first four metrics are those also described by Shah, Smith, and Vargas-Hernandez (2003) however Task Progress and Filtering Effectiveness were developed separately. Filtering Effectiveness is a measure of the number of ideas that were rejected during the session divided by the number of ideas considered minus the number of ideas desired. The resulting value provides an insight into the degree of elimination that can be achieved in the co-design sessions. Task Progress evaluates the progress made toward achieving the tasks set out before the co-design session as well as the participants' creation of new tasks during the session itself (Mombeshora *et al.*, 2017).

It is important to note that the capture of the metrics is built around a set of assistive templates: A Morphological Chart (Variety and Filtering Effectiveness), an Idea Chart (Quality and Quantity), a Novelty Chart, and a Task Progress Chart. The Morphological Chart collects the features of the design such as, for example, the position, number, and colour of buttons. The Idea Chart collects the completed ideas or concepts generated. The Novelty chart is populated using the ideas from the Idea Chart and novelty scores are awarded to each idea by the participants. The Task Progress Chart is populated at the start with a pre-session interview and then revisited after the session in a second interview to identify any tasks completed or new tasks generated during the session.

The major drawback of the version 2.1 metrics is the need for transcripts to populate the Morphological Chart and Idea Chart. This not only causes a delay for the researchers evaluating the session, slowing the research, but also causes the participants to forget the events of the session if any follow up questions are needed. This is especially true for the Novelty metric as the Novelty Chart can only be filled in and evaluated once the Idea Chart has been completed. It was with this reasoning in mind that the need for an equally robust set of metrics, which could be implemented on-the-fly, was identified.

In addition to the need for faster implementation, a failure in the variety metric was identified. Correct calculation of the metric requires knowledge of features that were already present prior to the beginning of the session. However, this proved difficult to implement as the participants often struggled to correctly identify a clear break between features that were new and those that had already been proposed in the past. As a result of these shortcomings a new set of metrics was proposed to aid in redressing the difficulties encountered.

]

2.6.5 Version 3 Metrics Development

The version 3 metrics sought to address the shortcomings of the 2.1 version by adjusting the Idea and Morphological Charts compilation process used to calculate the Quality, Quantity, and Variety metrics as well as, indirectly, the Novelty metric. This was because the compilation of the Idea and Morphological Charts was the most resource intensive. Due to the explorative nature of this development different approaches were implemented simultaneously to guarantee contingency.

Two contingent methods were used. First, the participants were asked to complete a questionnaire containing a set of Likert scales. The intention of this was to be able to calculate the Variety and Quantity of ideas as well as the Task Progress through the use of the questionnaire. The second method was to embed a researcher into the design session to make observations in real time and collect data directly and on-the-fly. They would compile and populate both the Idea and the Morphological Chart as the session progressed to avoid the need to rely on transcripts. Dinar et al. (2015) supports this approach as a method to strengthen any conclusions made during the comparison of the results obtained using the two different methods. Furthermore, the advantage of collecting data on-the-fly is that the post session interviews can be conducted immediately after the session concludes.

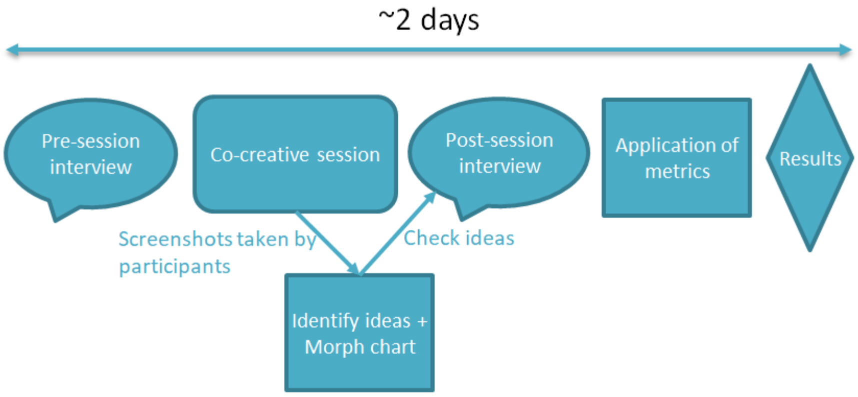


Figure 28 - Process for implementing v3 metrics (Boujut *et al.*, 2017)

Figure 28 shows the process for implementing the version 3 metrics. In summary, the version 3 metrics consist of three major elements: A self-reported questionnaire, a post-interview session with structured questions, and a live observer to record ideas as they are generated (Giunta, 2017). The comparison of the two methods is expanded upon in Section 2.6.6.

2.6.6 Design Research Metrics Validation Tests - Observations and Reflection on the Experiment Setup and Metrics

Two pilot studies were conducted using the version 3 metrics. Each pilot study consisted of one co-design activity. These were performed with industry partners Artefice and Stimulo. Both v3 and v2.1 metrics were deployed so comparisons could be drawn between the metrics. The lessons learned during their implementation aided in the formulation of a final set of metrics.

2.6.6.1 Introduction

Both studies featured a SAR system and followed the same setup: a physical prototype that was augmented using a table-top projector. The projector was connected to a computer which one of the designers from the respective design agency could use to manipulate the projection. This meant that the designers could change logos, colours, and projected textures. All the assets used in the projection were provided a priori by the respective design agencies.

The co-design sessions also consisted of a mix of designers and end-users. In the case of the Stimulo design session, the end-user was a student who met the target demographic. In the Artefice design session the participants were clients who had commissioned the work (Ben-Guefreche *et al.*, 2017). The room layout is shown in Figure 29, in addition to the cameras recording the session the participants were asked to wear individual microphones to record their voices.

When analysing the studies, it is important to understand what a co-design session normally consists of in the context of the SPARK platform experiments. Co-design sessions, as previously mentioned, allow designers and other stakeholders to share ideas and contribute to the design process. While the session is usually led by the designers, who may have some specific questions for the other participants, they are usually not strictly guided into a topic of conversation. The dialogue between participants is allowed to flow freely, thereby enabling all to express their opinions. The general setup for both companies is to explain the process and the needs that the client has expressed. In the case of Stimulo this usually involves a brief product pitch to explain the benefits and functionalities of the product being designed, whereas for Artefice it usually entails explaining the brand identity of their client as well as the message that the packaging should convey. The co-design sessions will usually last anywhere between one to two and a half hours, depending on the complexity of the design, the goals set out for the meeting by the designers, and design process stage.

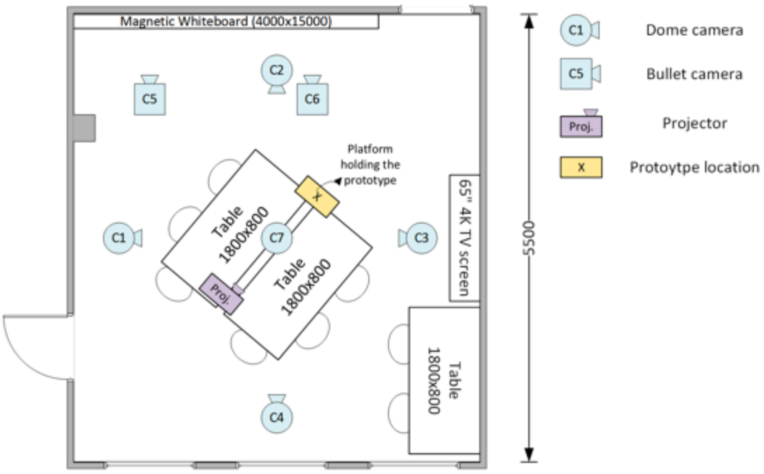


Figure 29 - Layout of SPARK system during the sessions (Ben-Guefreche *et al.*, 2017)

In both sessions the designers were allowed to conduct their design session as they normally would albeit with the SAR system in place. Three observers were placed in a corner of the room where the session was taking place to allow them to record the behaviour of the designers as well as collect data in real-time. At the end of the session the designers were asked to participate in a post-session interview where additional data was collected in line with the version 3 metrics (Giunta, 2017).

2.6.6.2 Results

A full review of the results obtained from the two sessions can be found in Giunta's (2017) MSc dissertation *"SPARK Assessment Metrics Refining and Applying Metrics for the Testing of the SPARK Platform*". This MSc dissertation preceded and provided some of the basis for the work presented in this PhD thesis, in particular, for the development of the metrics to analyse the design sessions. In summary, the results of the sessions for both Artefice and Stimulo found significant variations between the results obtained using the Likert scales and the results obtained by analysing the session using the version 2.1 metrics. In the case of the Artefice sessions, even when taking mitigating factors such as the variety between different types of design sessions, the difference between the values expected, those obtained using the v2.1 metrics, and the values obtained with the v3 metrics, ranged from 54% to 144%. The Stimulo session fared no better and, even with most mitigating circumstances taken into account, the percentage difference between expected values and obtained values ranged from 90% to 206%.

2.6.7 Conclusions

The conclusion of the pilot study was unequivocal: the v3 metrics, and their reliance on self-reported scores through the use of Likert scales, was not a functional approach. However, additional lessons were learned through the process of applying the v3 metrics. First, it was noted that the SAR research platform had the ability to "screenshot" ideas as they were being generated. The feature was originally implemented to allow the designers to better record their progress and look back at their ideas after the end of the session. It was noted that the designers made ample use of the feature and that it did not disrupt their design process at all. Furthermore, it was noted that compiling the Morphological Chart on-the-fly was relatively simple and that the results of doing so were markedly similar to the Morphological Chart compiled using the transcripts (Giunta, 2017). Lastly, the identification of old rows (feature rows in the Morphological Chart that had no new elements as a result of the design session) proved to be complex as the line became blurred between what consisted as a pre-existing feature and how the session added to each feature row.

2.6.8 Development of V4 Metrics

The version 4 metrics were developed based on the lessons learned from the pilot study used to evaluate the version 3 metrics. Much like the version 2.1 metrics, they still include the Quantity, Quality, Variety, Novelty, Task Progress, and Filtering Effectiveness categories (Giunta, 2017). One major change, however, was in how the Variety metric is calculated. By eliminating the old rows, it was possible to avoid the complication of identifying all the features that existed prior to the start of the session. The new approach splits the Variety metric into two separate subcategories: Variety Coverage and Variety New Rows. The former measures the number of old feature rows where new sub-solutions were created during the session. The latter subcategory is the number of completely new feature rows created during the design session. Variety New Rows is best understood by looking at a Morphological Chart and counting the number of rows with completely new features, which resulted from the discussion within the design session itself. Variety Coverage is instead the number of rows where additional features have been identified as a result of the design session. This eliminates the reliance on the old feature rows, albeit sacrificing the ability to normalise between the two subcategories to obtain a singular number for comparison (Giunta, 2017). This method also enables the simple conversion between version 2.1 metrics and version 4 metrics enabling continued comparison between the two.

One additional change made to the Morphological Chart was the redefinition of the feature rows and their comprising features into Idea Element Categories and Idea Elements respectively. This reduced the ambiguity in the vocabulary as well as clarified the specific function of the Morphological Chart (Giunta, 2017). The use of Idea Element instead of feature helps to clarify that the Idea Elements identified in the Morphological Chart, when assembled, allow for the creation of the ideas identified in the Idea Chart. This is in contrast to features which are simply subcomponents of a design and may not be shared between different ideas, thereby complicating the process of filling in the Morphological Chart when multiple ideas are discussed in sequence.

The version 4 metrics continue to implement a pre and post session interview but with a reviewed set of questionnaires (for a full analysis of the metrics and their implementation please refer to *"SPARK Assessment Metrics Refining and Applying Metrics for the Testing of the SPARK Platform*"(Giunta, 2017) and "*D4.1 Definition Of The Experimental Protocol For A Creative Design Process And Case Studies*" (Ben-Guefreche *et al.*, 2017)). A copy of the interview questions can be found in the appendix (section A). An illustrative copy of the Morphological Chart, task progress chart, and idea chart can also be found in the appendix (section B).

The process for the implementation of the metrics follows the same outline as in Figure 28. The Idea Chart is populated automatically by the designers as they make screenshots during their work and the Morphological Chart is populated by the researcher observing the session live. As mentioned, the questions asked in the pre-session interview have been changed to focus more on collecting the tasks to be tackled during the meeting as well as the number of desired ideas to be generated during the session. The post interview session serves as a review opportunity to ensure that no crucial elements of the session have been missed. The tasks completed and opened are reviewed, the Morphological Chart is analysed for discrepancies and Novelty scores are given by the designers to the various ideas.

2.7 SUMMARY

In summary, the review of the literature has revealed that:

- AR, and in particular SAR, has the potential to support design activities.
- There is a lack of research into how SAR could impact design sessions.
- It is currently not clear what, if any, causal links exist between the use of SAR and design session or design process outcomes.

In addition, the review of AR (section 2.3) highlighted the variety of systems that one could deploy. SAR, in particular, can be very varied; many platforms custom made to suit a specific objective leading to a considerable level of heterogeneity between SAR platforms. Therefore, it is better to consider an implementation of AR as a set of characteristics, that define the type of AR, and features, which provide functionality for the given application. By breaking down an SAR system into characteristics and features, it is possible to better generalise any findings, making them independent of the experimental platform used. Should any findings relating to the characteristics of the SAR system be highlighted then one could argue they would be applicable to any other SAR system. Any findings relating to a feature of the experimental platform used here, would be applicable to any other SAR system that shared the same feature.

The review of co-design reveals that co-design as an activity improves the design outcomes generated by participants (section 2.1.2). In addition, it has been shown that mixed-reality AR has considerable potential in supporting this activity with projective SAR being the most applicable, as discussed in section 2.3.4. The applicability of AR for design was further analysed and a collection of papers were reviewed to assess the current implementations of AR and SAR in design (section 2.4). The results of this analysis reinforced the claim that SAR was most suited to supporting co-design sessions. Section 2.5 highlighted the current limitations to the implementation of SAR as a tool to support design sessions, as well as highlighting the requirements that must be met for a wider adoption of the technology by the industry. The section provided a summary of methods for the analysis of SAR in co-design as well as some of the findings of previous research.

The review then continued into the metrics one would use to evaluate the impact of SAR on co-design sessions, as discussed in section 2.6. However, no standards or metrics specifically designed for this task exist and thus required pre-work by the SPARK project to create them ahead of the PhD. It has also been shown, in section 2.1.2.1, that tools/methods used to support co-design can support co-design irrespective of the product that is being designed. Therefore, if one provides evidence that a tool supports a co-design activity then it is highly likely to benefit other co-design activities.

3 METHODOLOGY

Chapter 3 details the research methodology to meet the Research Objectives, answer the Research Questions and achieve the Research Aim (discussed in section3.4). The chapter starts by detailing "What is Research?" so as to position the research with respect to the research landscape and approaches that could have been taken. It then continues into the methodologies typically taken in design research where a discussion ensues on their suitability to answer the Research Questions posed in this thesis. With this understanding, the methodology for this project is presented.

The methodologies adopted for each experiment will be discussed individually in the relevant sections describing the experiments themselves. This is as the methodologies adopted in each different experiment varied significantly to suit the different aims of each experiment.

3.1 WHAT IS RESEARCH

Understanding what research is will aid with setting the stage for the thesis. Furthermore, it evidences the decisions that have been made. The Oxford English Dictionary (2020) provides the following definition:

> *"Systematic investigation or inquiry aimed at contributing to knowledge of a theory, topic, etc., by careful consideration, observation, or study of a subject. In later use also: original critical or scientific investigation carried out under the auspices of an academic or other institution."* (OED Online, 2020)

This dictionary definition, while useful, fails to capture the more nuanced elements of research. Indeed, the definition provided by OED Online (2020) is abstract enough that it could apply to all forms of research, from mathematics to sociology. Robson and McCartan (2016) provide a more applied definition of the purposes of research, in the context of "real world" research conducted in the social sciences. They state:

> *"Three possible purposes of research are commonly put forward – to explore, to describe and/or to explain. Each of these might form the focus of a real world project. It is sometimes claimed that exploring or describing are inferior to explaining; that research worthy of the name should seek to provide explanations."* (Robson and McCartan, 2016)

Clough and Nutbrown (2012) make a distinction between proving and investigating. While the two terms can, and often are, used interchangeably it is important to highlight the point being made. Research is not simply conducted to prove or disprove the null hypothesis; rather research is an observational undertaking, aimed at expanding the knowledge space through observation and inquiry. By looking at research through this lens it becomes less binary and rather than simply seeking to find a yes/no answer to questions becomes more responsive to the realities of the studies being undertaken. New observations, often only possible due to the study itself being run, can be incorporated in the study. In this manner, the hypothesis evolves as the study is conducted and new information is discovered and incorporated.

This is, in part, already reflected in the research questions and objectives. For example, RQ-3 relies on the findings of RQ-2. In the process of answering RQ-2, the knowledge gained by examining and evaluating co-design sessions enables a more nuanced understanding of the characteristics and features of SAR and how they may impact the design process. This in turn allows a focused approach to answering RQ-3 as the most relevant characteristics and features can be selected for assessment

3.2 METHODS FROM SOCIAL SCIENCES: EPISTEMOLOGICAL POSITION

Design can be seen as a problem-solving exercise, aimed at tackling both social and technical challenges. As such, the realm covered by design research is more closely related to the social science fields than those of the harder sciences such as physics or mathematics. It thus makes sense to look at how this field has approached research to understand if inspiration can be gained from there. To understand the scientific approaches in social sciences one must first understand the epistemological foundations upon which they rest. This is no simple matter: there are a large number of competing schools of thought, many diametrically opposed to one-another. The description of the philosophical foundations of the social sciences provided here are then, by necessity, brief ones.

In their book, Real World Research, Robson and McCartan (2016) provide an overview of some of these competing schools of thought. They begin by discussing the differences between the Modernist and Post-Modernist schools of thought. In their description they provide the rationale behind the birth of Post-Modernism, how it has arisen as a critique of the certainty with which the acquisition of knowledge is pursued by some Modernists and how this can be seen as detrimental in the social sciences where a causal relationship is often hard to establish. While not rejecting this positivist view outright, Robson and McCartan (2016) provide some insight into quantitative and qualitative research paradigms. Ultimately, they recommend a realist approach to science as a “pragmatic” approach to research in the social sciences.

3.2.1 Quantitative Research

Robson and McCartan (2016) summarise quantitative research by focusing of the following aspects:

Quantitative research focuses on measurement and quantification, this is often achieved by turning available information or inputs into numbers. High accuracy and precision, both during the conversion and the processing, of this information into quantifiable data is desirable. There is often a focus on the behaviour of participants, be it groups or individuals.

Quantitative research strives to adhere to a scientific approach, attempting to closely match the processes and approaches found in natural sciences, such as chemistry or physics. As such, deductive reasoning is often applied. Pre-existing theories are tested as are concepts and other theoretical ideas.

Quantitative research studies are thus designed to support this approach. The research approach, methods, and outcomes are pre-defined at an early stage within the design of the study, and closely adhered to. The reliability and validity of the study are also often heavily emphasised within quantitative research studies. This means that the studies value consistency of the output data, regardless of time, observers, or the effectiveness of the data capture. Furthermore, the replicability of the study is often also emphasised, and as such a detailed procedure for the methodology is provided to enable reproducibility.

In addition Robson and McCartan (2016) highlight that quantitative studies will often include some form of post-processing. This is often to allow for a generalised application of the findings. As such, the samples used must be representative of the general population being studied. To aid this statistical analysis, as well as the validity and replicability of the results, objectivity is considered paramount. Distance between researchers and participants is thus desirable. Furthermore, in order to support the controllability of the study, its replicability, and its accuracy, the experiments are often standardised. This often results in some form of decontextualization and leads to some form of artificiality being injected into the study or condition being investigated.

In summary, Robson and McCartan (2016) state that quantitative research aims for a "neutral, value-free position". Quantitative research thus attempts to enumerate all input data, as well as formalise the collection of said data. This is in an attempt to maintain objectivity, replicability, and an adherence to the scientific approach.

3.2.2 Qualitative Research

Robson and McCartan (2016) also discuss qualitative research in social sciences as a contrast to quantitative research. Qualitative research differs from quantitative research by shifting the focus away from quantifiable metrics and towards attempting to understand problems at a system level, often through a thorough description of the context and scenarios. Their views are summarised as follows:

Qualitative research values the presentation of findings in the form of accounts of events. These are often presented verbally or through other non-numerical means. Qualitative research places little value on numerical data or any statistical analysis thereof (Robson and McCartan, 2016). Data analysis when using qualitative research techniques focuses on the use of inductive reasoning. Any data collected is analysed to form theories and concepts on which the research can then build or analyse further. As such, Robson and McCartan (2016) state, there is an increased focus on the meanings behind facts and events to understand the underlying patterns, links, and connections. Thus, context is seen as a vital element of the analysis process. Proponents of qualitative research argue that stripping data from the context within which it was obtained robs the researcher of the ability to gain a deeper understanding of the phenomena, their causes, and interactions (Robson and McCartan, 2016). For this reason, when data is gathered, it is often coded or transcribed from the perspective of the parties involved, avoiding the use of an impersonal third person perspective. Ultimately, proponents of qualitative research argue, the social world is a construct created by those who live in it, and are involved with it, and as such only a holistic approach can hope to fully understand its processes.

Within the qualitative research paradigm, research is not necessarily always fully planned out from the start. Exact goals and outcomes may not be predefined before studies are carried out. This lack of pre-set end conditions enables a more flexible and responsive approach to research design and methodology. This allows the research to shift focus as new factors come into play and respond to them holistically. As part of this holistic approach to research, qualitative researchers will accept the values of the participants and the researchers themselves, and that these may have an influence on the final outcome of the study. As a result, objectivity within qualitative research is not particularly valued, as it is seen as promoting detachment between the researcher and the participants (Robson and McCartan, 2016). This approach means that the results obtained from any one study are not easily generalizable, but this is often not seen as a major concern. Conversely, openness and receptivity to others is valued (Robson and McCartan, 2016). In addition, the personal commitment and self-awareness of the researchers are valued (Robson and McCartan, 2016).

In order to diminish any detachment from the reality of the scenarios being studied, qualitative research is rarely undertaken in-vitro or in abstracted situations, such as a laboratory. Research is often conducted on location and on a small scale, with few participants and a small number of cases being evaluated and assessed.

3.2.3 Realism

In an attempt to reconcile the two approaches of qualitative and quantitative research Robson and McCartan (2016) recommend a realist approach. Robson and McCartan (2016) argue that one of the major concerns when conducting real-world research is that problems and issues that include a social element are fluid. As such, any approach that attempts to address the inherent complexities of dealing with the fluidity of social behaviour, such as qualitative methods, is certainly worth considering. However, in the interest of maintaining a degree of replicability and rigour in the experiments conducted, a compromise must be found. The philosophy behind this approach is summarised by Robson and McCartan (2016) as follows:

Realists agree, to some extent, that science is not unquestionable. Facts are never beyond dispute and knowledge is merely a construct, a product of shared historical and societal interactions. As such, any fact will be steeped in theory and inextricable from the theory that spawned it. However, realists agree with positivists, to a degree, that the purpose of science is to create theories and test them empirically and rationally.

Thus, explanations of the world are based on describing how specific mechanisms act to obtain predetermined outcomes. Robson and McCartan (2016) paraphrase this by stating that “The guiding metaphors are of structures and mechanisms rather than phenomena and events”. A law is then seen as a “specific pattern of an activity” or the “tendency of a mechanism”. These laws act as a statement defining what is effectively happening and how things can develop from there.

Realists accept that the world is exceedingly complex and difficult to interpret effectively but are also aware that reality can often be subcategorised. In the context of social sciences, Robson and McCartan (2016) state that reality can incorporate individuals, groups or institutions, and entire societies. Causation is seen as the interactions of these entities as a function of their underlying structure. An explanation of events is then a clarification of how different factors have led to a specific outcome; even when it is not possible to predict outcomes, they can none the less be explained retroactively.

Realism thus serves to reconcile the critiques of positivist philosophy, which undermine quantitative research in social sciences, with the lack of subjectivity and repeatability of qualitative research approaches.

3.2.4 Summary

The research questions and research objectives, mentioned in sections 3.4.2 and 3.4.3 respectively, suggest a realist methodological approach. Indeed, RQ-1: “How can co-design sessions’ efficacy be measured?” and RQ-2: “How does an SAR system affect co-design sessions’ efficacy?” highlight the need for quantifiable metrics to evaluate design session outcomes as well as the impact that SAR can have on design sessions. Such metrics must, by necessity, be rigorous and repeatable across multiple different designs sessions. However, when attempting to further understand the system level impact of SAR as well as the requirements of the stakeholders, as listed in RQ-3: “How do specific SAR characteristics and features affect co-design sessions’ efficacy?” and RQ-4: “What are the industry requirements for an SAR system to support co-design?” a more holistic approach is required.

The decision to proceed with a realist approach to developing the methodology that supports the work presented in this thesis is thus a pragmatic one. There must be an acknowledgement of the limitations of quantitative research approaches in particular when a social context, such as design, is being analysed. However, the lack of rigour and poor repeatability present in qualitative driven approaches is not fully appropriate either. This is due, in part, to the fact that different partners will be contributing to the research from different institutions. The approach selected must enable the data collected at different locations to be comparable. Furthermore, by improving the rigour of the experiments conducted, it will be possible to compare the results from different experiments as the SAR platform is refined and developed. This will enable a better understanding of how the changes made to the platform affect the design session outcomes and will enable a more thorough understanding of which changes are having an impact.

3.3 RESEARCH IN ENGINEERING DESIGN

Approaches to research based on realism (section 3.2.3) fit the general practices of engineering design well. One example of this would be the Design Research Methodology (DRM) framework proposed by Blessing and Chakrabarti (2009). This handbook provides a broad discussion on the various approaches to address different types of design research. Rather than prescribing a specific approach to design research, the DRM discusses the different elements that should be included within design centred research. In doing so it acknowledges the nonstandard nature of design, thus allowing researchers to better customise and approach their specific design research problem in a manner that will yield valid and robust results.

In contrast, “A design science research methodology for information systems research” (Peffers *et al.*, 2007) and “Design Science in Information Systems Research” (Hevner *et al.*, 2004) provide a more focused look at design research methodology aimed specifically at the development of information systems. Their work provides a much more direct and tangible approach to research. However, their approaches tend to focus on an industry pull, the need to address a specific desire or problem posed by industry as well as the development of “artefacts” tangible hardware and software solutions. This approach, while somewhat reflective of the work presented in this thesis due to the nature of the SPARK project, was deemed overly restrictive. As mentioned in Section 1.3, and further highlighted in Figure 4, the work performed as part of this PhD extended beyond the original scope and goals of the SPARK project. As such, the proposed methodologies of Peffers *et al.* (2007) and Hevner *et al.* (2004) were deemed to be overly restrictive due to their focus on the development of artefacts as well as the direct focus on indurstry needs rather than a more generic approach aimed at answering design research questions.

In addition to the reasons mentioned above, the DRM framework proposed by Blessing and Chakrabarti (2009) meshed well with the realist approach to research discussed previously in section 3.2.3. As such, this framework was selected as a structure to develop the methodology used.

3.3.1 DRM, a Design Research Methodology

Using the DRM framework (Blessing and Chakrabarti, 2009) it is possible to analyse how the research questions may be answered. Figure 30 shows the basic method for implementing the DRM. The methodology starts with setting measurable criteria that are to be achieved by the end of the work. This first stage of the methodology concerns clarifying the research; as part of their description of the Research Clarification stage Blessing and Chakrabarti (2009) state that the purpose of this stage is to identify the goals that the research will achieve thereby pinpointing the focus of the research. This leads to determining the main research problems, questions, and hypotheses.

The second stage presents the first Descriptive Study where one observes and analyses the problem, as well as contextualises it. Methods often employed include additional reviews of the existing literature, in particular with regard to existing empirical studies performed in the area of interest. In addition, the second stage also focuses on running one's own empirical studies to collect data. Lastly, Blessing and Chakrabarti (2009) mention the use of reasoning, using data gathered either through reviewing literature or primary research to understand the "lay of the land". Literature is used to aid in the process of better defining and shaping the understanding of the challenges ahead.

The Prescriptive Study is then used to generate a scenario of the desired outcome. This enables the development of a method that will achieve this outcome. Furthermore, the Prescriptive Study enables the assumptions made during the development of the method, and by extension the method itself, to be tested. Lastly, the second Descriptive Study attempts to apply the method previously developed and evaluate it in two tests. The first focuses on the application itself, does the method affect the factors it should affect and are the expected results indeed as expected. The second test evaluates the success of the method by evaluating its overall achievement of the stated initial criteria (Blessing and Chakrabarti, 2002). The information obtained as a result of this second Descriptive Study can be fed back into the first or into the Prescriptive Study to refine the method or enhance the understanding of the literature.

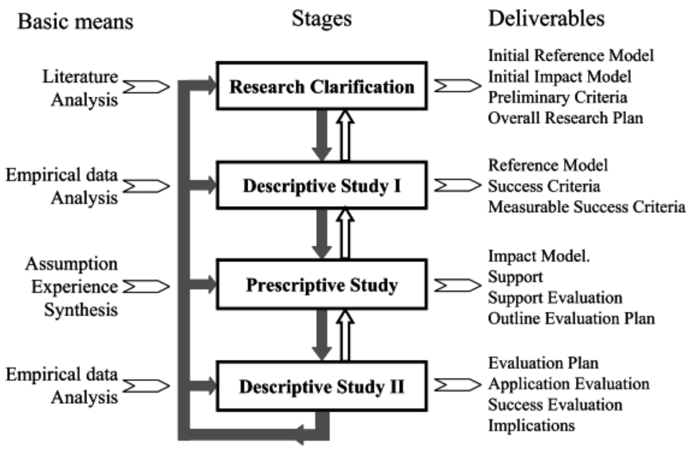
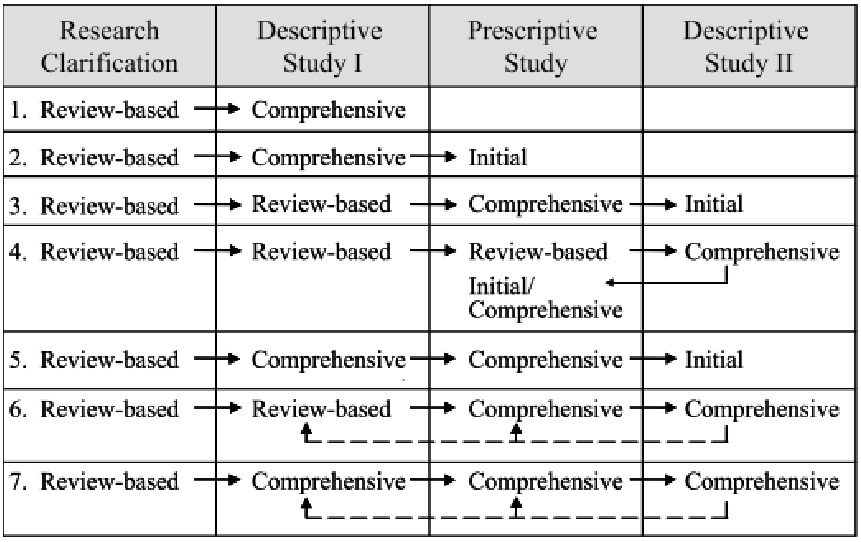


Figure 30. Basic method of DRM implementation (Blessing and Chakrabarti, 2009)

Figure 31 shows the types of design research projects as classified by Blessing and Chakrabarti (2009). Review-based studies are based on a review of the literature only, whereas a comprehensive study is where empirical research is performed by the researcher. It is important to note that all comprehensive studies will include a review-based study (Ibid).



Research Clarification	Descriptive Study I	Prescriptive Study	Descriptive Study II
1. Review-based →	Comprehensive		
2. Review-based →	Comprehensive →	Initial	
3. Review-based →	Review-based →	Comprehensive →	Initial
4. Review-based →	Review-based →	Review-based → Initial/ Comprehensive ←	Comprehensive
5. Review-based →	Comprehensive →	Comprehensive →	Initial
6. Review-based →	Review-based →	Comprehensive →	Comprehensive
7. Review-based →	Comprehensive →	Comprehensive →	Comprehensive

Figure 31. Types of Design Research Projects (Blessing and Chakrabarti, 2009)

In summary, Figure 30 shows the general methodological approach to designing design research studies as recommended by Blessing and Chakrabarti (2009). As can be seen in the figure they divide research into four main stages: Research Clarification, Descriptive Study I, Prescriptive Study, and Descriptive Study II.

Broadly speaking, the functions of each stage are the following: the Research Clarification sets the initial criteria to be achieved as part of the research as well as forming an initial model of the problem. The Descriptive Study I homes in on the problem to be analysed and aids with the development of tangible success criteria. The Prescriptive Study serves to provide empirical evidence to compare with the model created as part of the Descriptive Study and uses this evidence to test the assumptions made there. Finally, the Descriptive Study II focuses on recontextualising the findings of the previous studies to understand if they have actually had the intended impact as well as identifying any improvement to the research approach.

Not all research will include all four stages, however. Figure 31 highlights this by categorizing design research into seven categories, separated by which stages they adopt and how deeply they research each stage.

3.4 RESEARCH SCOPE AND METHODOLOGY

3.4.1 Aim

Having defined SAR and Collaborative Design, as well as the relevance of one to the other, one can now present the aim of the research. The research aims to investigate the applicability and usability of SAR tools to support co-design sessions, specifically the aim of this research is to:

INVESTIGATE SAR SYSTEMS' CHARACTERISTICS AND FEATURES AND HOW THESE AFFECT THE EFFICACY OF CO-DESIGN SESSIONS

The examination will analyse the impact of the SAR platform on co-design sessions not just at a systems level but also the features and characteristics of SAR.

The Research Aim, "Investigate SAR Systems' Characteristics and Features and how these Affect the Efficacy of Co-Design Sessions ", lends itself to being classified as a type five design research project: Development of Support Based on a Comprehensive Study of the Existing Situation.

Type five design research projects are recommended for situations where the scope is to develop support, but the understanding of the existing situation is poor. A comprehensive Descriptive Study I enables the development of understanding, a comprehensive Prescriptive Study enables the development of support and an initial Descriptive Study II evaluates the implementation (Blessing and Chakrabarti, 2009).

A type five design research project framework was adopted as the literature had highlighted a lack of a comprehensive understanding of the implementation of SAR in design, as well as a lack of information on the broader adoption of SAR outside the field of design (Sections 2.3.4, 2.4.3, and 2.5.2). However, Blessing and Chakrabarti (2009) themselves note that, as research progresses, it may become necessary to change the approach on the basis of the information collected through the literature and the results of any experiments conducted.

3.4.2 Research Questions

Based on the review, the aim was unpacked into the following RQs:

RQ-1: "HOW CAN CO-DESIGN SESSIONS' EFFICACY BE MEASURED?"

RQ-1 focuses on understanding the methods by which co-design sessions can be evaluated. This is both through a qualitative approach to analysing the direct session outcomes as well as understanding which quantitative approaches can provide additional insight into how the co-design session has progressed. This lays the foundations for future work, as by answering RQ-1, it is possible to analyse the impact that the use of SAR has on co-design. This in turn enables the achievement of the aim.

RQ-2: "HOW DOES AN SAR SYSTEM AFFECT CO-DESIGN SESSIONS' EFFICACY?"

This research question attempts to look at the overall impact of SAR on co-design. In answering RQ-2 it is possible to begin to understand the broader impact that SAR has on co-design. It also enables an explorative approach to the research. As the field of SAR, as applied to co-design, is relatively novel it is necessary to understand the bigger picture before any finer analysis can be performed.

RQ-3: "HOW DO SPECIFIC SAR CHARACTERISTICS AND FEATURES AFFECT CO-DESIGN SESSIONS' EFFICACY?"

RQ-3 builds upon RQ-2 in the pursuit of the aim. Based, in part, on the findings of RQ-2 it is possible to understand some of the characteristics and features that show promise for additional analysis. RQ-3 attempts to guide the research towards understanding the specific elements of SAR that may prove beneficial for co-design. This directly targets the aim of the thesis as by answering RQ-3 it will be possible to gain an understanding of how specific characteristics and features of SAR support or hinder co-design.

RQ-4: "WHAT ARE THE INDUSTRY REQUIREMENTS FOR AN SAR SYSTEM TO SUPPORT CO-DESIGN?"

In addition to understanding how the characteristics and features of SAR impact co-design, it is important to understand how the industry will react to SAR. Additionally, members of industry may request or veto specific features or characteristics of SAR. As such, understanding their perspective and integrating it into any final recommendations on how SAR may affect co-design is imperative to promote the adoption of the technology. This aids in guaranteeing the long-term impact of the research.

3.4.3 Research Objectives

To achieve the stated aim and answer the research questions, six research objectives were laid out. The desired outcome for these objectives is threefold. Firstly, the objectives should broadly investigate SAR for collaborative design. This should allow for the eventual development of guidelines for the effective application of SAR in collaborative design. Lastly, this should result in an improved understanding of the effect an SAR system's characteristics and features has on the efficacy of co-design sessions in product and packaging design.

These desired outcomes can be met through the following research objectives:

RO-1: "DEVELOP A METRIC FRAMEWORK FOR EVALUATING THE EFFICACY OF CO-DESIGN SESSIONS"

RO-1 is linked to RQ-1. It attempts to address the Research question by attempting to identify and select a metric framework by which the subsequent research can be conducted.

RO-2: "DESIGN AND DEVELOP AN SAR PLATFORM FOR USE IN CO-DESIGN SESSIONS"

This Research Objective attempts to address both RQ-1 and RQ-2. In the case of both research questions the presence of an SAR platform with which to conduct the studies is assumed. However, it is necessary to actively create a platform that meets the needs of the research.

RO-3: "EVALUATE THE EFFICACY OF AN SAR PLATFORM IN COMPLETE CO-DESIGN SESSIONS"

RO-3 addresses RQ-2 by analysing the effectiveness of an SAR platform in complete co-design sessions. In doing so, it is possible to understand the overall impact of SAR on co-design and to begin to understand the features and characteristics that most impact the process of design using SAR. Additionally, this Research Objective serves to expand upon the existing literature to confirm whether the pursuit of SAR as a tool for co-design is valid, or if a different approach is required.

RO-4: "ANALYSE THE IMPACT OF A SAMPLE OF SPECIFIC CHARACTERISTICS AND FEATURES OF THE SAR PLATFORM ON CO-DESIGN SESSIONS"

RO-4 addresses RQ-3, by utilizing the knowledge gained through the achievement of RO-3 to understand which features or characteristics of SAR are most valuable to realistically explore in detail.

RO-5: "CAPTURE INDUSTRY INPUT TO THE DEVELOPMENT OF AN SAR PLATFORM, AND ANALYSE THEIR RESPONSE TO THE IMPLEMENTATION OF THE SAR PLATFORM"

RO-5 addresses RQ-4 by investigating industry input to understand their requirements for SAR. In doing so, the Research Objective hopes to shed light on the features and characteristics of SAR that industry members believe are valuable. Additionally, the Research Objective hopes to understand how SAR can be integrated into the current workflow, thus allowing for recommendations that are better suited to promoting adoption of the technology.

RO-6: "PROPOSE RECOMMENDATIONS/GUIDELINES FOR THE DEVELOPMENT OF SAR PLATFORMS FOR CO-DESIGN"

This Research Objective's purpose is to collect and summarise all the knowledge generated throughout the studies in order to address the aim by showing the results of the investigation and collating them for future researchers to utilise.

Table 11 places these research objectives are in relation to the research questions. From the table, it is possible to see that each objective serves to address at least one of the research questions.

Table 11. Research Objectives vs Research Questions

	RQ-1: "How can co-design sessions' efficacy be measured?"	RQ-2: "How does an SAR system affect co-design sessions' efficacy?"	RQ-3: "How do specific SAR characteristics and features affect co-design sessions' efficacy?"	RQ-4: "What are the industry requirements for an SAR system to support co-design?"
RO-1: "Develop a metric framework for evaluating the efficacy of co-design sessions"	X			
RO-2: "Design and develop an SAR platform for use in co-design sessions"	X	X		
RO-3: "Evaluate the efficacy of an SAR platform in Complete Co-Design Sessions"		X		
RO-4: "Analyse the impact of a sample of specific characteristics and features of the SAR platform on co-design sessions"			X	
RO-5: "Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform"				X
RO-6: "Propose recommendations/guidelines for the development of SAR platforms for co-design"		X	X	X

3.4.4 Achievement of the Research Objectives

This section reviews the Research Objectives, first presented in section 3.4.3, with the intent of providing a method for achieving each Research Objective. An additional rationale, grounded in part in the methodology literature reviewed, is also provided.

RO-1: "Develop a metric framework for evaluating the efficacy of co-design sessions"

> *Method: literature review on existing metrics for design, leading to a proposed framework geared to the constraints of the available industrial setting for the co-design sessions. Metrics to be tested and iterated to refine their application.*

Rationale for the proposed method: metrics for the evaluation of design sessions exist but will, in all likelihood, need to be adjusted to the specific needs of SAR supported design sessions. Adjusting existing metrics and testing them to ensure their validity should lead to a robust metric framework with support both from the literature and empirical research.

RO-2: "DESIGN AND DEVELOP AN SAR PLATFORM FOR USE IN CO-DESIGN SESSIONS"

Method: develop an SAR platform based on the suggestions and feedback from the industry partners to closely match their use cases. Iterate the design as necessary to more closely match these needs while allowing research to be undertaken.

Rationale for the proposed method: by tailoring the platform to the use cases presented by the industry partners it will be easier to integrate the use of the research platform into their design process, enabling easier and more realistic data collection. Making necessary adjustments to the platform will allow for the continued use for research purposes while still allowing the results from different versions to be cross-comparable.

RO-3: "EVALUATE THE EFFICACY OF AN SAR PLATFORM IN COMPLETE CO-DESIGN SESSIONS"

Method: utilise the SAR platform to support a number of complete design sessions and then assess these sessions using the metrics developed in RO-1.

Rationale for the proposed method: complete design sessions provide a realistic scenario for the evaluation of SAR, both potentially highlighting specific characteristics and features of SAR to analyse in additional depth as well as providing insight into the benefits or drawbacks of SAR use. In addition, comparing the results from SAR supported sessions to those from non-SAR sessions will aid in understanding

RO-4: "ANALYSE THE IMPACT OF A SAMPLE OF SPECIFIC CHARACTERISTICS AND FEATURES OF THE SAR PLATFORM ON CO-DESIGN SESSIONS"

Method: based, in part, on the results from RO-3 identify specific characteristics or features of SAR that can be considered valuable to further explore. Targeted and controlled experiments will aim to reduce the number of independent variables in order to fully understand the impact of specific characteristics or features.

Rationale for the proposed method: the experiments mentioned for RO-3 are aimed at understanding the overall value of SAR as a tool to support design. More targeted studies could potentially identify causal or correlative relationships between desired outcomes and highlight specific characteristics or features of SAR. These could then also be adjusted in future SAR platforms in order to improve design session outcomes, thereby forming a basis for further research.

RO-5: "CAPTURE INDUSTRY INPUT TO THE DEVELOPMENT OF AN SAR PLATFORM, AND ANALYSE THEIR RESPONSE TO THE IMPLEMENTATION OF THE SAR PLATFORM"

Method: questionnaires, targeted studies, and interviews with a broader industry audience to capture their opinions on SAR as a tool to support design sessions.

Rationale for the proposed method: by understanding the input from industry, it will be possible to tailor recommendations for future SAR platforms to better suit their needs, enabling improved adoption of SAR.

RO-6: "PROPOSE RECOMMENDATIONS/GUIDELINES FOR THE DEVELOPMENT OF SAR PLATFORMS FOR CO-DESIGN"

Method: collate a list/table/guide of the major results as well as important or interesting findings from the studies collected in this thesis.

Rationale for the proposed method: by creating a single list or guide of important findings, it will be possible for future researchers to develop SAR platforms better suited and tailored to supporting design sessions. In addition, such a list will aid in preventing pitfalls as well as highlighting knowledge gaps for future research.

The planned contributions of this thesis are listed below in Table 12 and are linked to the above-mentioned research objectives as well as providing a link to the specific chapters where each research objective will be addressed.

Table 12. Research Objectives and planned work to address them

Research Objective	**CONTRIBUTION**	**RELEVANT CHAPTER(S)**
RO-1: "Develop a metric framework for evaluating the efficacy of co-design sessions**"**	Provide a robust, reliable, and simple approach that can provide relevant data on the output of design sessions that make use of SAR technologies	2.6
RO-2: "Design and develop an SAR platform for use in co-design sessions"	An SAR platform that can be utilised to support collaborative design sessions and whose functions and characteristics are known	4
RO-3: "Evaluate the efficacy of an SAR platform in Complete Co-Design Sessions"	Provide an analysis of realistic design sessions and develop a baseline that enables future comparison of SAR's impact on other design sessions	5
RO-4: "Analyse the impact of a sample of specific characteristics and features of the SAR platform on co-design sessions"	Determine the overall impact that SAR has had on collaborative design sessions and how specific features and characteristics of the SAR system have influenced this impact	6
RO-5: "Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform"	Provide insight into industry's response to the impact that SAR can have on collaborative design sessions	7
RO-6: "Propose recommendations/guidelines for the development of SAR platforms for co-design"	A set of recommendations for the future development of SAR systems for supporting collaborative design sessions	8

3.4.5 Thesis Structure and Methodological Implementation

Table 13 provides an overview of how the chapters of this thesis will address the previously stated research objectives and research questions in an attempt to achieve the aim. Table 15 expands on Table 13 by providing detailed insight into how each of the planned research activities aim to achieve the research objectives.

Table 14 expands on the information presented in Table 13. As mentioned, the work presented in this thesis was, in part, conducted in collaboration with the SPARK Consortium. Table 14 expands on this relationship by highlighting which elements of the work was led by the SPARK Consortium, and identifies where the author's contribution lies. It should be noted that some of the data collected by the SPARK Consortium was reanalysed in the context of this thesis. As a result, the conclusions drawn from some of the experiments, in particular where the author expanded on the original study conducted, may have diverged from those published by the SPARK Consortium as a whole.

Table 13. Chapters Mapped to Design Research Methodology, highlighting the contributions

Chapter	Description	Methodological Stage	Research Questions	Research Objectives
Section 1: Literature Review and Research Clarification				
2.1	Collaborative Design	Research Clarification		
2.3	Augmented Reality	Research Clarification		
2.4	Augmented Reality in Design	Research Clarification	RQ-1	
2.5	SAR in Co-Design	Research Clarification	RQ-1	
2.6	Developing Design Research Metrics	Descriptive Study I	RQ-1	RO-1
Section 2: The Development of an SAR Research Platform				
4.1	Requirements for an SAR Research Platform	Research Clarification	RQ-2	RO-2, RO-5
4.2	Description of The SAR Research Platform	Research Clarification		
4.3	Development of the Portable SAR System	Descriptive Study I		RO-2
4.3.2	Technical Challenges to Meet the Requirements	Research Clarification	RQ-1	RO-2
4.4	SAR Platforms' Suitability in Addressing Research Questions and Objectives	Descriptive Study I	RQ-3	RO-2, RO-5
Section 3: Analysing the Impact of SAR on Complete Design Sessions				
5.1	Comparing SAR and existing Co-Design Tools	Descriptive Study I	RQ-2	RO-3
5.2	SAR Platform Validation at End-Users' Premises	Descriptive Study I	RQ-2	RO-3
Section 4: Analysing the Impact of Key SAR Features and Characteristics				
6.1	Interface Comparison Study	Prescriptive Study	RQ-3	RO-4
6.2	Impact of SAR on Communication between Design Session Participants	Prescriptive Study	RQ-3	RO-4
6.3	Impact of Scale in Design Sessions Supported by an SAR Platform	Prescriptive Study	RQ-3	RO-4
Section 5: Feedback from Designers and Other Potential End-Users				
7.1	Industry Feedback from Trade Fairs	Descriptive Study II	RQ-4	RO-5
7.2	SAR Platform Pilots with Industry Members	Descriptive Study II	RQ-4	RO-5
7.3	Longitudinal Analysis of SAR Impact on the Design Process	Descriptive Study II	RQ-4	RO-3, RO-5
Section 6: Discussion and Conclusion				
8.4	Recommendations for SAR Platform Development	Descriptive Study II		RO-6

Table 14. Breakdown of Lead for the Work Presented within this Thesis

CHAPTER	DESCRIPTION	LEAD	AUTHOR'S CONTRIBUTION
SECTION 1: LITERATURE REVIEW AND RESEARCH CLARIFICATION			
2.1	Collaborative Design	Author	
2.3	Augmented Reality	Author	
2.4	Augmented Reality in Design	Author	
2.5	SAR in Co-Design	Author	
2.6	Developing Design Research Metrics	Author	
SECTION 2: THE DEVELOPMENT OF AN SAR RESEARCH PLATFORM			
4.1	Requirements for an SAR Research Platform	SPARK Consortium	Technical support and development
4.2	Description of The SAR Research Platform	SPARK Consortium	Technical support and development
4.3	Development of the Portable SAR System	Author	
4.3.2	Technical Challenges to Meet the Requirements	Author	
4.4	SAR Platforms' Suitability in Addressing Research Questions and Objectives	Author	
SECTION 3: ANALYSING THE IMPACT OF SAR ON COMPLETE DESIGN SESSIONS			
5.1	Comparing SAR and existing Co-Design Tools	SPARK Consortium	Data collection and analysis
5.2	SAR Platform Validation at End-Users' Premises	SPARK Consortium	Data collection and analysis
SECTION 4: ANALYSING THE IMPACT OF KEY SAR FEATURES AND CHARACTERISTICS			
6.1	Interface Comparison Study	SPARK Consortium	Data collection and analysis
6.2	Impact of SAR on Communication between Design Session Participants	Author	
6.3	Impact of Scale in Design Sessions Supported by an SAR Platform	Author	
SECTION 5: FEEDBACK FROM DESIGNERS AND OTHER POTENTIAL END-USERS			
7.1	Industry Feedback from Trade Fairs	SPARK Consortium	Data collection and analysis
7.2	SAR Platform Pilots with Industry Members	SPARK Consortium	Data collection and analysis
7.3	Longitudinal Analysis of SAR Impact on the Design Process	SPARK Consortium	Data collection and analysis
SECTION 6: DISCUSSION AND CONCLUSION			
8.4	Recommendations for SAR Platform Development	Author	

Table 15. Research Objectives Mapped to the Planned Research Activities to address them

RESEARCH OBJECTIVES	PLANNED RESEARCH ACTIVITIES				
	LITERATURE REVIEW AND RESEARCH CLARIFICATION	THE DEVELOPMENT OF AN SAR RESEARCH PLATFORM			
	2.6	**4.1**	**4.3**	**4.3.2**	**4.4**
RO-1	Development of the research metrics used to analyse the SAR supported design sessions				
RO-2		Develop a set of requirements for an SAR platform	Design and develop a portable SAR platform	Analyse the technical obstacles and limitations in developing a functional rig to support experiments.	Analyse the characteristics and features of the SAR platform and its impact on the research to be conducted
RO-3					
RO-4					
RO-5		Use industry input from previous experiments and literature review to guide development of requirements			Understand the industry input to aid in the development of future studies targeted at their needs
RO-6					

RESEARCH OBJECTIVES	PLANNED RESEARCH ACTIVITIES (CONT.)				
	SUPPORTING CO-DESIGN THROUGH SPATIAL AUGMENTED REALITY		STUDIES INTO SAR SYSTEM CHARACTERISTICS AND FEATURES		
	5.1	**5.2**	**6.1**	**6.2**	**6.3**
RO-1					
RO-2					
RO-3	Conduct experiments aimed at comparative testing between existing design tools and SAR using pre-set design tasks	Conduct experiments aimed at analysing the impact of SAR or open-ended design sessions			
RO-4			Conduct experiments analysing specific features and characteristics of SAR, focusing on interfaces	Conduct experiments analysing specific features and characteristics of SAR, focusing on collaboration	Conduct experiments analysing specific features and characteristics of SAR, focusing on Physical prototype scale
RO-5					
RO-6					

RESEARCH OBJECTIVES	PLANNED RESEARCH ACTIVITIES (CONT.)			
	STUDIES WITH INDUSTRY			DISCUSSION AND CONCLUSION
	7.1	7.2	7.3	8.4
RO-1				
RO-2				
RO-3			Provide additional evidence of SAR's efficacy, or lack thereof, in design sessions over the course of the entire design development process	
RO-4				
RO-5	Collect data to assess wider industry requirements for SAR adoption	Conduct open-ended design sessions with industry members who are not SPARK Consortium members	Analyse the long-term impact of SAR adoption on the design process	
RO-6				List the recommendations for future development of SAR platforms to support design

4 THE DEVELOPMENT OF AN SAR RESEARCH PLATFORM

The study of SAR's characteristics and features necessitates the development of a research platform that can enable their modification and manipulation. This chapter describes the development of two SAR systems based on industry feedback on their co-design needs. The chapter begins with a discussion of the requirements for the SAR research platform that has been derived from Chapter 2's literature review. This chapter then continues by describing the SAR research platform developed as part of the SPARK project, detailing the platform's components and setup. Thereafter the suitability of the platform developed is assessed within the context of the Research Questions and Objectives of this thesis. On the basis of this assessment a list of technical challenges is identified, and a further SAR research platform is proposed to fill the gaps. The chapter concludes with a discussion of the characteristics and features of SAR that the new platform can be used to investigate as well as an overall summary for the chapter.

4.1 REQUIREMENTS FOR AN SAR RESEARCH PLATFORM

The following section will focus on the development of both the SPARK platform and its derivative: miniSPARK. The functionalities, features, and characteristics of the SPARK platform are discussed here as well. The motivation for the development of miniSPARK as an offshoot of the SPARK platform is also discussed.

Section 2.3 highlighted that the definition for SAR is based predominately on the positioning of the AR platform in relation to the viewer. Section 2.3.3 provides additional clarity by providing examples of SAR technologies, in particular Projective SAR technologies. However, one crucial detail was noted for SAR in general and Projective SAR in particular: there is a lack of standardization in the deployment of SAR platforms.

It now becomes useful to outline the major characteristics that allow SAR to be considered as such, as well as take note of the features that can be implemented within SAR platforms to support a specific task. This distinction was first introduced in section 1.4, where the research focus was initially laid out. As noted in section 2.3.4, unless otherwise specified, references to SAR are specifically references to Projective SAR.

The distinction made here is thus: a characteristic is a core, defining element of SAR. Without such a characteristic, a platform cannot be considered to be a type of SAR technology. The categorization of these characteristics draws upon the knowledge assembled in sections 2.3 and 2.5, where augmented reality and SAR in design were, respectively, discussed.

The features of SAR are defined as those elements of an SAR platform that may be included within a typical SAR deployment, but whose presence is not necessary. As a result, if a feature were to be omitted from an SAR platform, it would still unmistakably be recognizable as an SAR platform.

4.1.1 Characteristics of SAR

Table 16 lists six characteristics found in an SAR platform. These are: Interface, Projection, Projection Surface, Real Time Interactivity, Hardware, and Software. For each characteristic, a description and rationale is provided. They have been distilled by way of logical and inductive analysis of a theoretical SAR platform based on the literature previously discussed in sections 2.3 and 2.5. The characteristics identified were identified as such due to the intrinsic need for an SAR platform to contain all of them, in some form, in order for the platform to be defined as SAR.

Table 16. Table listing characteristics of SAR, their definitions, and the underlying rationale

#	CHARACTERISTIC	DESCRIPTION AND RATIONALE
1	Interface	In order for any AR system, including SAR, to be considered a form of AR it must be interactive. Interactivity is what allows an AR system to be responsive to user input, varying based on said inputs and maintaining the AR illusion, modifying it to suit the users' needs. Without any form of interface an AR platform would be unable to respond to any type of user input. For example, a Projective SAR system that had no way of responding to user input would essentially be indistinguishable from a film projected in a cinema.
2	Projection	An SAR system must be able to display the SAR effect in some manner. Without a way of projecting the images that make up the SAR effect it would, by definition, be impossible to obtain. The projection represents the digital element, as highlighted in the Reality – Virtuality Continuum first seen in Figure 11.
3	Projection Surface	As mentioned, AR, and as a result SAR, are based on the blend of the Virtual and Real environments. Where the projection represented the Virtual environment, the projection surface represents the Real environment. Without a physical surface onto which to project, the SAR effect cannot be generated or maintained.
4	Real Time Interactivity	In order for the SAR illusion to be maintained the platform must be able to interpret any inputs and change the projection in (quasi) real time. Should too large a delay present itself between the input and the change in the projection the illusion may be broken, or the interface be considered non-functional.
5	Hardware	Though this may take many different forms, any SAR platform will require physical hardware. Be it a gantry to hold a projector or simply a PC onto which the software that runs the SAR platform is installed; hardware is a prerequisite to having a functional SAR platform
6	Software	Much like the hardware, this is a requirement as it represents the physical manifestation of the SAR platform, the software represents the "brains" of such a platform. Such software must be able to interpret the signals coming from the interface, whichever that may be, and transmit a signal to the hardware to update the projection.

4.1.2 Features of SAR

As features of SAR are being defined as the elements that can be included within an SAR platform but whose presence is not necessary for such a platform to be considered SAR, it is possible to extend this list indefinitely. As such, in the interest of brevity, to maintain the legibility, and the utility of the list for future researchers, the list of features provides top-level categories for features, as well as non-exhaustive examples of features.

Table 17 shows the top-level categories as well as the rationale for each category, followed by examples for each feature category.

Table 17. Table listing top-level feature categories of SAR and examples thereof as well as their definitions and the underlying rationale for their inclusion

FEATURE CATEGORY	EXAMPLE FEATURE	DESCRIPTION AND RATIONALE
Function	Colour Modification	An SAR platform is likely to be designed with a specific function in mind, such as supporting designers in co-creative sessions. Features integrated within the platform would then be geared towards attaining said function. For example, a platform aimed at supporting graphic designers would integrate the ability to place graphics or similar graphical assets onto the projection surface, as well as the ability to change the colour of the projection.
	Asset Placement	
Projection Support	Projector Type	As mentioned in Table 16, the presence of projectors is a fundamental requirement for SAR. However, the type of projector used, its luminosity, resolution, and colour vibrance can all change. This will often be based on the requirements being made of the SAR platform. Smaller platforms are more likely to require smaller projectors for example. The type of projector used will also influence the settings available, such as keystone correction or other.
	Projector Settings	
Tracking System	Infrared Tracking	In the interest of improving interactivity, it is possible to make the projection model manipulable by participants in the SAR session. However, as the model is moved, it is then necessary to track the movement to maintain the SAR effect. Different systems can be used to achieve this, from infrared tracking to visual tracking.
	Visual Tracking	
Projection Model	Surface Quality	A projection surface is, as discussed in Table 16, vital to the creation of an SAR platform. However, how this projection model is designed and developed can be changed. For example, the projection model can be made more reflective or matte. In addition, areas of the model can be painted in black or dark colours to reduce the visibility of the SAR effect in specific areas. Adding to this, the backdrop, behind the projection model can also be changed to suit the requirements. It may be desirable to make the backdrop as dark and matte as possible to reduce the visibility of any “spill-over”. In this way when the SAR projection does not line up perfectly with the projection model, this incongruity is better masked, and the illusion better preserved.
	Backdrop	

4.2 DESCRIPTION OF THE SAR RESEARCH PLATFORM

As previously mentioned in section 1.3, the SPARK platform’s development came about as a result of the European Union’s Horizon 2020 funding programme aimed at supporting and developing new technologies for economic development within the European Economic Area. The SPARK Consortium applied for, and received, funding for the development of a new Spatial Augmented Reality Platform aimed at supporting collaborative work and would, over the course of three years, develop and refine the SPARK platform. The exact goal of the SPARK Consortium was to develop a: **“Responsive ICT platform that exploits the potential of Spatial Augmented Reality for supporting and fostering collaborative creative thinking”**(SPARK Consortium, 2015).

The static SPARK platform was developed to be able to support the implementation of a room-scale SAR system (Caruso *et al.*, 2016a). With this technology, clients are able to participate in an SAR supported design session when visiting the design agency to discuss any projects or work. As such the projectors selected, were chosen with room-scale sizes in mind. In addition, the use of the Information system was based on the decision to have a centralised asset managing system that would allow multiple teams to view and review design sessions. It is important to note that assets, when used in the context of the SAR platform, refer to the graphics that are projected onto the physical model. These assets can be logos, textures, or other images used to represent specific elements of the design. Furthermore, the Information System could, theoretically, allow for a design session to be viewed remotely or for multiple SAR sessions to be coordinated.

However, as the project progressed and new knowledge was gained, it became apparent that a single system was not well suited to achieving all the goals as set out by the project. As discussed in further detail in sections 4.1 and 4.3, a portable system was also developed, branching off from the larger room-size system described in this section. The development of the portable SAR research platform was conducted at the University of Bath. It is important to note that, as the portable system was an offshoot of the static SPARK platform, they shared a number of similarities, in particular where the software and information system (Figure 32) were concerned.

Figure 32 provides an overview of the SAR platform configuration used by both the static and portable systems. The platform consists of three sub-systems: the Information System and Software sub-systems, and Projection Hardware. Additional details for each of the subsystems are provided in sections 4.2.1, 4.2.3, and 4.2.3 respectively.

The Information System acts as the “brains” of the platform. It stores all the data required to run the session. All the assets that can be selected need to be pre-uploaded to the Information System prior to the start of the session. The same is true for the digital models of the physical prototypes. In addition, the Information System enables the initial scene to be setup and the design session to be initiated as well as recorded.

The Projection Hardware consists of all the projectors, mounting brackets, and cables required to support the projection. Furthermore, the Projection Hardware includes the infrared tracking cameras used to track the movement of the physical projection model in 3D space. Lastly, the projection hardware also refers to the projection model, onto which the images are projected to achieve the SAR effect.

Finally, the Software refers to the system, including the GUI, used by the designers to interact with the Information System and the Projection Hardware. The Software enables the initial calibration that aligns the digital projection of the model onto the physical model itself, storing this data for further sessions. Furthermore, the Software holds additional information for the calibration of the tracking system. Additionally, the Software also records certain behaviours (such as placement, translation, scale, rotation and deletion of assets, the level of zoom, and the background colour). Lastly, the Software acts as the GUI during the design session, allowing the designer to make the desired changes to the projection.

The Information System and the Projection Hardware are in constant communication with the Software, which coordinates the overall system and allows the SAR illusion to be created and maintained. It should be noted that the Information System need not be in constant communication with the Software. Once the initial setup is complete and the software has downloaded all the assets and models necessary to setup the scene from the Information System the communication between the two can be severed although this results in a loss of some functionalities. A more detailed description of the three subsystems is provided below.

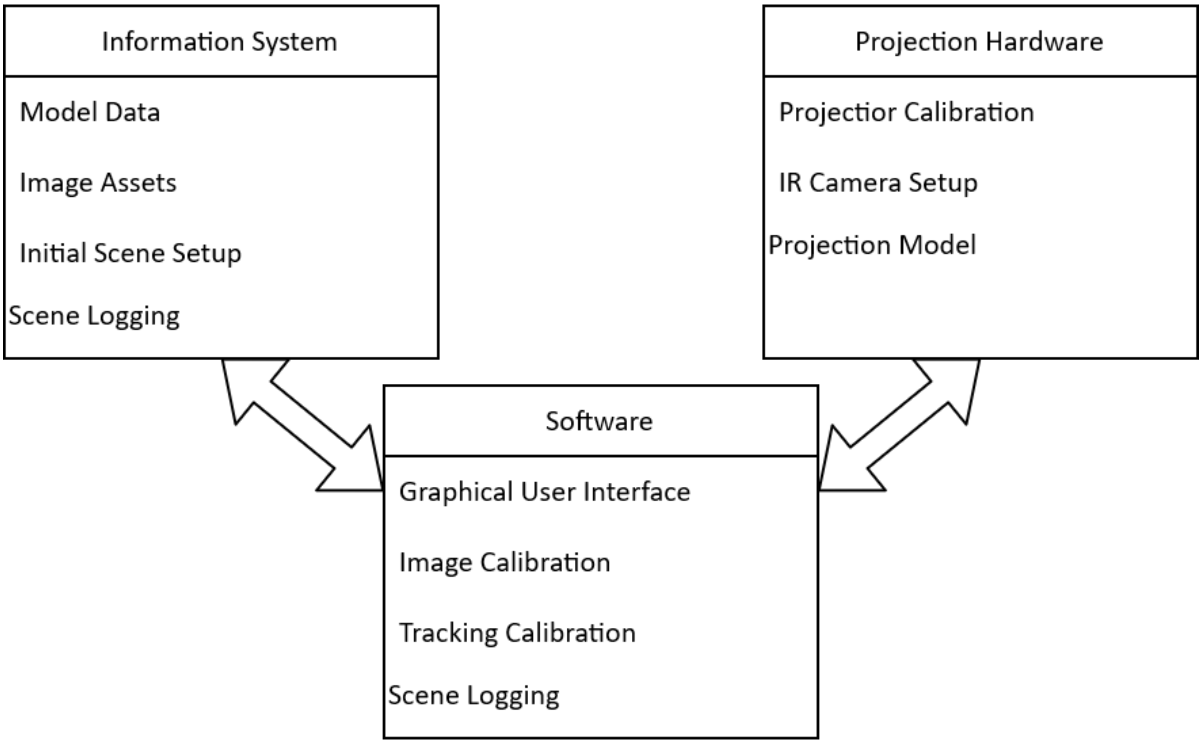


Figure 32. The SAR research platform System Architecture

Figure 33 shows the components that make up the SAR research platform and how each component communicates with the others to allow the SAR supported designs session to be run. As can be seen, the lynchpin of the system is the PC that runs the SPARK software. This allows it to communicate with the Information System's server, thus retrieving the necessary 3D models and assets. In addition, the PC communicates with the tablet being used by the designer as an input device via a router and with the IR cameras which constantly relay the position of the mixed prototype. All this information allows the PC to update the projection to match the inputs from the tablet and keep the projection centred on the mixed prototype, maintaining the SAR effect.

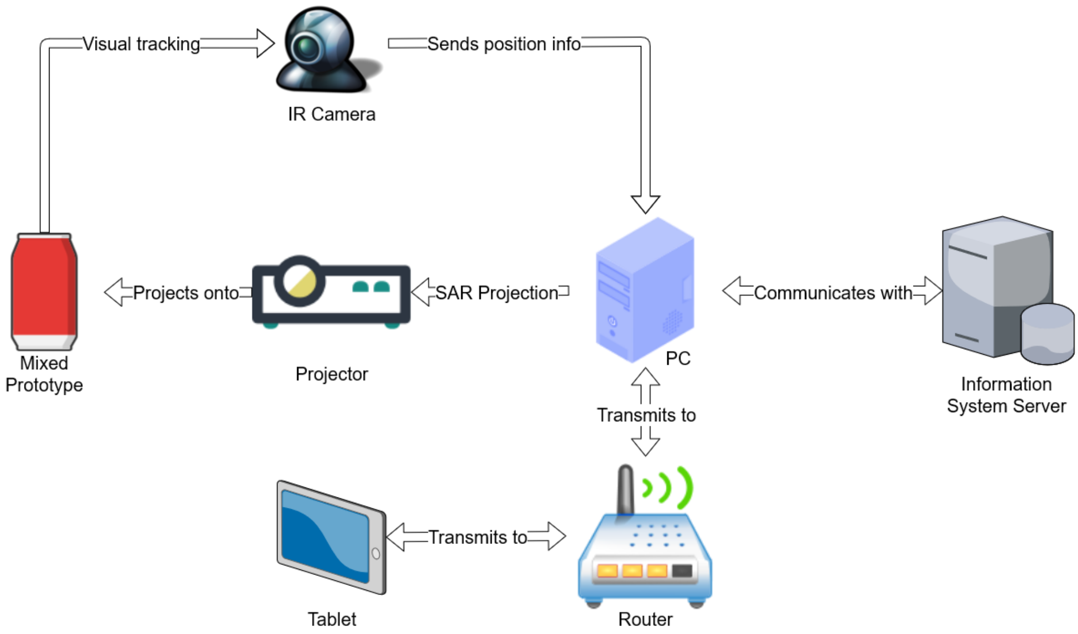


Figure 33. General overview of SPARK platform components and the communication between them

4.2.1 Information System

The Information System (IS) is a cloud-based database and server for storing and accessing the digital assets to run an SAR co-design session. Examples of the digital assets include: logos, textures, sprites, 3D models and their respective UV maps and scene configurations. This UV map tells the system how to unwrap the object and thus how to project images onto the 3D model. The initial scene setup allows the designers to start the design session with some pre-set assets already laid out onto the 3D model rather than having a blank canvas.

The letters "UV" do not refer to ultraviolet light, rather, they are used to refer to the X and Y coordinates of the 2D plane where texture is placed. The letters U and V are used instead of X and Y as these are already being used to define the XYZ coordinates of the 3D environment. As such, to highlight the fact that the two coordinate systems refer to different spaces, the X and Y coordinates of a 2D plane (such as that used in a texture), when used in a 3D environment, are called U and V respectively.

The process for interacting with the Information System is shown in Figure 34. The IS features four main screens used by the user during the setup of a co-design session. Users log into the IS, create a new session, upload a 3D model, upload the desired assets, and then modify the initial scene by adding the assets to the model. Once the user is satisfied, the session can be made live. This enables the SAR Software, first described in Figure 32, to select and use the created session. On selecting the session, the client software, which is running on the PC shown in Figure 33, downloads the required assets and runs the session. Figure 34 shows a brief overview of each of these screens as they appear to the user, showing how a session is created and managed in the IS prior to being downloaded for use by the software.

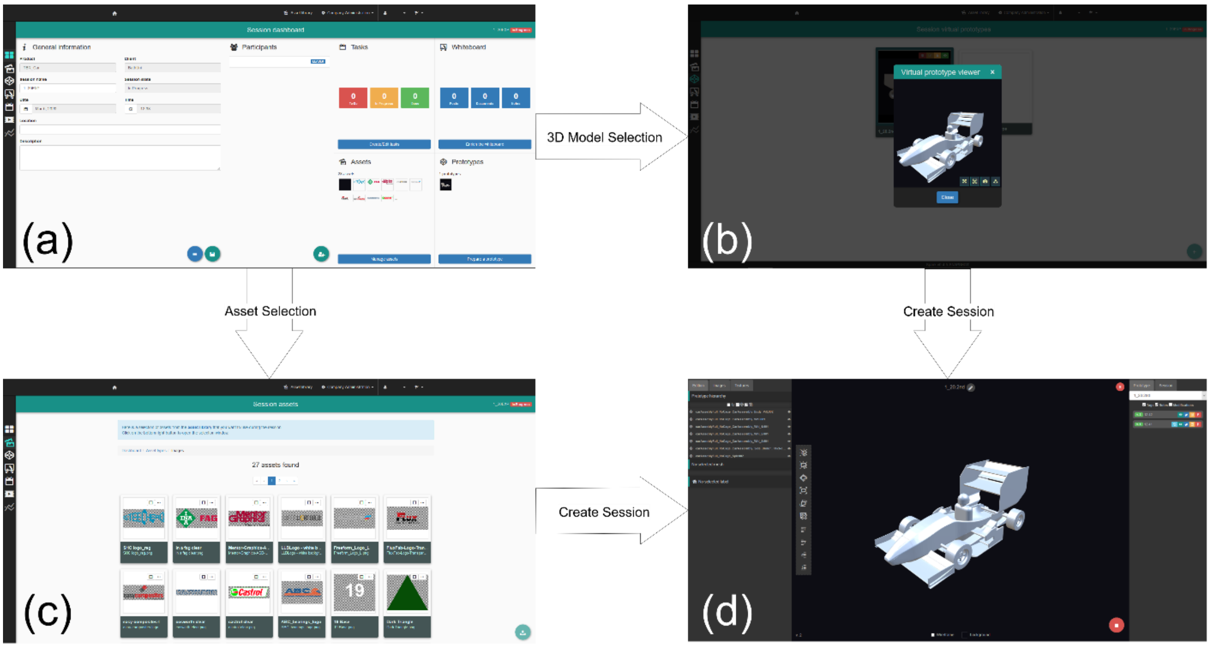


Figure 34. IS setup. (a) shows the initial splash-screen where all the general session information is shown and directs users to the 3D model and Asset Selection screens. (b) shows the 3D model selection. A 3D model is uploaded to the IS and is being previewed before being implemented in the session. (c) shows the asset selection screen. Here, the desired assets for the session are uploaded and can be previewed before being implemented. (d) shows the session creation screen. Here, the initial scene setup can be set, and the session made live once ready.

In addition, the IS logs the edits made to the scene as well as allowing remote viewing of the edits in real time. The log made by the IS tracks any and all committed changes to the scene enabling the research team to track the design process and design iterations. Participants can also tag the scene at any point in time thereby flagging that moment as a significant event (e.g., a final design concept). This can be seen in Figure 34(d) where the column on the right of the scene shows two iterations of the model. In addition to logging the edits, this feature also allows users to roll back and revisit past scenes (un-do functionality).

Additionally, the IS allows the users to review and replay a co-design session. These features are only enabled when the client SPARK software is in constant communication with the IS during the design session. This can be disabled to improve stability and decrease latency and allows the client to work where no internet connection is available.

4.2.2 Client Software

The client software serves as the nerve centre of the whole SAR setup acting as the local controller that enables the users to interact with the SAR projection. The software client was developed using the Unity game engine and consists of a Graphical User Interface (GUI), a calibration tool, and a logging system.

The GUI allows the users to modify the projection (and thus the SAR effect viewed) and is shown in Figure 35. The calibration tool allows the users to calibrate the overlays that are projected onto the physical model. The calibration tool requires the number of projectors to be set (and their overlap determined) and the tracking system (the IR cameras mentioned in Figure 33) to be calibrated and engaged (if desired). The tracking system allows the participants of the design session to interact with the model by translating and rotating the model, whilst the projection updates in real time to maintain the SAR illusion in place on the projection model.

Once a session is created and made live in the IS, the client software downloads all the assets, 3D models, and initial scene layout setup to begin the calibration. The digital model and the physical model are aligned during this calibration phase. Once all the calibration is completed the design session can be initialised and the GUI launched.

Figure 35. Early version of the Graphical User Interface (Giunta, 2017)

The logging system captures the user’s interaction during the session by recording which functions are selected, assets used, colours selected, movement of the digital model in the GUI, and other options (translation and rotation of assets, timestamps, user tags added, unique instances of assets). The logging system records the information in two ways. As part of the runtime the software takes a “screenshot” of all the assets, colour, camera position, etc. of the scene. Screenshots are periodically captured as the software loops within its runtime. Furthermore, when the user interacts with the interface, such as by pressing a button, it triggers an event that is also recorded in the same log file as a new line. All this information is stored locally on the PC hosting the session. Unlike the data sent to the IS, the log kept by the software tracks the individual interactions with the interface rather than recreating a twin of the current state of the session on a separate server.

Each new line in the log contains information identifying the session, the time since the start of the session, the system time. The log file also captures three main subcategories of information. These are: Activity, Saved Version, and Tracking. Each new line stored in the log file will capture information pertaining to each of these subcategories. The different categories are expanded upon below.

The Activity sub-category refers to the information pertaining to the user’s interactions with the GUI. The behaviour is further categorised into four sub-sub-categories: Selection and Manipulation of Assets, Activities Relating to the Use of the Interface, Change of Background Colour, and Change of Visualization or Viewpoint. Selection and Manipulation of Assets refers to the interactions the user has with the assets and tracks the placement, rotation, scaling, and layer order of these assets as they are changed by the user. When the user places a new asset into the scene, the asset is given a unique instance identifier alongside its asset identifier. This allows multiple assets of the same type to be tracked whether they appear concurrently or subsequently to one another. Activities Relating to the Use of the Interface tracks the user’s behaviour by recording how the assets are tagged, filtered using those tags, how assets are swapped for one another, and how assets are deleted. The Change of Background Colour tracks how the background colour of the model, or section of the model, is changed and the colour to which it is changed. Lastly the data captured as Change of Visualization or Viewpoint refers to how the user pans around the model in the user interface, as well as how they zoom in and out of the model. Furthermore, the selection of specific sections of the model is tracked through this.

The Saved Version sub-category records a snapshot of the system. This means that all the information regarding the Activity and the Tracking is saved immediately as a new line in the log file.

Lastly, the Tracking sub-category stores all the information pertaining to the physical model's position in space. Where the SAR system used on infrared tracking this meant the X, Y, and Z coordinates of the physical model as well as the Euler angles of the infrared markers. If the SAR system only used rotational tracking, then only the rotation about the Z axis was recorded.

4.2.3 Projection Hardware

The Projection Hardware consists of the projectors, the infrared (IR) cameras, and the physical projection model (mixed prototype). The projectors (one or more depending on the desired coverage for the physical model) need to be calibrated to provide the required projection envelop, focal length and focus. The IR cameras used to track the physical model need to be calibrated to follow the model and transmit the data to the client software. Lastly, the physical model itself needs to be prepared. Figure 36 shows a number of configurations for the SPARK platform hardware, as deployed at GINP, POLIMI, Artefice, and Stimulo. Figure 37 expands on Figure 36, showing a photo of the SPARK platform hardware as deployed at POLIMI. Not shown in either figure is the computer that runs the SPARK client software and co-ordinates the tablets with the projectors.

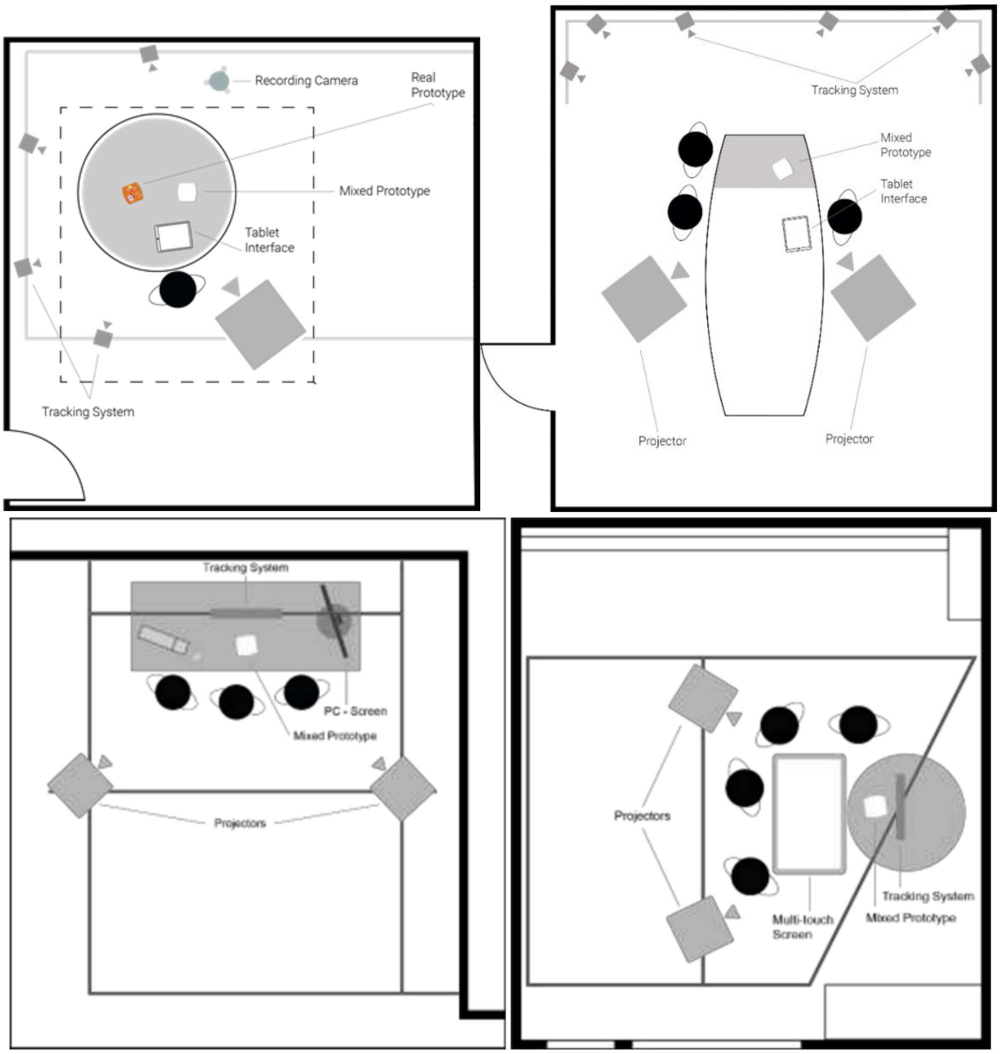


Figure 36. Multiple setups of the SPARK Hardware at various locations. While the hardware remains mostly unchanged between deployments the layout is adapted to fit the location. Clockwise from the top left the locations are: GINP, POLIMI, Artefice, Stimulo (Bellucci *et al.*, 2018a; Morosi *et al.*, 2018b)

Due to the modular nature of the SAR platform, there is no fixed way of setting the system up. Depending on the desired features of the co-design session, the platform can be setup with one or more projectors and make use of the IR tracking or not. Furthermore, the projectors and IR cameras can be placed in different layouts to optimise the projection envelope depending on the layout of the room. This was a key requirement for the SPARK project as the platform had to be deployed in different locations. Regardless of the location however, the functionalities enabled were the same across all tests conducted. Figure 37 shows the platform deployed at the Politechnico di Milano SPARK room used for some of the experiments. A stylised layout of the room is shown in Figure 36.

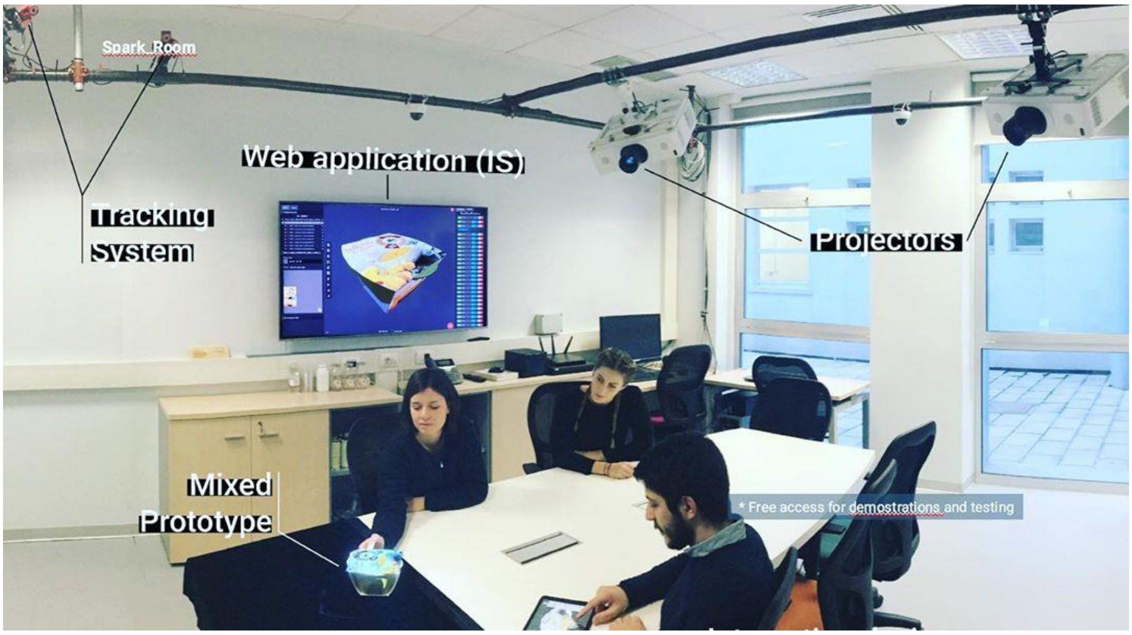


Figure 37.Picture of the setup shown in Figure 36 (SPARK Consortium, 2019b)

The mixed prototype itself must be built to scale to allow for the digital model to accurately map onto it during the software calibration phase. That is to say, the proportions of the digital model must match those of the physical prototype, for example, if a physical prototype is twice as tall as it is wide the digital model must reflect this as well. In addition, the model must be painted or otherwise coloured in such a way as to reflect light well enough for the SAR effect to be clearly visible.

During the development of the platform different surfaces were analysed. The approach was iterative where different types of surfaces and paints were used, and their effect was qualitatively assessed to see which gave the best projection performance. The most common type of coating used was white paint, with a fairly matte finish. The matte finish was found to be best suited as it gave a more “real” and “textured” feel to the physical prototype allowing the projection to look more realistic. However, in particular where the material being simulated was inherently glossy, a glossier, more reflective paint was found to be more suited. In general, as long as the surface was reflective and not too dark the SAR effect was achieved.

Indeed, different applications called for different approaches. The projection surface of the prototype shown in Figure 37, for example, is made of paper. As a result, it was unnecessary to apply any form of paint or colour as paper was found to be perfectly suitable for projection. Other prototypes, in particular those which were 3D printed, did not necessarily require much post-processing either. By printing using white or matte white filament it was possible to obtain a good projective surface. However, 3D printed prototypes did require post processing to remove any blemishes from the printing process.

Of the paints analysed during the development of the platform it soon appeared that matte paint enabled a diffuse reflection of light which aided in giving the physical prototypes a more realistic look. However, depending on the desired results, a glossier paint could also be used to increase the sheen of the mixed prototype. White is often used as the base colour as it will reflect all light well.

Again, depending on the application, a different colour may be desirable, for example, painting a feature in black to deliberately hide a specific area and make any projections there harder to see.

Lastly the IR trackers must be placed on the physical model to allow the IR cameras to follow the model as it is moved around. Figure 38 shows three implementations. The constellation of IR markers used for tracking also needs to be placed in such a way as to not be too obtrusive or obstructive for the projectors, yet still visible enough for the cameras to track, even when held by the design session participants. Later iterations of the IR trackers would improve on the constellations' obtrusiveness, reducing the risk of occlusions caused by the constellation. This was achieved, in part, through the use of IR markers placed directly on the model rather than as a constellation attached separately.

Figure 38. Three mixed prototypes. All have been painted matte white with paint and have been fitted with a "constellation" of IR trackers (SPARK Consortium, 2018d, 2019a)

4.3 DEVELOPMENT OF THE PORTABLE SAR SYSTEM

The portable SAR system, referred to as miniSPARK, was the second SAR system that was built to meet the requirements that could not be achieved in the first SAR system (e.g., portability). The decision to branch off from the base SPARK platform was in response to the feedback that was received from industrial partners of the SPARK Consortium (Stimulo and Artefice). This feedback highlighted how the intended use case for the SPARK Platform did not match the typical cases faced by the design agencies.

Furthermore, the industrial partners noted that the cost of the platform would prove prohibitive if the Consortium wished to commercialise it as small to medium design consultancies (such as Stimulo and Artefice) would not normally have been willing to part with the funds required to implement the SPARK platform had it not been for the support of the Horizon 2020 grant. While the cost of the platform did not play a large role in the development of the portable SAR system, it did remain a consideration during the design and development process.

4.3.1 Rationale

The original proposal, during the initial development of the SPARK platform, was that the industrial partners would have a SPARK platform installed at their location and use this to work with their clients. This would be consistent with the deployment of the other SAR platforms installed at the locations of the academic partners of the SPARK Consortium.

However, as the development of the SPARK platform progressed, the industrial partners revealed that this approach was a break with their standard operating procedure. Normally, when setting up a collaborative design session with one of their clients, both Stimulo and Artefice would travel to their clients' location and work with them there. The SPARK platform's immobility made this impossible: meaning that the clients would have to travel to them.

In addition to these discoveries, it quickly became apparent that setting up a SPARK platform at the University of Bath, as per the Consortium's agreement for academic testing of the platform, was proving to be needlessly complex due to a number of limitations (amongst which was a lack of available space to setup the platform).

Whilst it would certainly have been possible to work around these constraints, the decision was made to take the opportunity to explore the feasibility of a smaller "miniSPARK" system to address some of the concerns raised by the industrial partners.

The author sought, and obtained, permission from the other members of the consortium to explore this avenue. In doing so, the development of the miniSPARK remained linked to the larger SPARK platform, continuing to enable cross-comparison of data collected in experiments. However, it also enabled the research to address the research questions and objectives of this thesis in a more refined manner. The miniSPARK platform served the dual purpose of not only supporting the planned SPARK experiments, detailed in chapters 5 and 6, but also serving as an iterative platform to explore how the platform itself would influence designers. This more explorative approach led to a more iterative design of the miniSPARK platform when compared to its larger sibling.

Figure 39 shows some concept art created to help promote miniSPARK concept within the SPARK Consortium as well as to aid with storyboarding.

Figure 39. Concept art created to promote miniSPARK within the SPARK Consortium (SPARK Consortium, 2018c)

4.3.2 Technical Challenges to Meet the Requirements

The development of a miniaturised rig was done to address some of the challenges evidenced by the initial rollout and implementation of the SAR platform discussed in section 4.3.1. It thus is of interest to briefly explore some of the technical limitations that arise when using projectors, in particular as these problems are exacerbated by the shorter distance between the projector and projection surface. Indeed, when the distance between the projector and the projected surface is great enough, it can be safely assumed that the light arrives at the projection surface perpendicularly. However, as the distance decreases, this assumption becomes less true. As a result, some of the effects discussed in the subsequent paragraphs are of particular importance, as they play a role in maintaining the projection quality. Three topics relating to projector layout are of particular interest: Focus and Depth of Field, The Keystone Effect, and Resolution. All these technical challenges needed to be addressed in order to successfully achieve a functional miniaturised and portable SAR platform.

4.3.2.1 Focus and Depth of Field

Figure 40 shows a convex lens converging two light beams into a single point. This point (F) is known as the focal point. The distance from the lens to this focal point is known as the focal length (f). In order to achieve a sharp image when projecting, it is necessary for F to coincide with the projection surface.

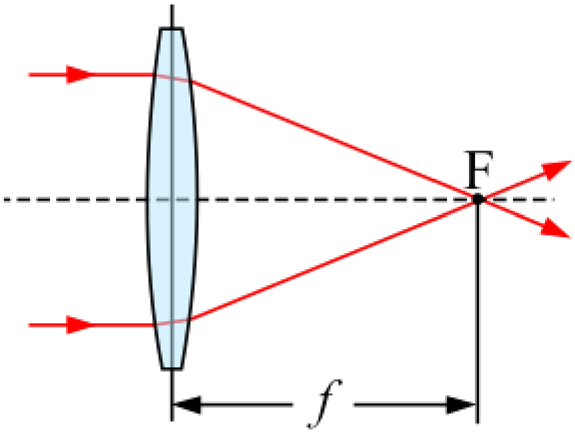


Figure 40. Illustration showing the focal point (F) and focal length (f) of a convex lens (Henrik, 2020)

When projecting a single point onto a plane, the angle of the plane to the lens does not matter. However, when projecting multiple points, as one would when projecting an image, the angle of the plane relative to the lens begins to have an impact on the ability to focus the image. Figure 41 shows this effect in action.

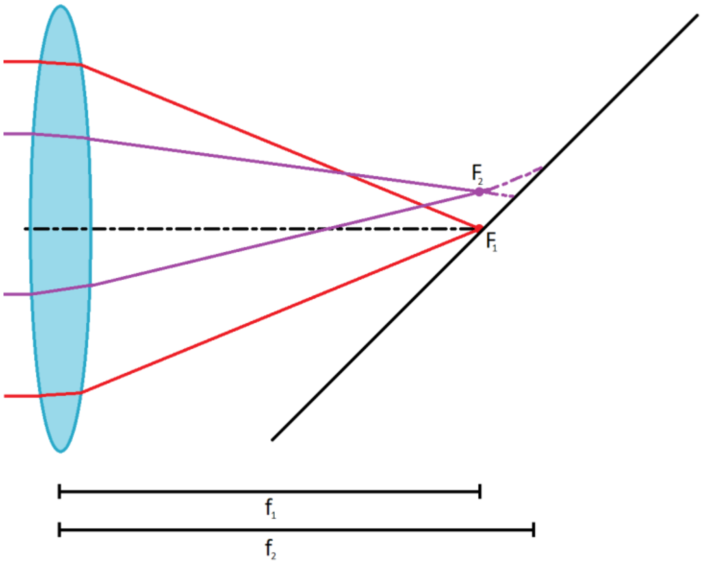


Figure 41. Two sets of light beams being projected onto a slanted surface. The red beams hit the surface at the correct focal length (f_1) and will thus appear in focus. The purple beams need to travel farther to hit the surface (f_2), as such the focal point F_2 is not coincident with the plane and will appear to be out of focus

Figure 42 shows a practical application of the theory described in Figure 41. As can be seen, the text is only partially in focus. As the page was photographed whilst slanted, the parts too close and too far from the camera lens are out of focus. However, the middle segment is in focus. As can be seen, the area that is in focus is not merely a line but a spectrum, this is known as the Depth of Field. As a human eye has a tolerance for focus, images that are not exactly at the focal length may still appear acceptably in focus, but as one moves away from the focal point, increasing or decreasing the distance, the image appears increasingly out of focus.

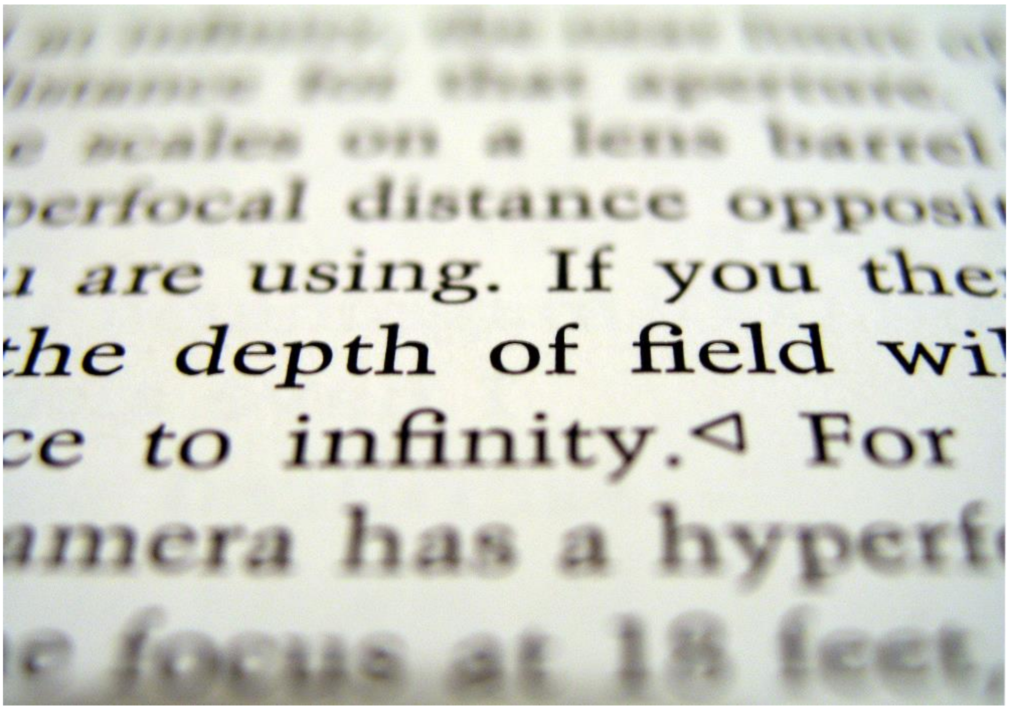

e scales on a lens barrel
perfocal distance opposi
u are using. If you the
the depth of field wi
ce to infinity.◁ For
amera has a hyperf
e focus at 18 feet,

Figure 42. Depth of field effect (Ligar, 2005)

4.3.2.2 *Keystone Effect*

Figure 43 and Figure 44show the keystone effect in action. This effect occurs when the projector is placed at an angle relative to the projection surface; by rotating the screen along its y-axis, as shown in Figure 44 (a horizontal rotation), or along its z-axis (a vertical rotation). In Figure 43 the projector is placed perfectly perpendicular to the screen and the projected image appears rectangular. Figure 44 shows the keystone effect due to the angle of the screen relative to the projector: the previously rectangular image is now shown as being trapezoidal. This distortion occurs due to the different distances that light has to travel to reach the screen and can lead to a series of difficulties when projecting, such as distorted images or poor resolution. It is possible to correct for this effect both manually and digitally. Manual correction involves moving the projector or the projection surface to guarantee they are perpendicular. Digital correction works by projecting a “pre-distorted” image; by enlarging the shorter side of the trapezoid and shrinking the larger one as a digital process before projection, thus the image will appear correctly on the screen.

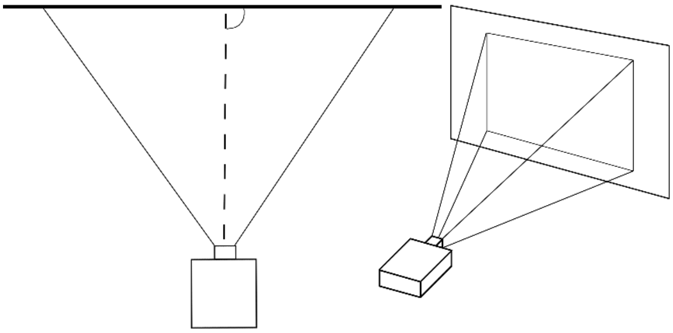


Figure 43. A projector projecting onto a perpendicular surface

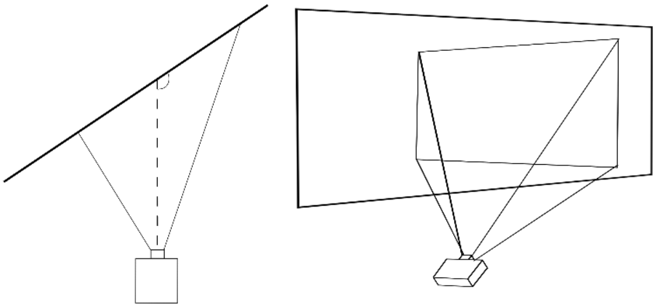


Figure 44. A projector projecting onto a surface at an angle

While it is possible to correct for the keystone effect, this fails to account for issues in focus. While the focus of most projectors can be adjusted, a specific focal point must be selected. In the case of a slanted projection surface the distance between the focal point and the actual area being projected on can lead to the extremities of the image being out of focus.

4.3.2.3 *Resolution*

Lastly, the issue of resolution comes into play. By attempting to project onto slanted surfaces warping begins to occur. This is both because pixels become stretched and compressed on the far and near side respectively. The use of digital keystone correction exacerbates the issue as the type of stretching used in the transformation relies on the modification of rasteried graphics which ultimately result in resolution errors.

4.3.3 First Iteration

The seeds for the development of miniSPARK had thus been sown. The stated intention behind the development of the miniSPARK platform was thus to provide a smaller, cheaper, and more portable alternative to the main SPARK platform. The miniSPARK platform would be expected to provide most, but not all, the functionalities of the main SPARK platform to achieve this. In addition, cross platform functionality was promoted, to allow the users of the main SPARK platform the ability to migrate to miniSPARK, and vice versa, as required. To this end the software and the Information System remained unchanged between the main SPARK platform and miniSPARK; only the hardware was modified to attain a more portable setup. The outline for this first iteration of the miniSPARK platform is shown in Figure 45.

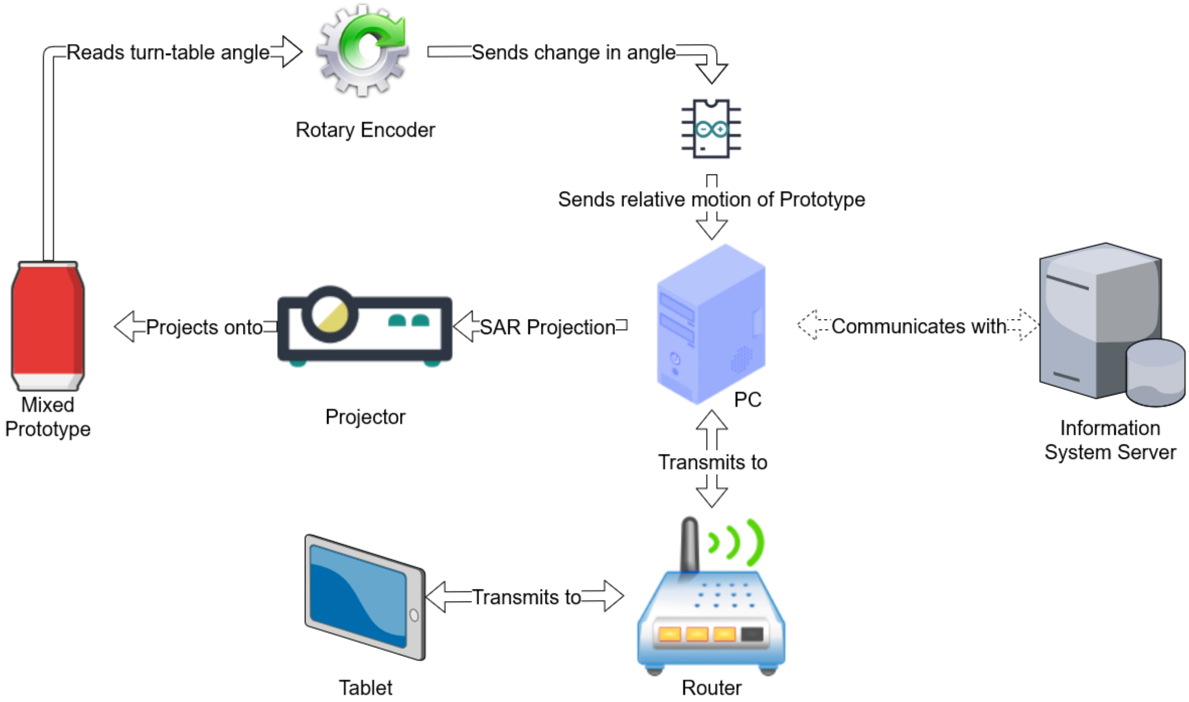
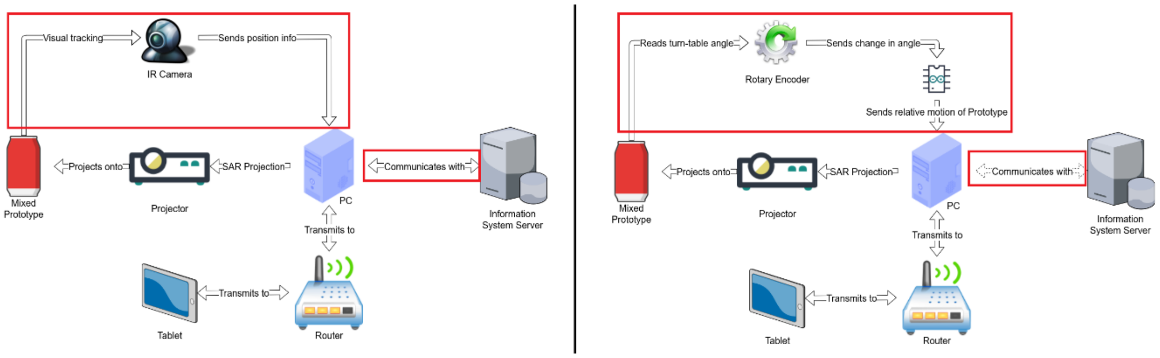


Figure 45. Overview of the initial miniSPARK platform components and the communication between them

Figure 47 shows one of the first functional iterations of miniSPARK on display at an exhibition in Belgium. Unlike its much larger sibling the entire system fits onto a single table. Much like the regular SPARK platform, miniSPARK still boasts a tablet interface and a projector as well as a mixed prototype. Figure 46 highlights the differences between the larger SPARK and miniSPARK.

As can be seen from Figure 46, the visual tracking system that made use of infrared cameras has been replaced with a rotary encoder and an Arduino. This results in a system which can only track the rotational position of the mixed prototype, sacrificing the ability to track translation. This sacrifices the participants’ ability to manipulate the physical model in 3D space, allowing them only to rotate the model. However, this rotational tracking is simpler to implement, does not rely on line of sight between tracking cameras and infrared markers, and allows the system to become portable.

In addition, the communication with the Information system has been reduced. In the larger SPARK setup, this was always on and the communication between the two remained on throughout the entire session. With miniSPARK, this was changed to only allow the communication at the session start-up. This enabled the sessions to be run even where an internet connection was unavailable, as well as improved the system's stability as the communication with the Information System was found to be a frequent point of failure.



Larger SPARK **miniSPARK**

Figure 46. Side by side comparison of the differences between the layouts shown in Figure 33 and Figure 45. The differences between the two setups are highlighted in red.

The configuration is thus very similar to that shown in Figure 36. Not shown in Figure 47 is the PC. This was stored underneath the table to reduce clutter during the exhibition but was still running the SAR projection. Three key differences exist, however. Firstly, miniSPARK's projectors are much smaller than those shown in the setup in Figure 37. To increase the portability of the system pico-projectors were used, greatly decreasing the weight of the application but sacrificing luminosity and resolution. Furthermore, the projectors, instead of being mounted on the ceiling, have their own armature holding them in place. The second key difference lies in the tracking mechanism used. While the setup shown in Figure 36 utilises multiple infrared cameras to track the movement of the mixed prototype in 3D space, miniSPARK can only track the movement of the object as it rotates along its z-axis. The pedestal the mixed prototype is resting on in Figure 47 can rotate along this axis and contains an optical encoder. This encoder generates pulses as the pedestal is turned; these are sent to an Arduino which converts the electrical pulses into values that the computer can read. This informs the computer how many degrees the platform has rotated, either clockwise or counterclockwise, thus allowing the projection to be adjusted accordingly to match the movement of the mixed prototype. The last major difference between the SPARK and miniSPARK platforms is in the way the Information System is used. Unlike the larger system, the standard operating procedure for miniSPARK is to only use the Information System to load the session. Thereafter the link to the Information System is disconnected and miniSPARK operates only locally.

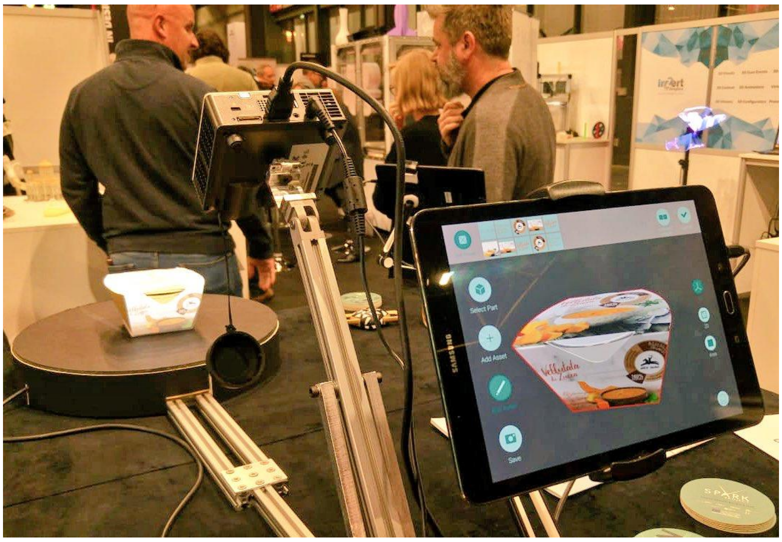


Figure 47.The first iteration of miniSPARK on display at the Prototyping 18 exhibition in Kortrijk, Belgium (SPARK Consortium, 2018b)

The first iteration of the miniSPARK platform had proven the concept as viable. However, after the first few deployments it quickly became apparent that this first version of the miniSPARK platform had some shortcomings. In addition to the immediate issues detected as part of the initial rollout of miniSPARK, the feedback and input of the industrial partners, Artefice and Stimulo, was sought. The main concerns raised are collected in Table 18.

Table 18. Issues Identified with first iteration of miniSPARK Platform

NR	ISSUE	DESCRIPTION
1	**PC TO TABLET CONNECTION.**	The connection between the PC and the Tablet was unstable and a common point of failure, in addition to being difficult (package drops and synchronising issues) to set up.
2	**PC SIZE.**	The PC used to run the SPARK software, often just a laptop, was large and unwieldly taking up valuable table space.
3	**ARMATURE MOUNTING POINTS.**	Only two mounting points were present to attach the projector armature to the pedestal/turntable.
4	**ARMATURE STABILITY.**	The armature holding the projectors in place tended to fail, leading to the projectors drifting. This was especially prevalent when the projector was placed in the far forward or far backward positions.
5	**ARMATURE DEGREES OF FREEDOM.**	Armature also lacked degrees of freedom, making the initial calibration phase more complex and difficult.
6	**TURNTABLE BUILD QUALITY.**	The turntable was 3D printed due to its unusual shape. Unfortunately, due to its size, the turntable had to be printed in segments. The combination of the inaccuracies innate to the 3D printing process and their impact on the assembly caused the turntable to not be perfectly stable. This problem was exacerbated by the light weight of the turntable; while desirable for transportation, the low weight meant the turntable tended to rock or move when people touched it.
7	**TURNTABLE ROTATION.**	The top part of the turntable, where the mixed prototype is placed, turned with difficulty. This was due, in part, to the poor tolerances inherent to the 3D printing process. One other cause was the bearing that attached the top of the turntable to the base.
8	**ARDUINO.**	The Arduino, used to convert the pulses from the encoder to generate code legible by the PC, was difficult to manage. While the task it performed was vital, the Arduino was large and took up considerable space as it required its own casing to protect it during use and transport.
9	**CABLE MANAGEMENT.**	The total number of cables was considerable and managing them proved to be a source of frustration and annoyance, especially during packing and deployment of the system.

4.3.4 Second Iteration

To address the issues identified and described in Table 18, an improved version of miniSPARK was developed. Figure 48 shows the general layout of the improved system.

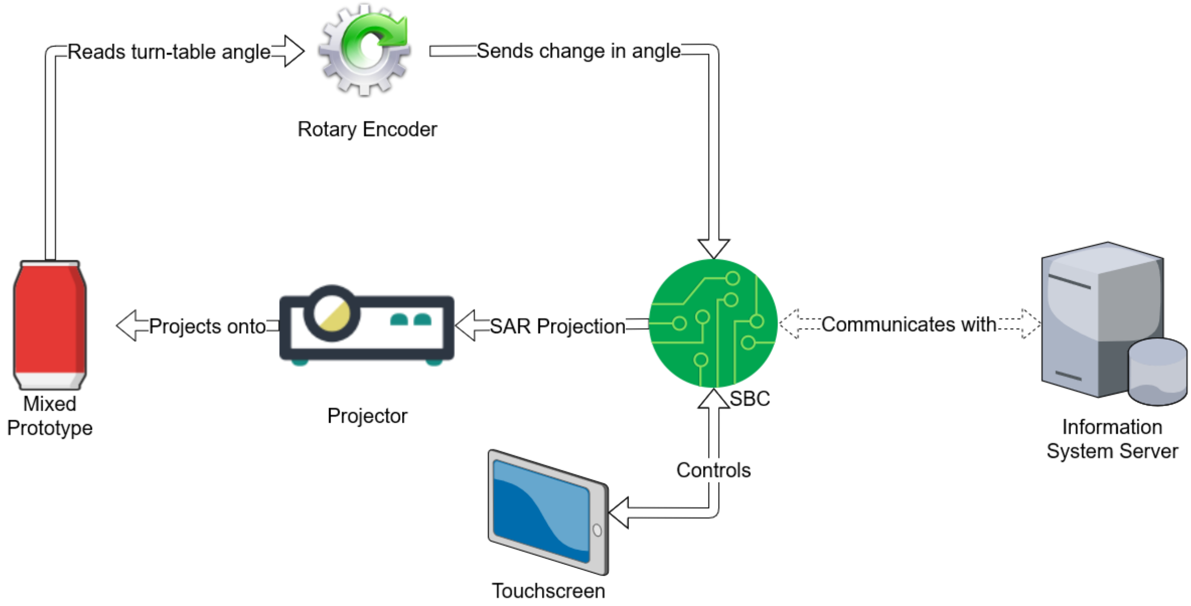


Figure 48. Overview of the second iteration of the miniSPARK platform components and the communication between them

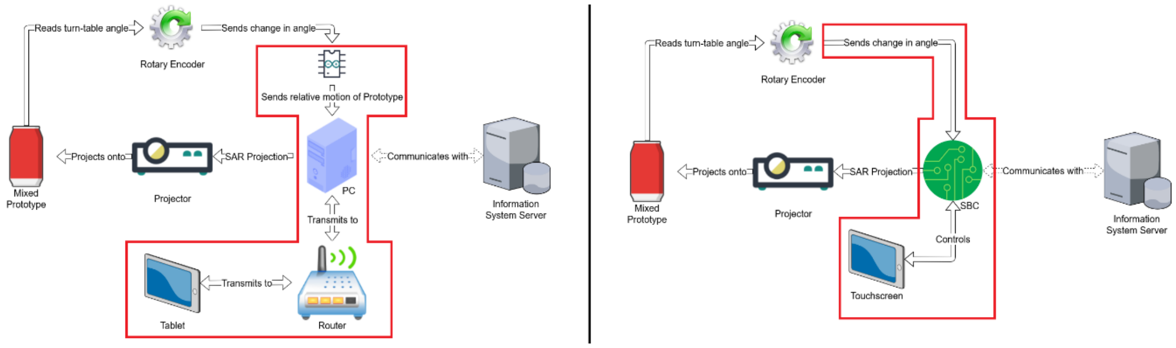
Figure 49 shows the two iterations of the miniSPARK platforms side by side. In order to improve clarity, the figures omit the cables required to connect all the hardware together. Both iterations consist of a turntable, a projector, a projector stand, and a mixed prototype. As can be seen from the figure, the second iteration of miniSPARK (Figure 49b) integrates most of the hardware inside the turntable, with only a touchscreen extending therefrom.

(a) **(b)**

Figure 49. First (left) and second (right) iterations of the miniSPARK platforms

However, the first iteration of the miniSPARK platform (Figure 49a) has most of the hardware outside the turntable. This results in a more complex setup as well as an increase in the number of cables required to power all the components. Lastly, as shown, the first iteration of miniSPARK uses a tablet rather than a connected touchscreen. In order to connect this to the computer shown, both need to connect to the same router (which also requires an internet connection).

Figure 50 expands on the details provided in Figure 49, by showing that the system diagram of both iterations of miniSPARK (first seen in Figure 45 and Figure 48 for the first and second version respectively). The red boxes highlight the components that have been changed across the two iterations. The PC was replaced with a Single Board Computer (SBC), eliminating both the need for an Arduino to interpret the signal from the rotary encoder as well as eliminating the need for a router to communicate with a tablet to act as an interface. The tablet was replaced with a touchscreen connected directly to the SBC.



(a) **(b)**

Figure 50. System diagrams of the first (left) and second (right) iterations of the miniSPARK platforms as shown in Figure 45 and Figure 48 respectively. Differences are highlighted in red.

The following changes, shown in Table 19, were implemented in this second version to address the previously identified problems.

Table 19. Improvements made by second iteration of miniSPARK addressing issues identified in Table 18

NR	ISSUE	IMPROVEMENT(S) MADE
1	PC TO TABLET CONNECTION.	The tablet was replaced with a touchscreen. This eliminated the need for a router that both the PC and the tablet needed to be connected to and instead allowed the touchscreen to be connected directly to the PC by means of a single cable.
2	PC SIZE.	The PC was replaced with an SBC. This greatly decreased the computer's footprint allowing the computer to be mounted directly inside the turntable, further decreasing clutter.
3	ARMATURE MOUNTING POINTS.	The method used to attach the armature to the turntable was redesigned providing four mounting points.
4	ARMATURE STABILITY.	The armature was redesigned to improve stability by adding more support points.
5	ARMATURE DEGREES OF FREEDOM.	The armature was improved by adding one additional degree of freedom
6	TURNTABLE BUILD QUALITY.	The turntable was manufactured out of aluminium increasing the tolerances. The added weight also eliminated the stability issues the turntable had previously faced.
7	TURNTABLE ROTATION.	A different bearing was used that went all around the edge of the turntable, greatly reducing friction. Furthermore, the platform was equipped with an electric motor to turn the turntable remotely.
8	ARDUINO.	The SBC used to replace the PC had an integrated Arduino, meaning that the need for a separate Arduino, and its cables, was eliminated. Furthermore, as the SBC was placed directly next to the encoder, the two could be hardwired together, greatly reducing issues with cabling.
9	CABLE MANAGEMENT.	The number of cables used was reduced. The elimination of the router, the integration of the Arduino into the SBC, and the placement of the SBC inside the turntable, helped to greatly reduce the number of cables needed.

4.4 SAR PLATFORMS' SUITABILITY IN ADDRESSING RESEARCH QUESTIONS AND OBJECTIVES

Previously, in sections 3.4.2 and 3.4.3, the research questions and research objectives to be answered and addressed within this thesis were laid out. This chapter has, in section 4.1, provided an overview of the requirements for a functional SAR platform. Subsequently, section 4.2, discussed an implementation of an SAR research platform to be used in the subsequent research and addressing the needs previously laid out. Section 4.3 built on this SAR platform, proposing a miniaturised portable SAR system which aims to address some of the shortcomings detected in the original SAR platform.

By cross referencing the research questions and the research objectives to the characteristics and features of the platform discussed in sections 4.2 and 4.3, it is possible to better understand how the research platform may need to be adjusted or re-evaluated. This is in order to address the specific needs of this thesis rather than the goals of the SPARK Project, for which the platform described in section 4.2 was designed.

4.4.1 Research Questions

Of the four research questions detailed in 3.4.2, RQ-2 and RQ-3 directly influence the construction of the research platform. RQ-1 and RQ-4 are excluded as they relate more to the methodology used to capture data (RQ-1) or to industry interests that influence the marketability of the research platform as a product rather than as a tool for research (RQ-4).

4.4.2 Research Objectives

Only three of the research objectives directly influence the creation of the SAR research platform. These are:

RO-2: "DESIGN AND DEVELOP AN SAR PLATFORM FOR USE IN CO-DESIGN SESSIONS"

RO-3: "EVALUATE THE EFFICACY OF AN SAR PLATFORM IN COMPLETE CO-DESIGN SESSIONS"

RO-4: "ANALYSE THE IMPACT OF A SAMPLE OF SPECIFIC CHARACTERISTICS AND FEATURES OF THE SAR PLATFORM ON CO-DESIGN SESSIONS"

These three research objectives necessitate the construction of a platform that can:

- Support co-design sessions (RO-2)
- Be used in co-design sessions in such a way as to capture meaningful data to evaluate the platform's efficacy at supporting co-design sessions (RO-3)
- Support a modular analysis of specific characteristics and features and their impact on the co-design process (RO-4)

Any future platform improvements or developments must keep these requirements in mind. In particular, it is imperative that the SAR platform enables the capture of meaningful data that integrates with the metrics used to evaluate the SAR supported design sessions. Presently, the static SPARK platform described in section 4.2 meets these requirements. The platform is able to support co-design sessions, the sessions themselves can be recorded, both through external means and through the use of the logs generated by the platform, and the platform enables a modular approach to analysing characteristics and features of SAR due to its modularity.

However, the SPARK platform is immovable. As such, there are space considerations that must be taken into account for the deployment of such a platform at the University of Bath campus. Furthermore, based on the feedback received from design practitioners, it became apparent that a movable platform would enable additional data capture at other locations. This led to the creation of miniSPARK as described in section 4.3. By relying on the framework of the original SPARK platform, miniSPARK also manages to support the achievement of all the relevant research objectives.

4.5 CHAPTER SUMMARY

This section sought to provide insight into the development of the SAR research platform used to conduct the studies reported on throughout this thesis. In order to achieve this, definitions for the characteristics and features of an SAR platform were created and summarised in Table 16 and Table 17 respectively. The distinction between the characteristics and features of an SAR platform were highlighted to distinguish between: the inherent properties that are required for an SAR platform to be considered such; and the additional features that may be added to improve functionality or support a specific use case.

Subsequently, a description of the original SAR platform was provided. This description discussed the characteristics and features, as well as the functions and subsystems of the SPARK platform. This was the research platform originally constructed by the members of the SPARK consortium to support their research. This initial platform provided the basis for future developments of the SAR research platform. This platform was also compared to the research questions and objectives originally laid out in sections 3.4.2 and 3.4.3 respectively. Through this comparison it was possible to see that the platform in its current state was theoretically adequate at addressing the research questions and objectives.

However, due to a number of technical challenges identified in the platform, as well as other considerations relating to the size and ease of deployment of the platform, an updated platform was required. This section concludes with an overview of the development process of a miniaturised version of the SPARK research platform, miniSPARK, in order to address these considerations.

5 SUPPORTING CO-DESIGN THROUGH SPATIAL AUGMENTED REALITY

Chapter 5 details the studies to meet Research Objective three (Section 3.4.4) by examining how SAR can support, or hinder, the activities conducted in co-design sessions. The experiments detailed here analyse the overall impact that SAR has on co-design sessions rather than look at the impact of a specific feature or characteristic of SAR.

The first set of studies, Comparing SAR and existing Co-Design Tools described in section 5.1, sought to benchmark SAR relative to other tools used to support co-design sessions in controlled lab settings. Namely, SAR was compared to a Handheld AR system as well as non-ICT/standard tools, such as paper cut-outs (as would be used in a regular, non-augmented, design session). The co-design sessions used in this study were simulated using past products that had been developed by SPARK's industrial partners.

The second set of studies, SAR Platform Validation at End-Users' Premises discussed in section 5.2, sought to examine SAR in a set of real-world scenarios. SAR supported co-design sessions hosted by SPARK's industry partners were analysed. Through these studies, the impact of SAR on open-ended co-design sessions was analysed. In addition, insights into the industry needs for supporting co-design through SAR were examined. Unlike the first set of studies, the experiments were run using actual clients and products in development. It should be noted that both sets of experiments were conducted using the SPARK platform described in section 4.2 and not using miniSPARK.

The results from each set of experiments, together with their individual methodologies, are reported below. The chapter then concludes by summarizing the key findings in relations to the Research Objectives and Research Questions.

5.1 COMPARING SAR AND EXISTING CO-DESIGN TOOLS

To examine SAR's ability to support or hinder co-design, a study was devised that compared SAR with existing co-design support tools. Three tools were compared and analysed as part of this study. These were: the static SPARK platform described in section 4.2, a Handheld AR utilising the same SPARK GUI interface described in 4.2.2, and non-ICT/standard tools such as paper cut-outs, colour swatches, and pen and paper sketches. The three tools formed the three conditions for the study.

The data gathered as part of the experiment reported here was aimed at answering Research Objective three. As discussed in Table 12, the expected research contribution of Research Objective three is to provide an analysis of design sessions in order to develop a baseline for future comparison of the impact of SAR. The experiments discussed in this section attempt to provide a contextualization for SAR. By comparing SAR-supported design sessions to sessions supported using other tools, it is possible to better understand the value of SAR as well as substantiate the validity of the overall research presented in this thesis.

The types of co-design reflected the activities performed by the industry partners to enable comparisons with their real-world co-design sessions. Sessions where Artefice participated were focused on packaging design, whereas the sessions with Stimulo focused on product design; more specifically the sessions with Stimulo focused on what they describe as the "colours, materials, and finishes" stage of their design process.

The products evaluated during these design sessions were products that had previously been developed by Artefice and Stimulo. Participants in these design cases were stakeholders in the product being developed, but had not been involved in the design of the product prior to the study.

5.1.1 Experimental Conditions

Table 20 and Table 21 describe the experimental conditions of the co-design involving Artefice and Stimulo respectively. For each design agency, the breakdown provides the type of product used, session brief, and information on the end-users and designers who participated.

The co-design sessions involving Artefice (Table 20) used the same product and session brief. In order to reduce the risk of any learning effect, the designers provided by Artefice were rotated: no designer participated in more than one design condition. For the Stimulo design sessions, constraints including travel to and from the SAR equipped co-design room (Barcelona to Grenoble) and company size led to an alternate approach to mitigating learning effects. The designers for the Stimulo sessions were kept constant throughout all three conditions. However, the products that they worked on changed. The products were selected so that the design briefs were as similar as possible. This was achieved by having the design brief for of all three sessions focus on the colour, materials, and finishes stage.

End-users were permitted to attend a single co-design session out of all the conditions and agencies, and were selected from the general population of student and staff at the universities where the experiments were conducted. None of the end-users had any affiliation to the SPARK Project.

Table 20. Experimental conditions for design sessions involving Artefice (Ben-Guefreche et al., 2018)

DESIGN AGENCY		CONDITION ONE (SAR)	CONDITION TWO (AR)	CONDITION THREE (STANDARD)
ARTEFICE	PRODUCT DESCRIPTION	Fresh soup — single serving in plastic bowl with film lid and cardboard sleeve		
	SESSION BRIEF	Further develop three pre-prepared alternative designs for the cardboard sleeve graphics and layout by combining graphical elements (colours, logos, text, images etc) in order to propose a complete packaging design		
	END-USERS	Female, age 18-30 Male, age 18-30	Male, age 18-30 Male, age 18-30	Female, age 30-45 Female, age 30-45
	DESIGNERS	Digital Creative Director, 16 years of experience, female Art Director, 18 years of experience, female	Senior Art Director, 19 years of experience, male Graphic Designer, 10 years of experience, male	Art Director, 10 years of experience, female Junior Art Director, 1 year of experience, female

Table 21. Experimental conditions for design sessions involving Stimulo (Ben-Guefreche et al., 2018)

DESIGN AGENCY		CONDITION ONE (SAR)	CONDITION TWO (AR)	CONDITION THREE (STANDARD)
STIMULO	PRODUCT DESCRIPTION	Handheld device for assessment of human exposure to electromagnetic fields	Smart fitness product to monitor performance when using gym equipment	Handheld device for communicating your location in an emergency
	SESSION BRIEF	Define the colours, materials, and finish of the main housing. Define the location and pattern of LED status lights and speaker. Location of logo.	Define the colours, materials, and finish of the main housing. Location of logo	Define the colours, materials, and finish of the main housing for specific environments. Define the location and pattern of LED status lights
	END-USERS	Female, age 18-30 Male, age 18-30 Male, age 45-60	Female, age 18-30 Male, age 18-30 Male, age 45-60	Female, age 18-30 Male, age 18-30 Female, age 45-60
	DESIGNERS	Creative Director, 14 years of experience, male Designer and Business Developer, 15 years of experience, male		

Condition one featured the static SAR research platform. As the experiments with Stimulo and Artefice were conducted at two different locations, GINP in Grenoble and POLIMI in Milan respectively, the setups varied slightly between the two locations. However, these variations were mostly limited to seating arrangements. The Information System and Software of the SAR Research platform remained identical between the two locations. The hardware changed slightly in the make and model of projector used, but the final outputs between the projectors were calibrated to obtain similar levels of resolution, brightness, and contrast. An overview of the two SAR platform setups as deployed at GINP and POLIMI is presented in Figure 36, found in 4.2.3.

Condition two featured a Handheld – Object augmentation AR platform. The AR platform consisted of a tablet PC with an 8” screen that was capable of displaying an overlay over physical objects. To achieve this, the prototype used in the co-design session was packaged in white featuring an irregular pattern (Figure 51a). This irregular pattern allows the tablet PC to know the position of the physical prototype in 3D space and in real time, allowing the tablet PC to accurately track the physical prototype as it is translated and rotated. This process is shown in Figure 51. As can be seen from the figure, the distinctive markings enable the AR system to position the overlay in the correct orientation, even as the prototype is moved. The tablet screen also provides the interface by which the designers can modify the design of the packaging. It should be noted however that the digital overlay does not account for ambient lighting conditions. As such, if the AR platform is deployed in areas with uneven or extreme lighting conditions the digital overlay may look out of place. When indoors and even in moderately well-lit conditions, the effect is not particularly unrealistic.

(a) (b) (c)

Figure 51. Image (a) is how the physical prototype appears to the naked eye. Images (b) and (c) show how the physical prototype appear through the tablet PC: with an augmented digital skin. (Ben-Guefreche et al., 2018)

Condition three featured tools that the design companies normally utilise during their co-design sessions. These consisted of, but were not limited to: Pantone® swatches, diagrams, paper prototypes, 3D printed prototypes (Figure 53). The tools taken into the session were chosen by the designers as tools representative of those they would normally have used in such a session. During the sessions, designers and participants made use of sketches; both ones prepared before the sessions and sketches created as part of the discussion during the sessions. The tools provided to support condition three sessions were not made available during sessions in other conditions. The use of sketches to support the design session only occurred in condition three. While the designers and participants were not explicitly forbidden from making use of sketches in the other conditions, they were also not encouraged to do so nor were they explicitly provided with the material to do so.

Figure 52. Examples of the "standard" tools used (Ben-Guefreche *et al.*, 2018)

5.1.2 Experimental setup

Each design company conducted three co-design sessions, giving a total of six sessions. The sessions were all conducted with different end-users. The participants were not given an explicit time limit to complete the activity. Rather, the designers were tasked with managing the session. As a result, they chose when to bring the session to a close, often based on their feeling of how the discussion had evolved.

In addition, the designers were cognisant that multiple sessions were to be completed on the same day. Thus, the designers would have to deploy their time management strategies when interacting with clients. This should, nonetheless, remain representative of their daily activities.

Italian and English were, respectively, the languages used for the Artifce and Stimulo sessions. Both participants and designers in the sessions involving Artefice felt more comfortable using their own native language. In the sessions involving Stimulo the language used was English. However, as one of the two designers felt less confident speaking in English, they often spoke to the end-users via their colleague; speaking in Catalan and having their colleague translate. Figure 53 shows the rooms where the experiments took place.

Figure 53. Participant layout for the experiments. Leftmost image shows the setup at POLIMI, the rightmost two images show the setup at GINP. The end-users and designers are seated at the table, the session observers are seated in the background making annotations on the session.

In each session, one of the researchers compiled a Morphological Chart during the session and conducted pre- and post-session interviews with the designers. This was all performed in accordance with the guidelines for the version 4 metrics as described in section 2.6.8.

The aforementioned metrics specified a four-stage approach to evaluating co-design sessions. These stages consisted of, firstly, a pre-session interview. This was then followed by in vivo data collection. Thereafter, a post-session interview was conducted. Finally, there was a data post-processing step to arrive at the final results.

The pre-session interview was used to collect information regarding the desired session outcomes, any tasks that the designers had in mind that should have been addressed during the session, and the expected number of concepts to needed generating. Here the observer asked the designers a number of questions using a structured interview template. Questions relating to the previous work done on the project as well as the desired session outcomes aided the observer in understanding and contextualising the work they would later see performed during the session. Questions regarding the tasks to complete and their importance aided in assessing the task-progressing metric.

The in vivo data collection was conducted by allowing the observer to view the design session as it progressed. The observer was instructed to avoid disturbing the end-users and designers as much as possible during the data collection and to not interrupt the session.

During the session, the observer compiled a Morphological Chart adding new rows and columns to said chart based on the discussions observed between the designers and end-users. The information gathered during the pre-session interview aided in contextualising the design and allowed the observer to more accurately complete the chart. It is important to note the distinction between the addition of new rows and new columns in the Morphological Chart. Each row in the chart refers to an “idea element category”. These are unique to each design session. At the start of the design session the observer compiled the Morphological Chart by noting down which idea element categories were already present within the design. As the design session progressed, new rows may have been added based on the ideas that emerged from the discussions between the end-users and the designers. Similarly, the columns in the Morphological Chart represent the individual idea elements for each row. New idea elements may be added to the chart as the session progressed. Existing rows and columns and new rows and columns were marked differently in the Morphological Chart to show how the session led to the further exploration of the design space.

In addition, during the in vivo data collection stage, the designers were tasked with taking a screenshot of their ideas. Whenever the designers and end-users arrived at an idea they felt comfortable with and would like to save for future use, they took a screenshot using the appropriate button in the SAR/AR software interface. In the case of condition three (standard) design sessions, the participants were provided with a camera to take pictures of their work.

Subsequently, the post-session interview was conducted with the designers. As with the pre-session interviews, a template was used to structure the discussion. During the post-session interview, the designers were shown the screenshots taken by them during the session. The designers were then asked to confirm or discard any ideas collected through those screenshots in order to compile an idea chart. The ideas collected in this chart were then rated by the designers for novelty. Where more than one designer was present in the interview, the novelty rating was obtained by consensus between the designers. Furthermore, the designers were asked whether each idea would be taken forward, that is to say whether the idea generated would be considered in future sessions to be in order to be developed further.

After the idea chart was completed, the designers were asked to review the Morphological Chart compiled during the session. They were asked to add or remove any rows or columns they felt were incorrect. Moreover, any columns or rows which were marked as new (or as old) could be switched by the designers to better reflect their experiences in the design session.

Thereafter, the designers were asked to review the task chart completed during the pre-session interview. They were asked if any of the tasks noted down had been completed. In addition, the designers were asked to add any new tasks they had identified during the session and, in addition, these on importance.

The final step was then to utilise all the data gathered as part of the interviews and the in vivo data collection to compile the scores for each of the six metrics: Quantity of Ideas, Variety of Ideas, Quality of Ideas, Novelty of Ideas, Task Progress, and Filtering Effectiveness. The calculations required for each of these metrics were laid out in section 2.6.

5.1.3 Results

Table 22 presents a summary of the results. In the Stimulo design cases, AR and SAR scored higher than traditional design methods in Quantity, Quality, and Novelty of Ideas as well as Task Progress. SAR, but not AR, scored higher than standard methods for Variety of ideas. Conversely, Filtering Effectiveness was higher for AR, but not SAR, when compared to standard design practices. The Stimulo study shows that HHD AR and SAR consistently outperform standard design practices.

The Artefice results were somewhat mixed. Standard design practices outperformed both AR and SAR in Novelty of Ideas and Variety of Ideas. However, SAR, but not AR, outperformed standard practices for Quantity and Quality of Ideas as well as tying for Task Progress. AR, but not SAR, tied with standard practices for Filtering Effectiveness.

Table 22. Summary of the results. Higher scores are better, Filtering Effectiveness should approach one. The cells have been shaded to rank scores for the Stimulo and Artefice led sessions separately. In each scenario green indicates the best session, orange the middling one, and red the poorest one. (Ben-Guefreche et al., 2018)

Metric Title	**Stimulo**			**Artefice**		
	SAR	**AR**	**Standard**	**SAR**	**AR**	**Standard**
Quantity of Ideas	8	8	6	11	4	5
Variety of Ideas	Coverage = 5 New Rows = 1	Coverage = 1 New Rows = 1	Coverage = 4 New Rows = 1	Coverage = 2 New Rows = 0	Coverage = 4 New Rows = 0	Coverage = 5 New Rows = 1
Quality of Ideas	4	5	1	3	1	2
Novelty of Ideas	=44/8 = 5.5	=51/8 = 6.4	=23/6 = 3.83	=7/3 = 2.3	=9/4 =2.3	=19/5 = 3.8
Task Progress	1 x High = 3 Total = 3	2 x High = 6 1 x Med = 2 Total = 8	1 x Med = 2 Total = 2	2 x High = 6 Total = 6	1 x High = 3 Total = 3	1 x High = 3 1 x Med = 2 1 x Low = 1 Total = 6
Filtering Effectiveness	= 4/(8-1) = 0.57	= 3/(8-5) = 1	= 5/(6-1) = 1	= 8/(11-1) = 0.8	= 3/(4-2) = 1.5	= 3/(5-3) = 1.5

5.1.4 Observations on the use of SAR in collaborative design

While the results cannot confirm the supremacy of SAR over HHD AR, they do indicate that SAR performs better overall when compared to conventional co-design support tools. As such, the results substantiate the argument that both AR and SAR are viable solutions as tools for co-design. Furthermore, the results identify opportunities for improving the SAR system. From the results it is possible to see that SAR systems underperform when it comes to using them as a tool for filtering designs.

There were some key points to take away from the sessions. Hardware issues with the SAR platform interrupted the sessions and interfered with its progression. The hardware issues encountered mostly concerned the connection between the Information System and Software. On occasion, the two systems would go out of sync necessitating a reset of the system. This typically occurred on longer sessions. The resultant downtime was approximately fifteen minutes.

Furthermore, the participants noted that there was a considerable, noticeable difference between the colours being projected onto the physical prototype and GUI. This was attributed to ambient light interference as well as the lack of a process to calibrate the colour being transmitted to the projector to ensure that they were equivalent. As the tablets used for the GUI were not colour calibrated to a known standard there was some variation in how they displayed colour. The projectors were also not colour calibrated to the same standard as the tablets (as the manufacturing specifications for either were not known). In addition to these complexities, the very nature of projection meant that dark colours could only be as dark as the ambient light in the room, as it is not possible to project black. As a result, the colours projected often appeared lighter than the GUI.

It also became apparent that the participants were limiting their interaction with the physical prototype and that they did not move or rotate it much during the sessions.

Overall, it was observed that SAR has potential as a tool for supporting co-design, but the SAR platform requires improvements before the full potential of SAR is achieved.

5.1.5 Summary of Findings

The key findings identified were:

KF-1. SAR supported co-design sessions generally achieve better outcomes than conventional tool supported co-design sessions. This is particularly true for the Quantity and Quality of Ideas metrics.
KF-2. SAR struggles to support Idea Filtering.
KF-3. Hardware instability causes considerable delays and breaks up the creative flow of the co-design sessions.
KF-4. SAR colour correction and calibration are difficult to implement due to ambient light interference, different colour standards between devices, and the inherent difficulty of projecting dark colours.
KF-5. Minimal participant interaction with the physical model was observed.

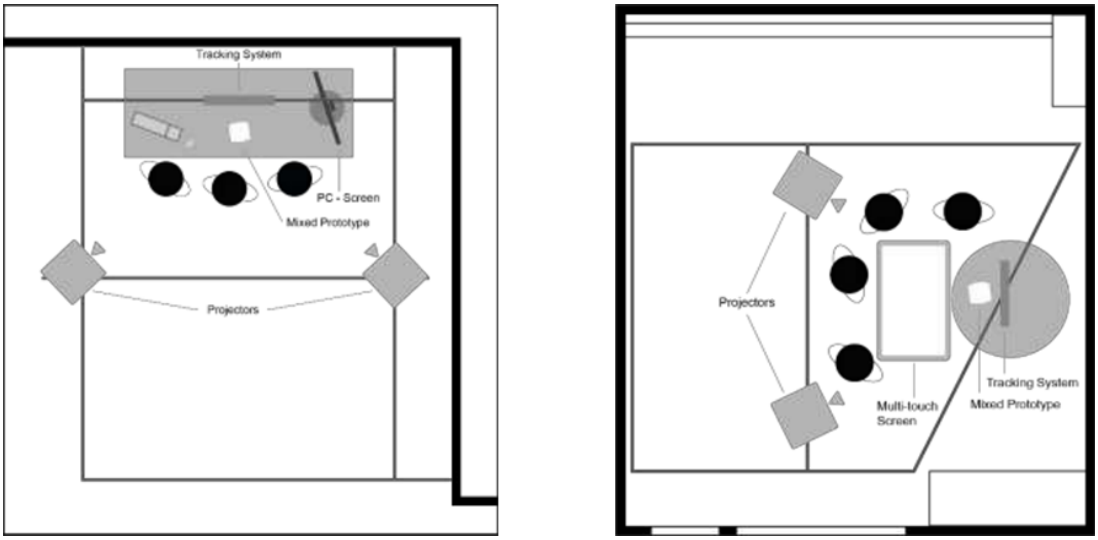
5.2 SAR PLATFORM VALIDATION AT END-USERS' PREMISES

Study Two placed SAR in the less controlled, albeit more realistic, environment of end-user premises for in vivo experimentation. This is to contrast Study One’s in vitro experiments. However, both studies are targeted towards RO-3: “Evaluate the efficacy of an SAR platform in Complete Co-Design Sessions” as discussed in section 3.4.5 of the Methodology.

5.2.1 Experimental Setup

The experimental setup for the analysis of the co-design sessions did not change significantly from that used for Study One as this would allow for cross-comparison of the results. The same version 4 metrics were implemented following the same procedure. The major difference between the experiments reported here and those from Study One was that the users were no longer simulated, nor were the design tasks. The experiments were conducted at the user premises providing a much more natural setting. As such, the participants communicated in their native languages, Italian and Catalan/Spanish for Artefice and Stimulo respectively.

The clients who commissioned the design work were invited to participate, as they normally would, in the co-design sessions with the intent of making decisions that would affect their projects. As such, the co-design sessions closely align with the conditions that would be found in a regular design agency interacting with a design client. One example being the sessions led by Artefice. These all involved the same partner, working on three different products, the sessions were conducted back-to-back over the course of one day.



(a) (b)

Figure 54. SPARK Setup at Stimulo (a) and at Artefice (b) (Bellucci et al., 2018a)

Figure 54 shows how SPARK was setup at both Stimulo and Artefice. The Stimulo setup was essentially identical to that used in Study One while Artefice's setup included a "Multi-touch Screen". This is in essence a large tablet that the designer can use to make the necessary design changes. Both locations had tracking systems equipped with infrared markers to track the physical prototypes and both locations made use of two projectors.

Five sessions were conducted. Three of these sessions were led by Artefice and conducted at their location using the setup shown in Figure 54a. All three of the sessions led by Artefice involved a single design client who had commissioned Artefice to design the packaging for three separate food products. The sessions led by Stimulo, using the setup shown in Figure 54b, involved two separate and unrelated design clients. Confidentiality agreements limit the amount of information that can be disclosed with regards to these products and packages. All the sessions conducted were at roughly the same stage in the design process, enabling the cross-comparison of the sessions.

All experiments were run so as to match a real design session as closely as possible and without impeding the partner's co-design session process. The only additions were the pre- and post-session interviews and the availability of the SAR platform. No time limits were set, and breaks were permitted when and where necessary, with both the clients and the designers free to call for one at any point. As the sessions with Stimulo involved two separate clients, the clients were invited at different times and did not see the SAR platform in use by the other client. In the case of Artefice, as the client was the same, working on different products, the apportionment of time for each product was left to the designers and the clients. However, as the sessions were due to take place all on the same day, all parties were conscious of time and the need to complete the sessions by the end of the day.

5.2.2 Results

Stimulo conducted design sessions with two different clients: LISN and Neosonics. Artefice worked on three different products: a frozen Pizza package, a Mozzarella container, and a Parmigiano wrapper. Figure 55 shows two of these co-design sessions, one for Artefice and one for Stimulo.

(a) **(b)**

Figure 55. Design session in progress at Artefice (a) and Stimulo (b) (Bellucci et al., 2018a)

Table 23 shows the results of the co-design sessions. The results for Artefice indicate that the Mozzarella packaging co-design session scored worst of all with the Parmigiano and the Pizza having relatively comparable results. The Stimulo sessions see LISN performing marginally better but overall, the results are fairly comparable.

Table 23. Stimulo and Artefice results according to v4 Metrics (Bellucci *et al.*, 2018a). Higher scores are better, Filtering Effectiveness should approach one. The cells have been shaded to rank scores for the Stimulo and Artefice led sessions separately. In each scenario green indicates the best session, orange the middling one, and red the poorest one.

Metric Title	**Stimulo**		**Artefice**		
	LISN	**Neosonics**	**Pizza**	**Mozzarella**	**Parmigiano**
Quantity of Ideas	🟩 8	🟥 7	🟧 4	🟥 3	🟩 5
Variety of Ideas	🟩 Coverage = 8 New Rows = 2	🟥 Coverage = 5 New Rows = 0	🟧 Coverage = 3 New Rows = 0	🟩 Coverage = 3 New Rows = 1	🟧 Coverage = 1 New Rows = 1
Quality of Ideas	🟥 4	🟩 7	🟩 4	🟧 2	🟧 2
Novelty of Ideas	🟩 =39/8 = 4.9	🟩 =34/7 = 4.9	🟩 =25/4 = 6.3	🟧 =12/3 =4	🟥 =13/5 = 2.6
Task Progress	🟩 0	🟩 0	🟩 1 x High = 3 4 x Med = 8 2 x Low = 2 Total = 13	🟥 2 x Med= 4 Total = 4	🟧 2 x High = 6 Total = 6
Filtering Effectiveness	🟩 = 4/(8-1) = 0.57	🟥 0	🟥 = 0/(4-2) = 0	🟧 = 1/(3-1) = 0.5	🟩 = 3/(5-2) = 1

5.2.3 Observations

Of the sessions led by Artefice, the Mozzarella session was overall the worst whereas the Pizza session performed marginally better than the Parmigiano session. The Quantity of Ideas was lowest among all the sessions as was the Task Progress. Variety, Novelty, and Filtering Effectiveness were middling. The Task Progress metric revealed that the Mozzarella session made limited progress and not only were few ideas generated, but their quality was quite low. The failure of the Mozzarella session can be attributed to the fact that the packaging was predominately white. Since projecting shades of white is rather difficult, due to the inherent brightness of the projector, the participants complained of poor rendering quality and of difficulties interpreting the images. Furthermore, the Mozzarella physical prototype was the smallest of any of the physical prototypes used, making text rendering difficult due to interpolation of raster images caused by scaling. Lastly the Mozzarella session took place at the end of the day, and it is possible that fatigue reduced the interest of the participants.

In contrast, the Stimulo sessions proceeded smoothly. The main point of concern is that both sessions scored a zero for Task Progress. Post-interviews and follow-up discussions did not provide any insights into why this was the case.

5.2.4 Summary of Findings

The key findings identified as part of this study were:

KF-6. Colour accuracy and visibility continue to interfere with session progress.
KF-7. Physical prototype size plays a role in how the product is designed and interpreted during the session.
KF-8. Changing the size of the physical prototype to scale up may aid in improving the resolution and quality of the rendered images, as well as the colour rendition.
KF-9. Fatigue should be taken into account when running co-design sessions, participants should be allowed sufficient breaks and time.

5.3 CHAPTER SUMMARY

Chapter 5 examined how SAR can support, or hinder, the activities conducted in co-design sessions. Section 5.1 sought to benchmark SAR relative to other tools used to support co-design sessions in a controlled lab setting. Section 5.2 set out to examine SAR in a set of real-world scenarios.

Table 24 shows the research outcomes for each of the experiments mapped back to the research questions first described in section 3.4.2. The experiments detailed in this chapter were, as discussed in section 3.4.4, Table 12, aimed at addressing RO-3: “Evaluate the efficacy of an SAR platform in Complete Co-Design Sessions”. The experiments achieved this objective as they enabled the testing of an SAR platform in realistic design scenarios.

Table 24. Research Outcomes Compared to Research Questions

RESEARCH QUESTIONS	RESEARCH ACTIVITIES	
	5.1 Comparing SAR and existing Co-Design Tools	5.2 SAR Platform Validation at End-Users' Premises
RQ-1: “How can co-design sessions’ efficacy be measured?”		
RQ-2: “How does an SAR system affect co-design sessions’ efficacy?”	• Compared to other tools, such as AR and standard design tools, SAR provides better outcomes particularly for the Quantity and Quality of Ideas. • Idea Filtering seems to compare poorly to other technologies.	• Low Task Progress noted in some sessions. Other results consistent with previous findings. • Physical prototype size plays a role in how the product is designed and interpreted during the session.
RQ-3: “How do specific SAR characteristics and features affect co-design sessions’ efficacy?”	• Participant interaction with the SAR physical model was less than expected.	
RQ-4: “What are the industry requirements for an SAR system to support co-design?”	• Hardware instability must be avoided, elsewise it will cause frustration in design session participants. • Colour accuracy is desirable for improved immersion.	• Physical prototype size and colour rendition limitations, must be accounted for to avoid decreased quality of projection, which impacts the session.

6 STUDIES INTO SAR SYSTEM CHARACTERISTICS AND FEATURES

Chapter 6 presents the studies conducted to answer Research Question three via Research Objective four. Two characteristics and one feature of SAR, as defined in section 4.1, were selected for analysis and further study. These were:

1. User Interfaces (section 6.1);
2. Communication Between Participants (section 6.2); and,
3. Physical Prototype Scale (section 6.3).

Section 6.1, Interface Comparison Study, describes a controlled experiment to compare four different types of interfaces used to interact with the SAR platform GUI (as described in section 4.2.2). Data was collected from participants across three locations: University of Bath (UBATH), Politecnico di Milano (POLIMI) and Grenoble INP (GINP). The results reported here were originally published in a report by Morosi *et al.* (2018b). Table 14 in section 3.4.5 provides an overview of the authorship of the works presented throughout this thesis. As noted in the table, the main lead for the work presented in this section was the SPARK Consortium. Thus, the experiments discussed in this section were not specifically designed to answer one of the research questions laid out in this thesis. However, they do provide a valuable and unique source of experimental data that this PhD has been fortunate enough to have access to. As such, the results presented here are secondary analysis of the study data that sought to uncover insights to address the Research Questions and Objectives. Thus, the studies will be discussed alongside the secondary analysis that has been performed.

The Impact of SAR on Communication between Design Session Participants (section 6.2) aimed to investigate the impact that an SAR platform had on the ability of participants to communicate effectively. The controlled study consisted of one participant having to communicate a pre-defined concept to the other participant, who controlled the SAR platform's GUI. The results discussed in this section were first published in a paper by Giunta *et al.* (Giunta *et al.*, 2019)

Finally, the Impact of Scale in Design Sessions Supported by an SAR Platform (section 6.3) discussed a controlled experiment designed to analyse the impact the scale of the physical prototype used in an SAR platform would have on the design process and outcomes. This was a direct response to the learnings reported on in section 5.2.4, where scale was identified as a potential element that influenced the designers' ability to perform. This study was also first published as a conference paper by Giunta *et al.* (2020).

The results from each set of experiments, together with their individual methodologies, are reported below. Table 14 and Table 15, found in section 3.4.5, show how these experiments are positioned to answer the research question and achieve the research objectives.

6.1 INTERFACE COMPARISON STUDY

The study was performed at the University of Bath (UBATH), Politecnico di Milano (POLIMI) and Grenoble INP (GINP). The conclusions presented in this section are the result of a combined analysis performed by members of all three universities based on the data collected from each partner university. The use of multiple locations enabled the recruitment of more participants for the experiment, at the cost of increased variance between the participants.

The results from the studies reported in section 5.1 highlighted a lack of understanding regarding the impact that the interface used by the designers had on the design process and the SAR system. Furthermore, as a result of the developmental nature of the SAR platform, the SAR platform had received multiple updates to its interface. Whilst these updates took place after the studies reported in sections 5.1 and 5.2, the question was raised regarding the efficacy of the changes made.

For these reasons, an analysis of the impact of the SAR platform's interface was undertaken. To allow for a better focus on the impact of the interface itself, and to reduce the number of extraneous variables, the decision was made to run the studies using a single participant in each session. This was in contrast to previous experiments, which had all used multiple participants collaborating together. Since the SAR platform was not developed with the use case of multiple individuals interacting with the interface, this was considered an acceptable simplification.

The experiment analysed the impact of three different types of interface on a set design task. The design task consisted of replicating the design of a packaging using the SAR platform, using provided assets. It is important to note that assets, when used in the context of the SAR platform, refer to the graphics that are projected onto the physical model. These assets can be logos, textures, or other images used to represent specific elements of the design. The metrics used to evaluate the efficacy of each type of interface were:

- Accuracy of placement (position), rotation, and scaling of the assets
- The participants' efficiency in making use of the interface: measured in time taken to complete the task.
- The users' reported usability of the system based on the Creativity Support Index (CSI) and the System Usability Scale (SUS); these metrics are discussed in further detail in section 6.1.2.2

6.1.1 Experimental Setup

The experimental conditions consisted of one training condition (A), three experimental conditions (B1, B2, and B3), and one control condition (B4).

Condition A consisted of a traditional mouse and keyboard interface with the visualization of the graphical user interface (GUI) occurring on a regular computer monitor. No SAR features were included in this condition. As such, the only way for the participants to view the edits made was through the GUI itself. Figure 56 shows the condition as it would have appeared to the participants.

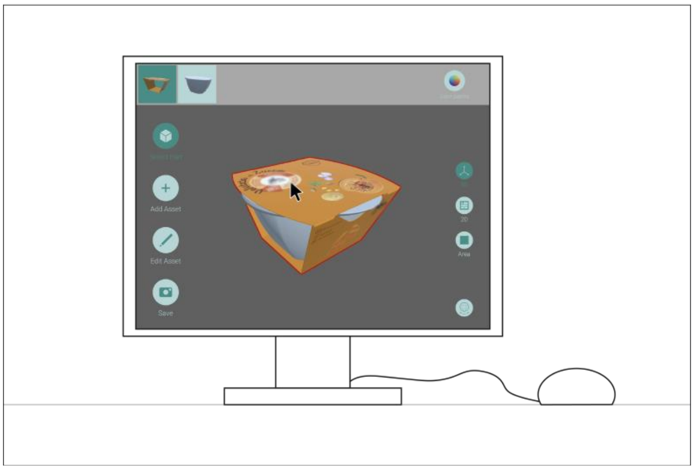


Figure 56. Conditions A and B4 showing the user interface displaying a digital model for edit by the participants (Morosi et al., 2018b)

In conditions B1, B2, and B3 (Figure 57) the GUI was displayed on a Samsung Galaxy Tab S2 (2016) tablet with a screen size of 9.7” (24.64cm).

Condition B1 consisted of a 2D representation of the 3D model (Figure 57a). This 2D representation was the UV map used by the SAR system to “unwrap” the 3D model to understand how to texture the 3D model. UV maps are discussed at greater length in section 4.2.1. In this configuration the user can see all the assets as they are placed on the UV map. The computer then interprets this information to display the assets onto the mixed prototype.

Figure 57b shows the GUI used in condition B2. This GUI only displayed the options for selection and placement of assets but, unlike all other conditions, did not display the assets during or after placement. The assets were only visible on the mixed prototype and could be edited once placed down but could not be viewed using the GUI. All other features of this GUI were identical to the other conditions.

Condition B3, shown in Figure 57c, provided the user with a full 3D model of the mixed prototype. This is similar to conditions A and B4. However, unlike with conditions A and B4, the interface is touch based.

The touch-based interface used in the three main experimental conditions (B1, B2, and B3) enabled the use of multi-touch gestures to control the interface. These gestures were as follows:

- Pinch: this scaled the selected asset.
- Rotate: a clockwise or counter-clockwise rotation using two fingers allowed the asset to be rotated in the respective direction.
- Drag: using one finger to translate an asset.

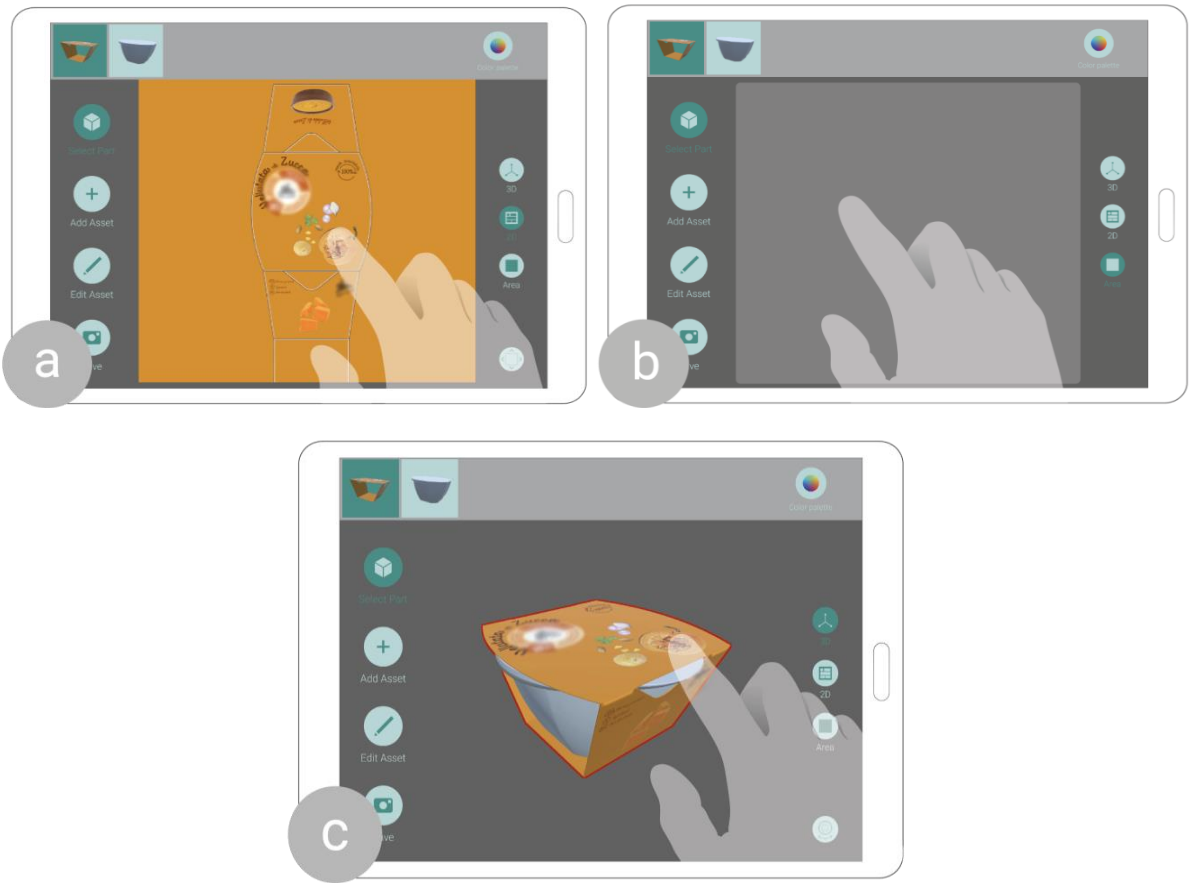


Figure 57. a. UV Map (B1), b. Touch Area (B2), c. 3D View (B3) (Morosi et al., 2018b)

In conditions B1 through to B3 the SAR platform projected the images displayed on the interface onto a mixed prototype. This occurred in real time as changes were made within the GUI. Figure 58 shows one such session in progress. The mixed prototype appears bright white to the camera due to the projection. The participant can be seen making changes to the projection by means of the tablet interface.

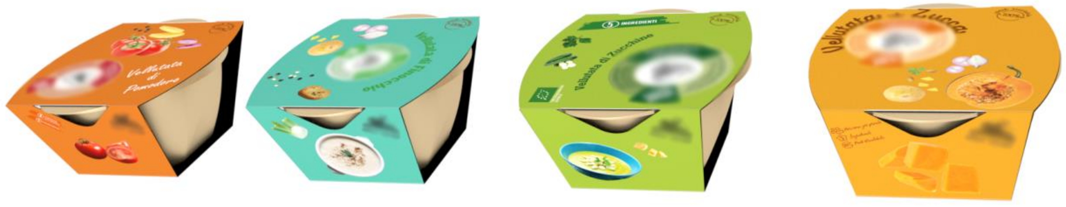
Figure 58. Session in progress showing the tablet interface in use while the projector is projecting onto the mixed prototype

Condition B4 (Figure 56) is identical to condition A. Both use a mouse and keyboard interface to interact with the GUI and use a monitor to display the GUI. This condition was included as a control to understand the impact of a potential learning effect caused by the use of condition A by all participants as a training session to understand how to use the GUI.

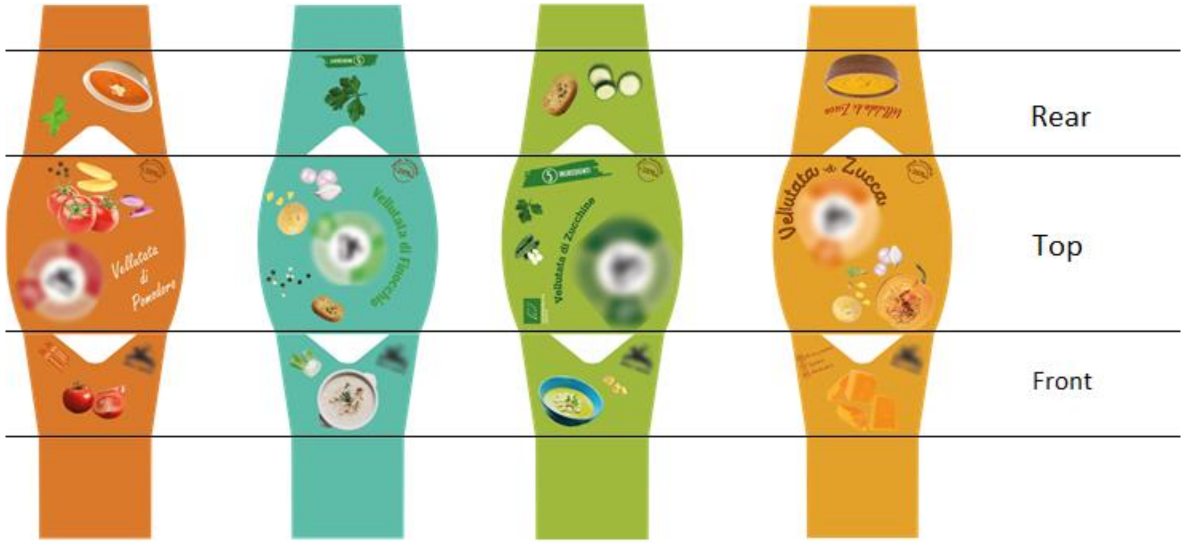
6.1.1.1 Design Task

Before testing any condition, the participant was provided with a physical print-out of one of the layouts shown in Figure 59. The layouts represent four different cardboard sleeves for a pre-packaged soup. The participant was then informed that, using the interface provided, they had to recreate the layout as accurately as possible. This meant placing the same assets as shown on the reference layout they were provided in the same position, with the same rotation, and to the same scale. The background colour of the sleeve also had to match that of the reference layout.

To enable session comparisons, each layout featured the same number of assets distributed over the three main faces of the sleeve (front, top, and rear) (Figure 59b). All layouts had three assets on the front face, seven on the top face, and two on the rear face. All the assets across all conditions are placed and rotated at random with two exceptions: one asset placed in the top right corner of the top face and one asset placed in the top right corner of the front face. These two assets are identical across all layouts (position, rotation, and scale) and acted as controls.



(a) 3D view of the four alternative layouts



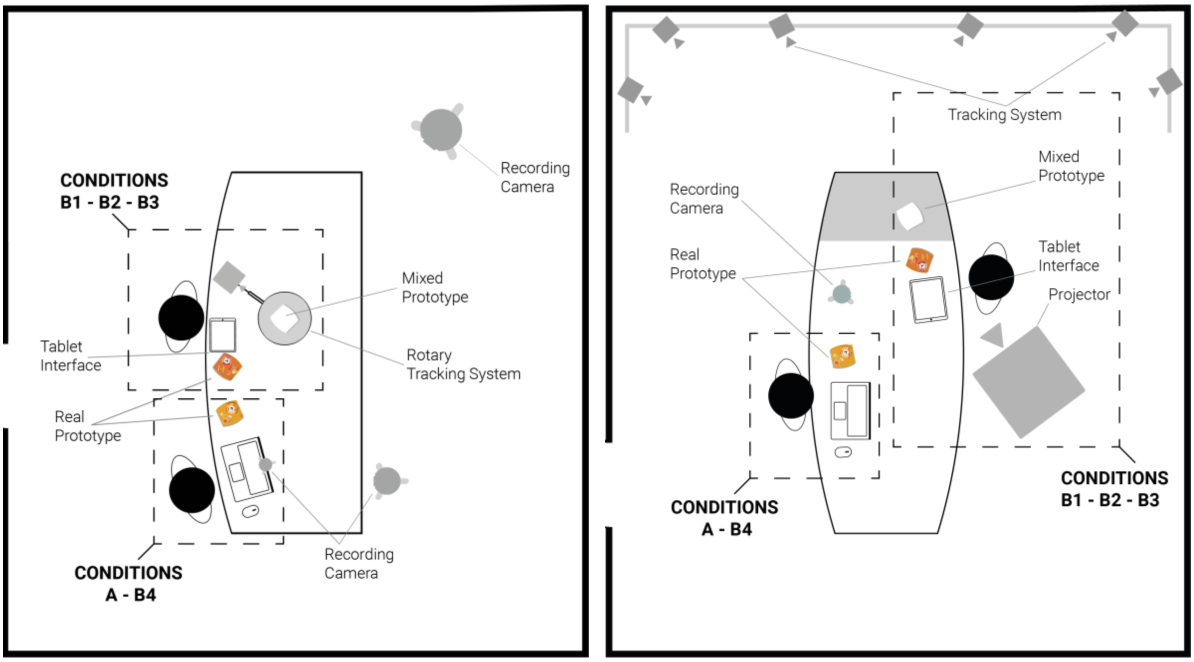
(b) Cardboard sleeves flattened out to highlight front, top, and rear sections.

Figure 59. The four alternative layouts for the cardboard sleeve. Some assets have been blurred to preserve confidentiality (Morosi et al., 2018b)

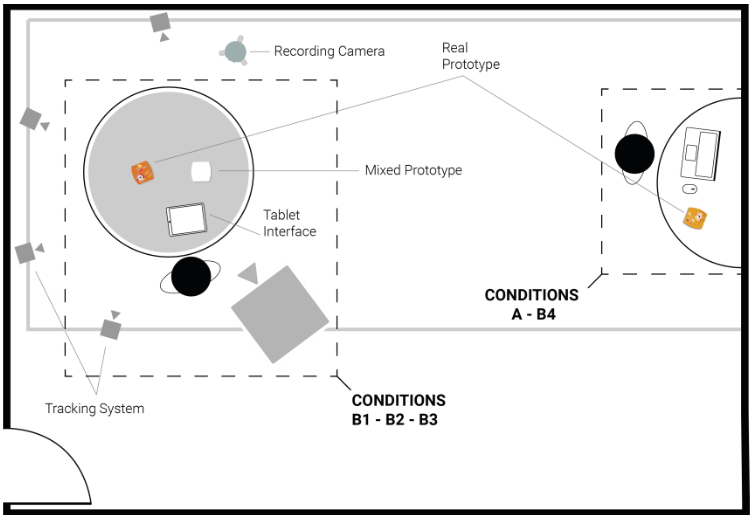
6.1.1.2 Experimental Procedure

Figure 60 shows the experimental setup at the three test locations. Two test stations were set up in each location. One station was used for the testing of conditions B1, B2, and B3 (conditions using the tablet interface). The other test station was used for testing conditions A and B4 (conditions using the PC and mouse). It should be further highlighted that the SAR system used in the experiments at UBATH featured portable projectors and rotational tracking. The setup used at UBATH, and its capabilities, were first described in section 4.3.3. This differed from the SAR system used at both POLIMI and GINP, who made use of standard size ceiling-mounted projectors and infrared tracking for the mixed prototype.

The use of a different projection at UBATH was expected to greatly influence the results of the experiment. While portable projectors usually have lower resolutions and brightness, this is compensated for by having the projector closer to the projection surface. The lack of translational tracking may have played a larger role. This is, however, not entirely certain as the experiments discussed in section 5.1 highlighted that the participants often did not interact much with the physical prototype.



(a) UBATH Setup **(b) POLIMI Setup**



(c) GINP Setup

Figure 60. Experimental setups across the three test locations (Morosi *et al.*, 2018b)

Each session was run as follows:

1. **Briefing:** The task was described to the participant, and they were invited to sign the relevant consent forms as well as ask questions.
2. **Training for condition A:** A short presentation was given to guide the participant through the interface.
3. **Free learning:** The participant was given a couple minutes to familiarise themselves with the interface.
4. **Condition A:** On starting condition A, the participant was informed that they had 10 minutes to recreate one of the four layouts (Figure 59). The layout they were provided with was random and the participant did not see the other layouts. The participant started the session with a blank model displayed on the interface and had to recreate the layout provided as faithfully as possible within the time given. If the participant required more time, beyond the original 10 minutes, they were permitted to continue until they felt they had completed the task. Once the task was complete, the participant was asked to save their work.
5. **1st Questionnaire Round:** The participant was then asked to fill out both a CSI and SUS questionnaire.
6. **Training for condition B:** The functionality of the interface used in the B conditions were identical. This step was taken to ensure that the participant was familiar with the new visualization and interaction. The exact training varied based on the type of condition B the participant was to be tested for.
7. **Condition B:** Similarly, to the condition A test, the user had to recreate a layout within a 10 minute period. A new layout the participant had not previously encountered was provided for this task. In the event that the participant was selected for condition B4 the participant was asked to repeat the task with a new layout.
8. **2nd Questionnaire Round:** As in the 1st questionnaire round, the participant was asked to fill out a new set of SUS and CSI surveys.

6.1.1.3 Session Evaluation Metrics

Five metrics were used to evaluate the impact of the interface on the participants' ability to complete the design task. These metrics were: Placement Accuracy, Rotation Accuracy, Scaling Accuracy, Participant Efficiency, and Usability. These are defined in Table 25.

Table 25. Metric Definitions (Morosi et al., 2018b)

Metric		Definition
Accuracy	Placement	Difference between asset location as placed by the participant and asset location as defined by the reference location.
	Rotation	Difference between asset angle as placed by the participant and asset angle as defined by the reference location.
	Scaling	Difference between asset scale as placed by the participant and asset scale as defined by the reference location.
Participant Efficiency		Time taken to complete the task.
Usability		Based on the results from the CSI and SUS surveys.

It should be noted that in the case of the Placement, Rotation, and Scaling Accuracies, the difference was calculated as the mean difference between the asset placement as placed by the participant, and the reference location across all assets for that specific layout.

6.1.2 Methodology for Data Collection

Quantitative and qualitative data was collected during the course of the experiments. Qualitative data was predominantly collected through observation of participant behaviour. An example of this would be the questions asked by them and the uncertainties they had when completing the task. All sessions were video recorded to capture these interactions and, upon completion of the experiments, a reflection was compiled to better synthesise these observations.

Quantitative data was tracked through the SAR system's own logging capabilities as discussed in section 6.1.2.1. In addition to the logging system, the SUS and CSI surveys allowed to better quantify the opinions of the participants and capture their thoughts on the SAR platform's performance. This is expanded upon in section 6.1.2.2.

6.1.2.1 Log File

The SAR system is equipped with a logging feature. This allows the system to record the actions taken by the participants as well as the current state of the system, the assets, and the model. The log file is stored as a Comma Separated Value (CSV) file. As the software that runs the SAR system runs, and loops through, a new line is added to the CSV file recording the relevant information. This provides redundancy in the event of a system crash.

Additionally, participants were asked to press the button to trigger the Saved Version (as discussed in section 4.2.2) when they felt confident they had completed their given task. In this way it was possible to track the time taken to complete the task, as this would be saved in the log along with all the other pertinent information. Furthermore, as this command saves all the activity information, it was possible to use the information from this line in the log file to evaluate the accuracy of the placement of the assets by the users.

6.1.2.2 Participants' Perception of Usability

Two surveys were used to capture the participants' opinions on the usability of the SAR system. These were the Creativity Support Index (CSI) (Cherry and Latulipe, 2014) and the System Usability Scale (SUS) (Brooke, 1996). The CSI survey consists of two sections, the first allows the user to rate the performance of the SAR system in terms of the system's ability to support them creatively. Users are asked to rate creativity factors such as “exploration” and “immersion in the task”. In the second section of the CSI survey the user is asked to rate the importance of the creativity factors. These ratings are then used to generate weightings to determine the importance of the creativity factors. Cherry and Latulipe’s paper (2014) also provides instructions on how to use their digital tool that automatically calculates CSI scores.

It should be noted that the CSI survey was modified prior to the implementation in this study. As the task the participants were asked to complete was not particularly creative, because they simply had to replicate a given design on their own, the CSI survey was modified to reflect this. Two creativity factors were removed from the questionnaire, the “Expressiveness” and “Collaboration” factors as the task did not include any element of either. The remaining factors were: “Results Worth the Effort”, “Exploration”, “Immersion”, and “Enjoyment”. The scoring calculation was then adjusted to reflect this change. Without the exclusion of the “Expressiveness” and “Collaboration” factors the maximum raw score for the CSI is 300. This is then divided by 3 to obtain a score out of 100. The omission of the two factors leads to a maximum total score of 120. To maintain a final score out of 100, the score was divided by 1.2 instead of 3. A copy of the full CSI survey questionnaire is available in section E of the appendix.

In the SUS survey participants are asked to complete a five-point Likert scale scoring participants’ agreement with ten statements. A complete copy of the survey can be found in section F of the appendix. Five of the ten statements are phrased in a positive manner (e.g.: “I think that I would like to use this system frequently”) and five statements are framed negatively (e.g.: “I found the system unnecessarily complex”). The responses of the participants are then tallied and a score out of one hundred is calculated. The SUS final score calculation is as follows: a score is calculated for questions 1,3,5,7, and 9; this is calculated by taking the response value (from 1-5 with 1 representing strongly disagree and 5 representing strongly agree) and subtracting 1. The scores for questions 2,4,6,8, and 10 are calculated as 5 minus the value of the response (from 1-5 with 1 representing strongly disagree and 5 representing strongly agree). The sums from all the scores are then summed together and multiplied by 2.5 in order to obtain a final score ranging from 0-100 (Brooke, 1996).

6.1.3 Data Analysis Methodology

The data collected was used to generate three metrics to evaluate participant performance. These were Accuracy, Participant Efficiency, and Usability. The following section expounds on how each parameter was calculated based on the data collected.

6.1.3.1 Accuracy

Using the information captured in the log through the Saved Version it was possible to evaluate the accuracy with which the assets were placed on the physical model. The placement, rotation, and scaling of each asset was stored within a Saved Version line in the log file. The results from each participant’s session for the placement, rotation and scaling of each asset were extracted from the session’s log file and were evaluated by comparing the results of their session against the target values. The difference between the expected placement of the assets and the effective placement of the assets by the participants provided a measure of how accurate they were in completing the task.

The target value was not a preordained value. Rather it was calculated as the mean position (separately for rotation, placement, and scale), for each asset, in every layout across all the sessions in one location. In other words, each location (UBATH, POLIMI, and GINP) calculated the average placement of each asset for each layout and used this as the target value. By using a mean value rather than a preordained one, it was possible to mitigate possible issues that could arise due to calibration difficulties. This enabled better comparison between the various test locations.

The difference between the target value and the actual value recorded in each session was taken as the error. In the case of the placement error the error in both the X and Y axes was combined into a total position error. This was done by calculating the Euclidean distance i.e., the vector between the target position and the actual position. The placement error was measured in millimetres.

For rotation, a similar method was applied, however, the target and actual position were compared through vectors rather than simple subtraction. This was because the smallest difference between the target and actual angle was required. Otherwise comparing a target value of 350°, to the actual value of 5°, would return an error of 345° rather than 15°. This is because 360° and 0° are the same value from a rotational perspective.

In the case of scale, recorded scale and target scale were compared as percentages of the total canvas. The difference in percentage was then used to measure the accuracy.

This process was repeated for all three Accuracy metrics across all sessions for a given condition. The values were then used to create the mean error and standard deviation for each condition.

6.1.3.2 Participant Efficiency

Participant Efficiency was measured by evaluating the time taken to complete the task. The log file was used to calculate the time taken. At the beginning of each session, the Saved Version button was triggered to indicate the start of the session in the log file. Once the participant was satisfied they had completed the task, they were then asked to trigger the Saved Version once more. The difference in the timestamps was then used to calculate the exact duration of the task. In addition to recording the total time of the session, the log categorised how the time was spent by the participant into one of four categories. These four activities are:

- **UI use:** This activity refers to time the participant spent interacting with the user interface in some way, such as by adding/removing assets, viewing the assets available, or modifying existing assets.
- **Virtual prototype manipulation:** This refers to time spent interacting with the digital representation of the model in the user interface. This was possible for all conditions except B2.
- **Colour selection:** This activity occurs when the participant selects the background colour of the model to match that of the target layout provided.
- **Real prototype manipulation:** This refers to any time spent manipulating the physical prototype. This was not possible for conditions A and B4 as they did not include a physical prototype.

6.1.3.3 Participants' Perception of Usability

The SUS survey was scored according to its predefined metrics. In this way the survey returns a score out of 100. To evaluate the results, the guidelines laid out by Bangor *et al.*(2009) were used.

In the case of the CSI survey some adjustments had to be made to the scoring system. As two factors were excluded ("Expressiveness" and "Collaboration") the scores needed to be adjusted to reflect the change. The unmodified CSI can produce a maximum raw score of 300, which is normally then divided by 3 to obtain a final score out of 100. The modified CSI can generate a maximum score of 120. For this reason, it was decided to divide the results of the modified CSI survey by 1.2 to maintain a final score out of 100.

6.1.4 Results

The following section details the results of the study according to the three parameters (Accuracy, Participant Efficiency, and Usability). Unfortunately, the log files collected at GINP were unusable. This meant that the results for Accuracy and Participant Efficiency could only be calculated for UBATH and POLIMI. However, the experiments were conducted identically across all three locations, with the exception of the tracking system used at UBATH. As the rig used in the UBATH experiments differed from that used in the two other locations, though following all other parameters, the decision was made to present the data separately. For this reason, it was still possible to use the CSI and SUS scores from all three partners to calculate the Usability of the SAR platform.

6.1.4.1 Participant Information

All participants were students attending one of the three universities running the experiments (UBATH, GINP, POLIMI). The level of design proficiency varied across the participant population. In order to reduce the level of variation, only final year undergraduates, postgraduates, and doctoral students were selected to participate in the experiments. As such, their level of knowledge was assumed to be sufficient to be classed as adept at design. All participants would be expected to have a basic grasp of design and design practice. Additionally, participants were asked to state their CAD experience in number of years as well as to self-identify as either a designer or an engineer, in order to gauge the participants' overall proficiency with computer-based design. A summary of the participants is provided in Table 25.

Table 26. Participant data per university

	UBATH	GINP	POLIMI
NUMBER OF PARTICIPANTS	19	28	37
OF WHICH FEMALE	5	10	6
NUMBER SELF-IDENTIFYING AS A 'DESIGNER'	7	-	8
MEAN YEARS OF EXPERIENCE USING CAD SOFTWARE	6	-	5

6.1.4.2 Accuracy

Figure 61 shows a box plot of the error across all the tested conditions at both UBATH and POLIMI. The accuracy of placement does not appear to vary significantly between conditions. This is true for both the experiments conducted at UBATH and at POLIMI. It is interesting to note that the experiments conducted at UBATH resulted in a marginal improvement in the average accuracy, despite an increase in the variance. This may indicate that the platform used may play a larger role than originally imagined. Furthermore, it would appear that condition B3, the 3D view, had the least amount of variance. Overall, the accuracy across all conditions tested was high, with an average error of less than 2mm in all conditions except B4.

One additional finding is that condition B4 was less accurate, and showed more variance than condition A. This is particularly true for the UBATH cases. The data does, however, support the claim that there is no learning effect caused by condition A, as condition B4 would be expected to be more accurate had this been the case. One potential explanation for this could be that the participants become bored with repeating the same task a second time and are less attentive as a result.

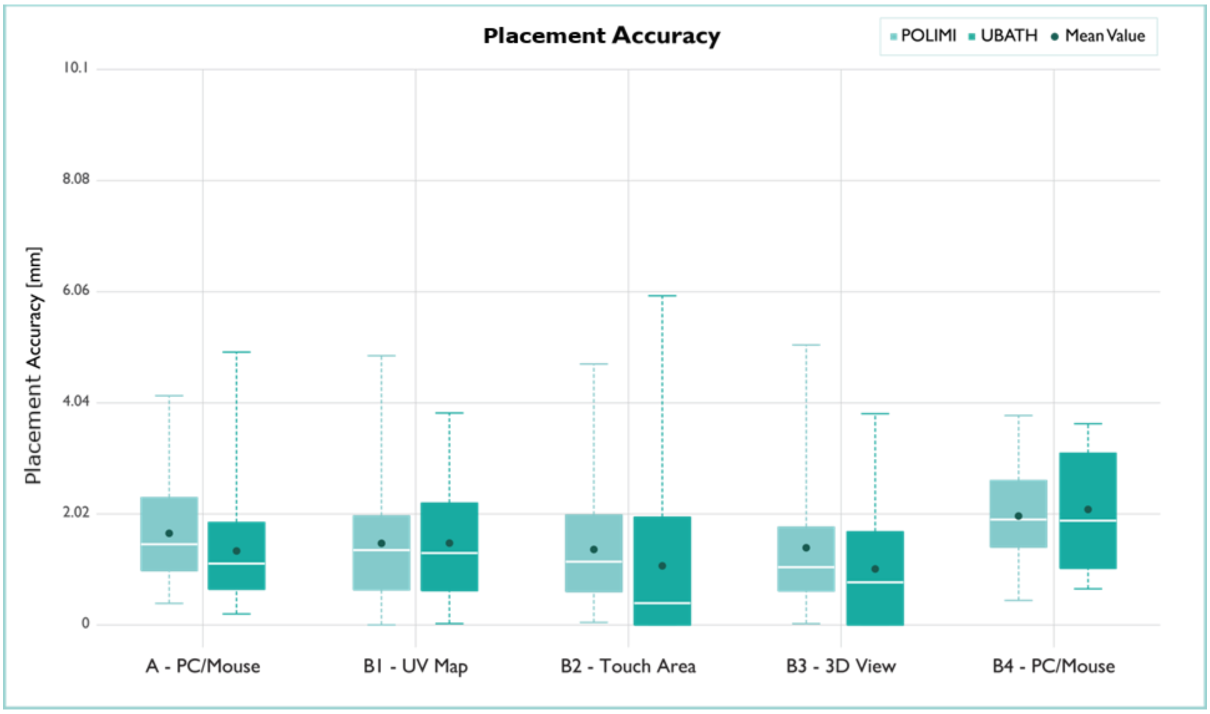


Figure 61. Box plot showing the position accuracy in millimetres for each condition at UBATH and POLIMI (lower is better) (Morosi et al., 2018b)

Rotation accuracy, shown in Figure 62, does not appear to vary significantly between conditions. As with Placement Accuracy, no learning effect was observed since the condition B4 did not appear to be more accurate than the condition A. Results for conditions B1 and B2 appear consistent across the two locations. Condition B3 saw greater variance than that of conditions B1 and B2 in the case of POLIMI but more concentrated for UBATH. The results therefore indicate that the interface does not play a major role in the Rotation Accuracy.

Of particular interest for the UBATH experiments is the increased variance in condition B4 when compared to condition A, with errors of up to 35° recorded. As condition B4 was identical between POLIMI and UBATH, since no projection occurred, it is hard to understand what underlying factors may have influenced this outcome.

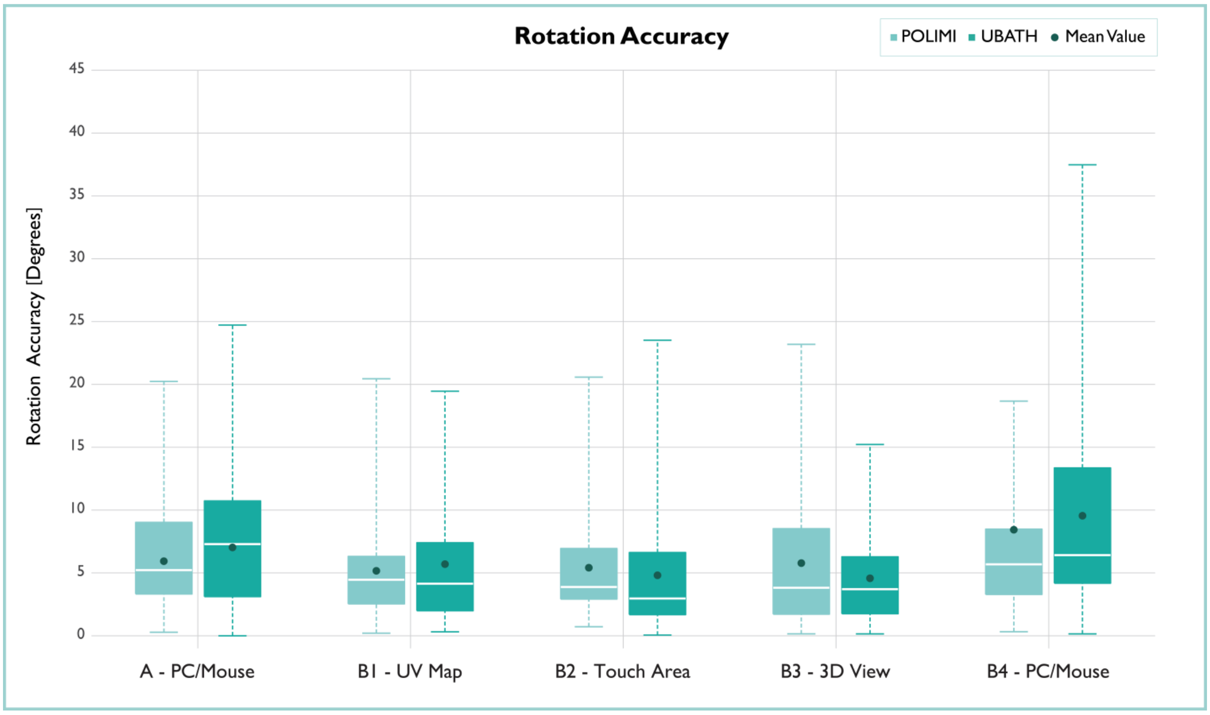


Figure 62. Box plot of the rotation accuracy, in degrees, for each of the conditions at both UBATH and POLIMI (Morosi et al., 2018b)

Figure 63 shows the accuracy of Scale as a percentage of the total canvas size across all the conditions for both the experiments undertaken at POLIMI and UBATH. At both test locations, the results for condition A showed higher levels of accuracy and reduced variance, when compared to results from condition B4. This is particularly true for the experiments conducted at POLIMI. The fact that the accuracy decreased in the B4 condition clearly indicates that the participants were not influenced by a learning effect when completing their second condition.

The results collected from conditions B1, B2, and B3 are similarly inconclusive. While at both testing locations the B2 condition appears to have provided good levels of accuracy, there is discord between the data collected at UBATH and POLIMI regarding the accuracy of conditions B1 and B3. The data from POLIMI indicates that condition B1 was the most accurate condition, whereas the data collected at UBATH shows this as being the least accurate and with the largest variance. The results from UBATH are also at odds with those from POLIMI regarding the variance in condition B3. UBATH had the least variance in this condition, but POLIMI had the most.

Of particular interest is how the participants at UBATH seemed to be more accurate and have less variance in conditions A and B4, when compared to participants from POLIMI. However, the results from conditions B1, B2, and B3 show participants from UBATH were consistently less accurate and showed greater variance in their accuracy when compared to participants from POLIMI. This is likely a result from the type of tracking and projection technology used, highlighting the inaccuracy of the UBATH setup, where scale is concerned. This could be due to the setup itself or its calibration.

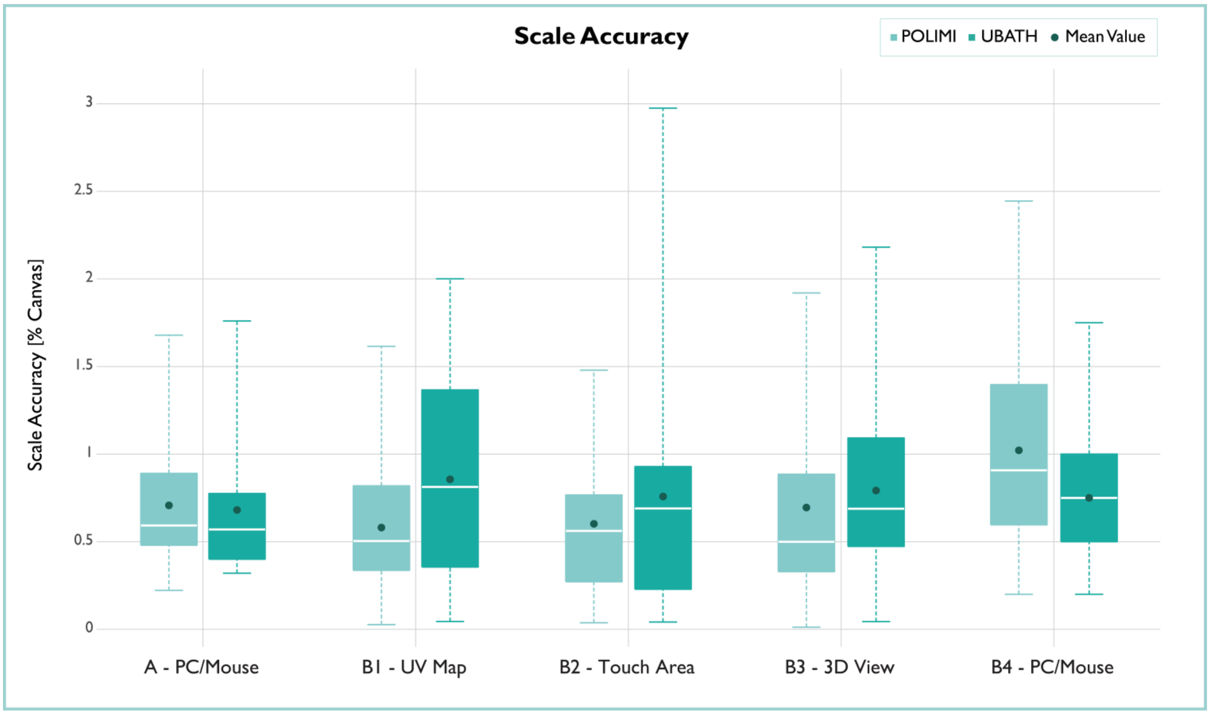


Figure 63. Box plot showing scaling accuracy as a percentage of canvas size for UBATH and POLIMI. The canvas size is equal across the layouts and refers to the area given to place the assets (Morosi et al., 2018b)

6.1.4.3 Participant Efficiency

Figure 64 shows the time taken at UBATH and POLIMI to complete sessions across all conditions. The graph illustrates that both UBATH and POLIMI participants were significantly faster at completing the task in the B4 condition when compared to the time required to complete the task in condition A. This is very likely due to the learning effect. As conditions A and B4 are identical, the difference in the layouts provided is minimal, and the task itself is fairly repetitive, meaning it is likely that participants were quick in determining how to use the interface to recreate the new layout they had been provided. The reduced time taken to complete the task may also provide an explanation for the previously unexplained phenomenon that showed reduced accuracy of placement, rotation, and scale for the B4 conditions. It is possible that the participants, faced with a similar task as in the just completed condition A, find themselves less engaged in the task and perform it by rote rather than attempting to take greater care. This results in reduced accuracy but also reduces the time taken.

While the results for condition A and B4 are consistent across both universities, the results from conditions B1-3 are much less harmonised. Not only are there significant discrepancies between the times taken for each of these three conditions at either university, but it is also less clear which condition is most advantageous. Based on the data collected it is not possible to claim one condition as superior to the others. Results from UBATH suggest that condition B1 shows the fastest average completion times yet has the highest variance. Furthermore, the discrepancy between the data collected at UBATH and at POLIMI is greatest for this condition. Results from condition B2 show the most overlap between POLIMI and UBATH. Of the three conditions (B1, B2, and B3) this is the one with the least variance for UBATH. The data from condition B3 shows the lowest variance of any of the three conditions for POLIMI, as well as the lowest average. When comparing the results from conditions B1, B2, and B3 for a single university, it is hard to claim a significant difference between any condition; as a result, it is hard to draw any solid conclusions based solely on this data. It thus stands to reason that the type of interface may not play as large a role as initially thought, so long as the necessary functionalities are present.

The discrepancies between universities can be, in part, explained by the difference in time taken in conditions A and B4. Participants from UBATH were significantly faster at completing both conditions, in particular condition B4. This may indicate that participants from UBATH were faster at acclimatising themselves with the interface and applying all its functions to the task at hand. It is likely, however, that the differences in setup between the two universities also played a role in the discrepancy. It may be that, as the participants at UBATH were more constrained in their interaction with the physical prototype, they perceived the task as being less complex and thus experienced less doubt. The participants at POLIMI had the option of translating and rotating the physical prototype in three dimensions and may have had more difficulty setting up before they felt comfortable beginning the exercise.

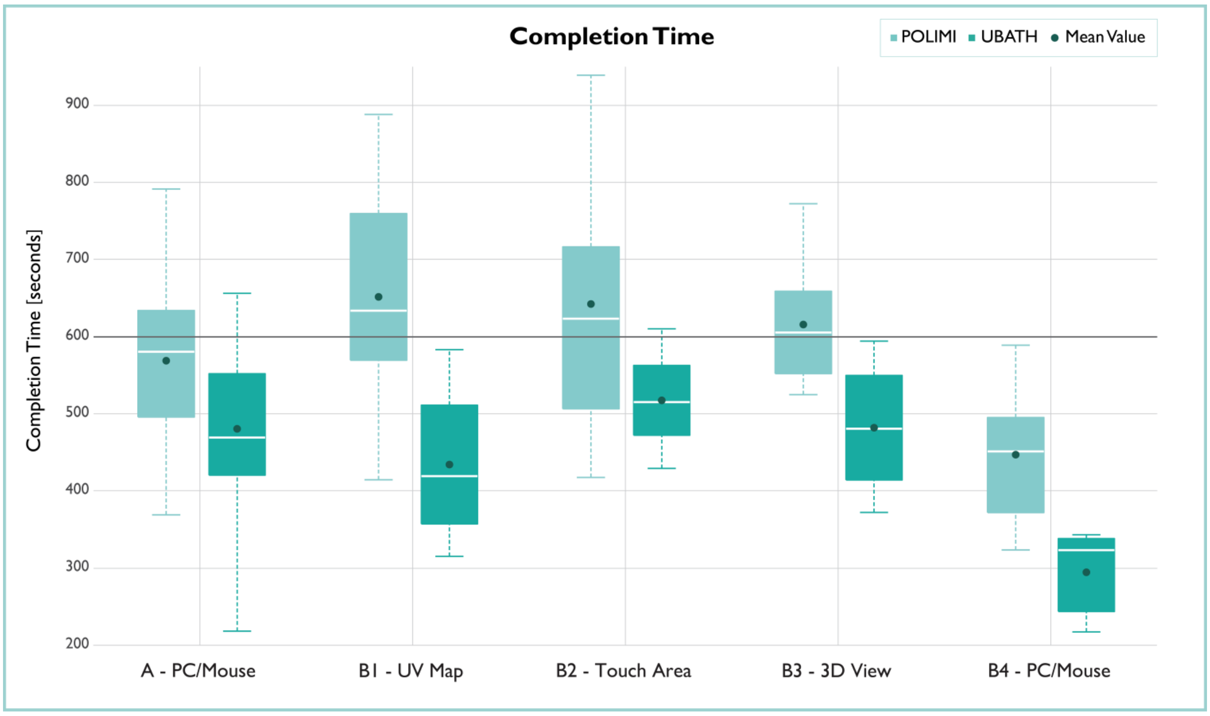


Figure 64. Time taken to complete sessions, in seconds, across all conditions for UBATH and POLIMI (Morosi et al., 2018b)

This last point is supported by Figure 65. This figure shows the average time spent on each of four activities for both UBATH and POLIMI. As can be seen from Figure 65, all participants spent the vast majority of their time interacting with the UI. This is to be expected given the nature of the task, in particular because the participants would have to spend time identifying the correct assets to place. Furthermore, participants in condition A were much less familiar with the task and this may have caused them to be more explorative as they attempted to understand all the functionalities of the system. This would explain why participants in condition A spent, proportionally, more time on activities other than the UI when compared to all other conditions. Once the participants were more familiar with both the system and the task, they could work in a more targeted manner to complete their task, clearly demonstrating the presence of a learning effect.

It should also be noted that the participants from UBATH spent more time, proportionally, on the UI than participants from POLIMI. This is particularly true for conditions B1-3. This phenomenon may be explained by a number of different factors. Firstly, in condition A, participants from UBATH were more explorative, spending a larger percentage of time on virtual prototype manipulation when compared to participants from POLIMI. This higher level of exploration might have allowed the participants from UBATH to feel more confident when working in their second condition; thus, focusing only on those activities required to complete the task, and exploring the functionalities of the system less. One additional consideration is that the system used at UBATH only allowed the physical prototype to be rotated along its z-axis. This limited range of movement may have deterred participants from interacting with the physical prototype, due to its limited utility. Thus explaining why the interaction therewith was so limited when compared to participants from POLIMI. Furthermore, the method by which the physical prototypes were rotated and translated varied greatly between the two conditions. Participants at UBATH could only rotate the physical prototype by rotating a turntable. Participants at POLIMI were able to freely pick up and translate/rotate their physical prototype. The difference in the ease of this interaction likely also contributed a role in deterring participants at UBATH from interacting with the physical prototype.

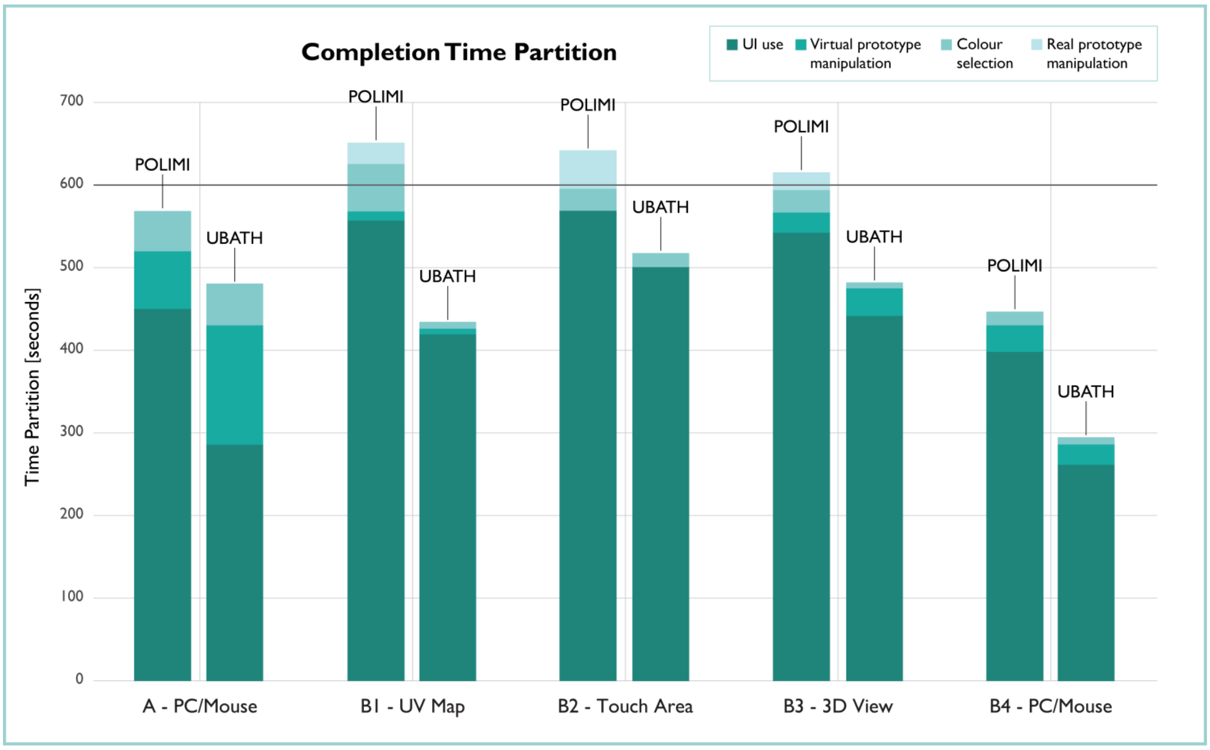


Figure 65. Average time spent, in seconds, on specific activities across all conditions for UBATH and POLIMI. The different activities are divided into four categories: "UI use", "Virtual prototype manipulation", "Colour selection" (for background), and "Real prototype manipulation". The line at 600s represents the nominal time limit for the session (Morosi et al., 2018b)

6.1.4.4 Participants' Perception of Usability

The SUS surveys were conducted at all three participating universities (GINP, POLIMI, and UBATH). However, researchers at GINP did not administer the SUS survey to participants after the completion of condition B4 as they reasoned that, since the experiment was identical to condition A, this was unnecessary. Furthermore, unaware of the failure of the logging system at GINP, the researchers elected to omit the CSI survey. The reasoning was that preliminary studies had shown the similarities of the results between the SUS and CSI scores. This was aimed at streamlining the process to be able to run the experiment with more participants.

Figure 66 shows the results collected from all three sites. The graph is spilt into three sections (“Excellent”, “Good”, and “OK”). These divisions are based on the work of Bangor *et al.* (2009) and provide a guideline to contextualise the results obtained from the SUS surveys. In their paper, Bangor *et al.* (2009) had attempted to provide descriptors to aid in contextualising the scores obtained from SUS surveys in order to aid with the understanding of the values obtained. As part of their work, they were able to find that descriptors such as “Excellent”, “Good”, and “OK” correlated with SUS scores of, on average, 85.5, 71.4, and 50.9 respectively. These descriptors were added to Figure 66 to aid with digesting the information presented in the graph by contextualising the results.

The results obtained are encouraging yet somewhat mixed. While the mean for every condition did not go below the “OK” threshold for any condition at any testing location, the confidence interval from the results at UBATH does dip below this level in two conditions: B1 and B2. However, it should be noted that the confidence interval for UBATH was quite large for all the test conditions, in particular B1, B2, and B3. The results from the test locations where the confidence interval was smaller return more encouraging results.

The results for conditions A and B4 are quite high, near or at the “Excellent” level. The fact that this did not decrease for condition B4 shows participants were fairly certain of their level of satisfaction and that no learning effects were at play. This should however not be overly surprising; the participants were all well versed in the use of CAD software. As a result, it is likely they found the setup quite familiar and easy to access. However, it is interesting to note, that, overall, the participants preferred the setup from conditions A and B4 above all others. In particular, it is worth noting that the B3 condition, which closely resembles the setup of the A and B4 conditions (as a virtual 3D model is available to rotate in the tablet) scored lowest. This could be attributed to the participants being more accustomed to a mouse and keyboard setup compared to a tablet setup. Using their fingers to directly control the interface, and thus obstructing their view of the 3D model, may well have been perceived as disruptive. Additionally, the observers noted that the participants struggled with orbiting the digital 3D model, and this may have contributed to their frustration.

The lower results seen in the UBATH cases, when compared to the results from POLIMI and GINP, in particular for conditions B1 and B2, are likely attributable to the different tracking technology used. These two conditions required the participants to interact more with the physical prototype, due to the reduced number of degrees of freedom as well as the increased difficulty; this interaction may have overly frustrated the participants.

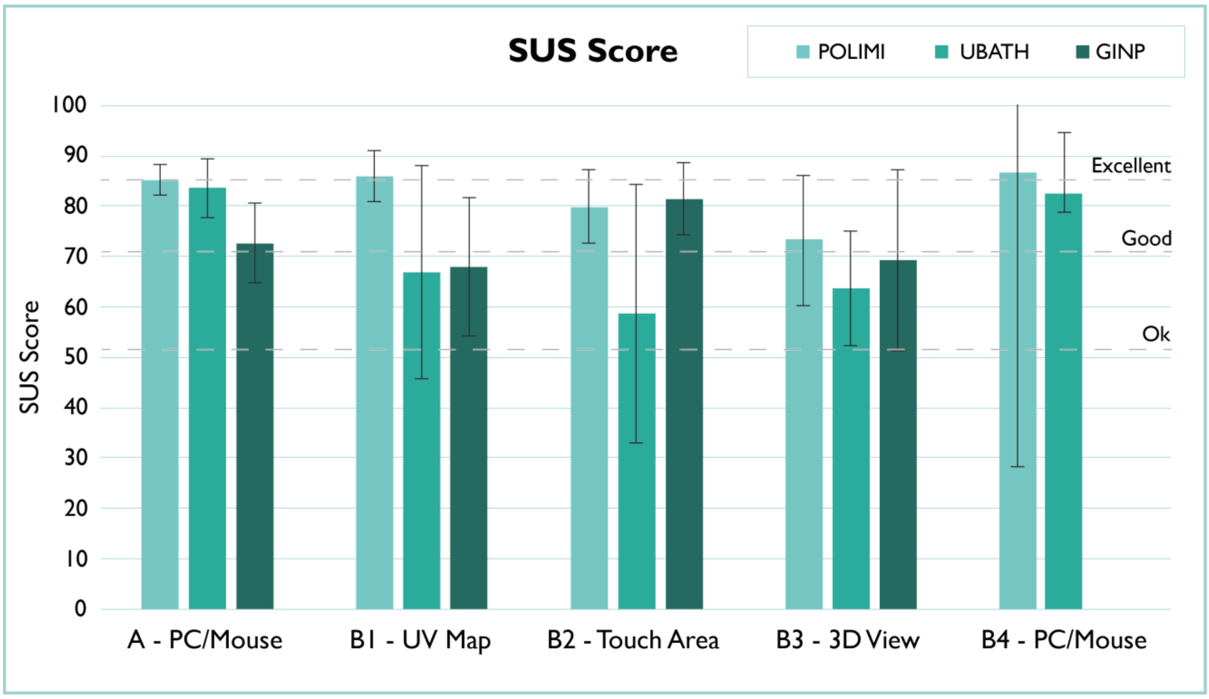


Figure 66. Mean SUS scores across all universities with confidence intervals (Morosi et al., 2018b)

The results from the CSI surveys, shown in Figure 67, are consistent with those of the SUS surveys (Figure 66). Again, it appears that conditions A and B4 (PC/Mouse) have scored the highest with condition B3 (3D View) scoring quite low. However, B3 (3D View) did score better, according to the CSI survey, than B2 (Touch Area) for POLIMI. This is in contrast to the SUS survey which showed the B2 (Touch Area) condition as being preferred by POLIMI participants. This mirrors the results from UBATH for the SUS but, due to the large confidence interval in the UBATH results, it is hard to draw firm conclusions. Overall, the results from the CSI survey serve to show that there is a consistent pattern between UBATH and POLIMI. At both locations, participants ranked the conditions in the same order: preferring B4 (PC/Mouse) the most and B2 (Touch Area) the least. This is highlighted in Table 27.

Table 27. Participant rankings of test conditions from best to worst according to CSI Scores

RANK	CONDITION
1	B4 (PC/Mouse)
2	A (PC/Mouse)
3	B1 (UV Map)
4	B3 (3D View)
5	B2 (Touch Area)

While participants at POLIMI gave higher scores in all the conditions, the discrepancy between the two testing locations seems relatively constant across all conditions.

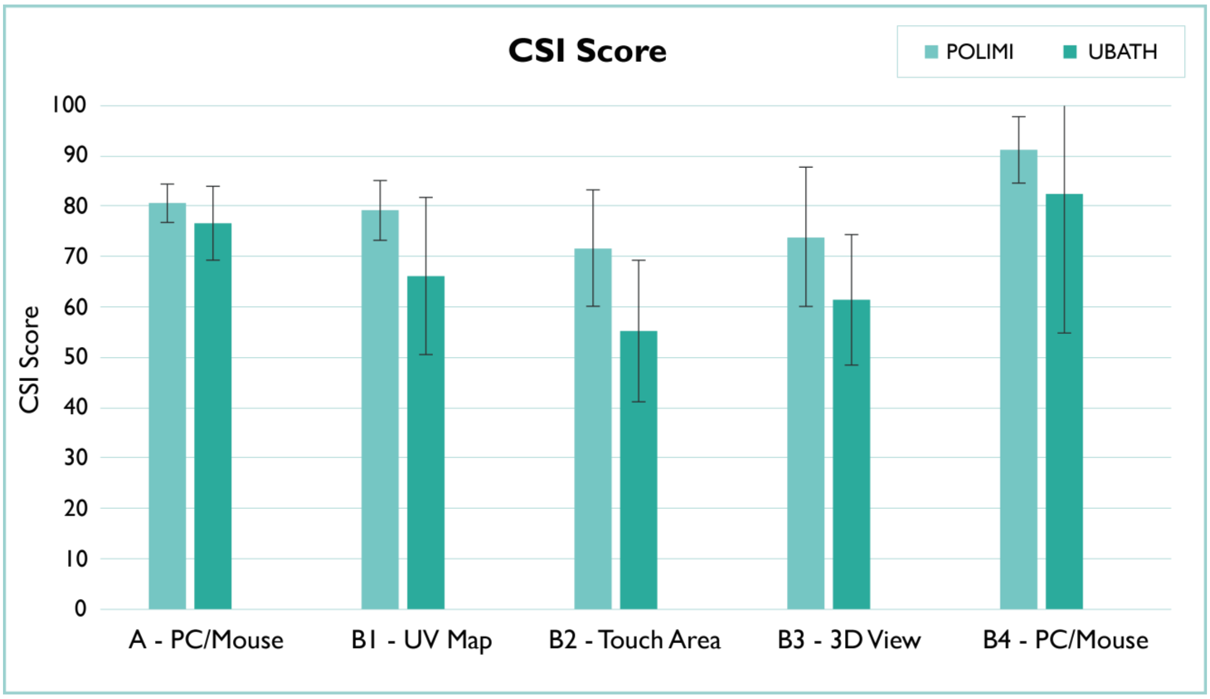


Figure 67. CSI scores across UBATH and POLIMI with confidence intervals (Morosi et al., 2018b)

Overall, based on the results from both the SUS and CSI surveys, it is clear to see that the participants ranked the usability of conditions A and B4 highest. Out of all the tablet-based interfaces, B1 (UV Map) was the preferred option whilst the B3 (3D View) was the least. This dislike may be linked to the participants' difficulty with rotating the digital model as well as the inherent difficulty of viewing the screen whilst making modifications.

6.1.5 Summary

This section presents the results of an experiment to better inform the development of an SAR system. In particular, this study sought to address RQ-3: "How do specific SAR characteristics and features affect co-design sessions' efficacy?". A total of eighty-four participants, across three institutions, took part in the experiment which investigated the impact of SAR interaction modalities on the ability to complete a replication design task. The experiments assessed the participants' perception of usability, the participants' efficiency, and the accuracy with which the design task could be completed. These interaction modalities were cross referenced against a benchmark consisting of a simple computer and mouse interface, the most common alternative.

The main findings made as a result of these experiments were:

KF-10. Accuracy of position, rotation, and scale was comparable between conditions relying on the use of a touch interface (conditions B1 (UV Map), B2 (Touch Area), and B3 (3D View)) to that of the conditions that made use of a mouse interface.

KF-11. The discrepancy in accuracy, in particular the larger variance, found between UBATH and POLIMI could be attributed to the different SAR platform setups used at the two locations. However, the results for conditions A (PC/Mouse) and B4 (PC/Mouse), which were identical at both locations, show a larger variance at UBATH. This may indicate that the SAR platform setup is not the sole contributor to the discrepancy.

KF-12. Conditions B1 (UV Map), B2 (Touch Area), and B3 (3D View) showed comparable levels of accuracy. This indicates that accuracy is independent of the method of interaction.

KF-13. The analysis of the Participant Efficiency was inconclusive at determining the most efficient interface. There was discrepancy between the results from the two universities, UBATH and POLIMI, which renders a firm conclusion impossible. It was interesting to note, however, that participants from UBATH completed the task in less time than those from POLIMI in all conditions. Whether this is linked to the different SAR setup is, as of yet, unclear.

KF-14. By analysing the time taken and the accuracy it was possible to note that participants completing condition B4 (PC/Mouse) were less engaged in the task. This is likely linked to the repetitiveness of the task which caused them to feel uninvested and uninterested.

KF-15. The participants' perception of usability was significantly higher for conditions A (PC/Mouse) and B4 (PC/Mouse). This is to be expected as the mouse and computer interface would have been the most familiar interface with which to complete the task.

KF-16. Of the touch interfaces, condition B1 (UV Map) was rated with the highest perceived usability.

KF-17. The experiments showed that providing a training scenario to the participants to familiarise them with the interface and task did not create a learning effect that would impact the results of a comparative test. This enables future experiments to be undertaken with participants completing multiple conditions or training cases.

KF-18. The results from both the CSI and SUS surveys yielded closely matching results, with neither providing additional insight over the other. In order to reduce the time taken to administer subsequent tests, it would be advisable to utilise only one.

6.2 IMPACT OF SAR ON COMMUNICATION BETWEEN DESIGN SESSION PARTICIPANTS

The following study sought to contribute to RQ-3: "How do specific SAR characteristics and features affect co-design sessions' efficacy?". This section explores how SAR influences the communication between participants in a design session.

The data included in this section was originally published in the paper "Investigating the Impact of Spatial Augmented Reality on Communication between Design Session Participants—A Pilot Study" (Giunta *et al.*, 2019) published in "Proceedings of the Design Society: International Conference on Engineering Design" by the author.

To investigate the effect of SAR on the communication behaviour between a client and designer, a controlled study was developed to mimic a co-design activity. The investigation of client-designer interaction was based on the results of the studies detailed in Chapter 5. There the designers had always used the SAR platform to communicate with the clients, rather than amongst themselves, as designer-designer interaction was expected to require less facilitation. In addition, the structured interviews conducted as part of the studies reported on in sections 5.1 and 5.2, revealed that the designers often approach a design session with a present plan for the session that they have already agreed upon amongst themselves. The design session, for the designers, is an opportunity to gather feedback and better dial in on the clients' expectations. Various validation studies conducted seem to support the utility of the SPARK platform (Bellucci *et al.*, 2018a; Ben-Guefreche *et al.*, 2018;

O’Hare *et al.*, 2018a). Understanding which underlying characteristics of SAR may provide these benefits is thus of interest.

As discussed in section 2.5.1, the improvement in communication between design session participants could lead to improved designs session outcomes. However, it is unclear whether SAR affects the communication between design session participants.

The aim of the study was to identify whether there are significant differences in the number of interactions within the design activity that uses an SAR or non-SAR setup. If identified, then this will highlight the need for further research into the nature of this change as well as provide additional insight towards answering RQ-3. This section continues by:

1. Describing the emulated co-design activity.
2. Detailing the experimental setup and conditions along with the rationale in relation to how it enables researchers to investigate communication behaviour.
3. The data capture and subsequent post-processing of the data to provide insights into the communication behaviour discussed.

6.2.1 The Co-Design Activity

The emulated co-design activity was based on a packaging design meeting involving a client and a designer. In such a design session, the client attempts to share their packaging design idea by communicating it to the designer. The designer then attempts to recreate this idea within the packaging design software. In such a scenario, it would often be the case that a shared screen would be used to show the design to the client. This scenario was based largely on the experiences with the experiments discussed in Chapter 5.

To emulate this activity, the design sessions used in the experiment featured two participants: one representing the client and one representing the designer. The choice of using a single client and designer was made to simplify the experiment; both rendering the experiment easier to run and reducing the potential number of extraneous interactions (i.e., those not between designer and client). The client was given a packaging design that only they are permitted to see. The design represents their idea and acts as the final result they wish to achieve by the end of the design session. The designer was given a packaging design tool that only they can use within the session to create the design.

It is worth noting that this approach does not precisely mimic a design session; clients will rarely, if ever, have a precise or clear-cut idea they wish to achieve. Thus, the designer’s task often includes interpreting the vague directions of the client.

6.2.2 Participants

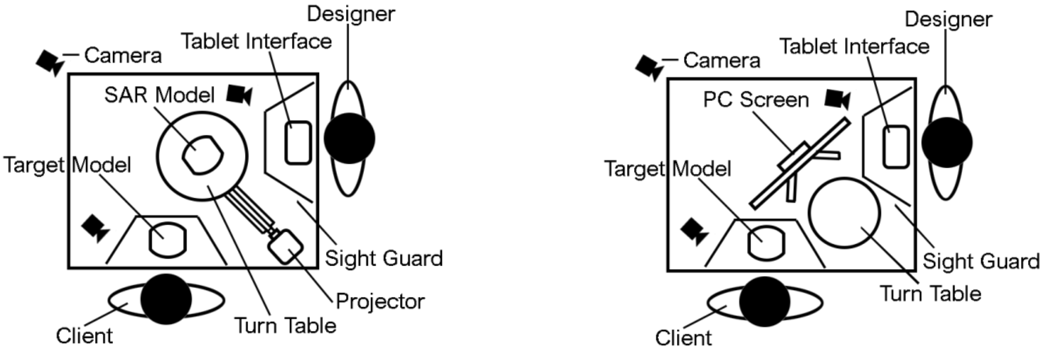
Participants were recruited from the Faculty of Mechanical Engineering and Faculty of Electrical Engineering at the University of Bath. All participants fulfilling the role of designer had experience with Computer Aided Design (CAD). No assessment was made as to the participants' knowledge or competence, but only postgraduate or final year bachelor students were permitted to participate, guaranteeing a consistent level of experience and education.

6.2.3 Experimental Setup

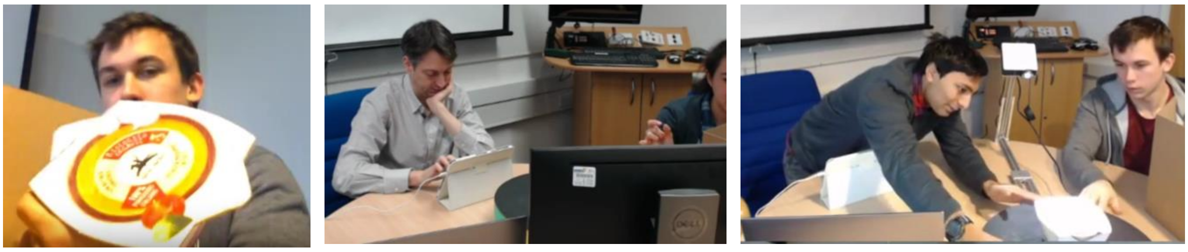
The mobile SAR research platform was used to conduct the study. Two conditions were tested: Figure 68 shows the experimental setup for Condition A, the SAR platform (Figure 68a), whereas Figure 68b shows Condition B, the traditional shared PC screen. In the former setup, Condition A, the shared representation is a physical prototype (SAR Model) placed on a turntable. The SAR projection completes the setup by projecting the images chosen using the tablet interface. In the latter setup, Condition B, the shared representation is a PC screen. In both conditions a turntable is present and is the only means of rotating the shared representation along its y-axis.

Regardless of the condition, each client is given a Target Model that is shielded from the designer's view by means of a sight guard. Each designer is given the packaging software, which is shielded from the client's view. To control the scenario further, the client was informed that they were not permitted to show the Target Model to the designer and the designer was informed that they were not permitted to show the interface of the tablet for editing the shared design representation. The Target Model was shielded in order for the client to build their own cognitive model of what they see and thus, need to communicate to the designer. The objective was to mimic the cognitive model of a client coming to a designer with an ‘idea’. The packaging software was shielded to mimic the designers’ competency in using a tool that a client would be unfamiliar with and thus, not be able to support in its use.

The objective of the session was for the designer to accurately replicate the Target Model held by the client. Video recordings were made of each session using three cameras; one per participant and one wide view camera to capture the scene as a whole (Figure 68 a, b). The experiments made use of two separate technologies: the SPARK platform and the "Observer" software, used to analyse the interactions between participants during design sessions (Ben-Guefrache *et al.*, 2018).



(a) SAR **(b) Traditional (PC)**



(c) Client looking at Target Model behind sight guard

(d) Traditional (PC) condition. Monitor is seen in foreground; behind it is the turn table

(e) SAR condition. Client and Designer in the process of turning turn table (SAR model appears white due to brightness disparity)

Figure 68. Schematics of Experimental Setup (a-b) and implementation (c-e)

Figure 68(a, e) shows the SAR condition. The designer was provided with a tablet (Figure 69a) that controlled the SAR system that was placed between the designer and client. The SAR system could be seen by both participants and was capable of projecting images onto the physical model fixed to the top of a turntable. When the turntable was rotated, the images projected appeared to remain static in relation to the model itself through rotational tracking. Both client and designer were informed that they could interact with the turntable and the SAR model as they saw fit. The tablet the designer was provided with contained all the necessary digital art assets to successfully obtain the desired final result, as well as additional spurious art assets. These spurious assets were added to prevent the designer and client from simply iterating through the list of assets to complete the task. By including assets that closely resembled, but did not match, those on the Target Model, the client and designer would have to communicate and evaluate the assets together, rather than simply having the client ask for one asset after the other. For example, multiple different images of a tomato were included in the session, but only one of the images was the correct one shown on the Target Model. This prevented the client from simply saying: “place a tomato here”.

Condition B used a conventional PC screen, as shown in Figure 68(b, d), to display the shared design representation between the two participants. As with the SAR condition, the designer was provided with a tablet containing the same set of assets. Both the client and the designer could freely rotate the digital model along the y-axis. This rotation was achieved by means of a turntable placed in front of the screen. Rotating the turntable turned the virtual model of the packaging displayed on the monitor. Both participants were informed that they could interact with the digital model on the monitor and the turntable as they saw fit. The experiment was controlled by keeping the Target Model and design software constant for all design sessions (Figure 69).

(a) Tablet Interface used by "Designer" (Morosi *et al.*, 2018a)

(b) Target Model used by "Client"

Figure 69. Items provided to "Designer" and "Client"

Due to the reuse of assets and the final design to be recreated, participants were only permitted to join one design session in order to avoid potential learning effects. Additionally, at the start of each session, the designer was allowed some time to familiarise themselves with the tablet interface that would be used to manipulate either the SAR model or the 3D representation. The amount of this practice time provided was not set but rather, until the designer felt comfortable using the interface. A different model with different assets was used for training.

6.2.4 Data Processing and Analysis

To identify whether there are significant differences in design activity interactions, the communication transactions between the client and designer were examined. The framework defined by Ben Guefrache *et al.* (2018) was used to give an insight into the *types* of interactions occurring between the session's participants. The efficiency of communication analysis involves the examination of the overall *time taken* to complete the session. These two techniques are now discussed.

6.2.4.1 Interaction Analysis

Scientific literature has analysed the interactions in design and co-design sessions and has predominately relied on the protocol analysis method (Cross *et al.*, 1996). Protocol analysis is based on the analysis of verbalisations and/or gestures associated with these verbalisations.

In order to understand the design practice that occurs in co-creative sessions, the different typologies of interactions between participants need to be observed and analysed. The interaction-centric framework used (Ben-Guefrache *et al.*, 2018) is based on the capture of verbal and non-verbal interactions between the participants and the materials used in the sessions (such as physical prototypes, digital, etc.).

The coding scheme is based on a number of elements: the client(s) and designer(s) who form the Actors, and the interaction(s) that occur between them. The interactions that can occur between Actors can be classified as: Verbal, Digital, Mixed; or Ephemeral (Figure 70). The analysis of the participants' interactions provides insight into the ability of the client and the designer to effectively share ideas and move closer to the desired final result.

(a) Interactions present in SAR sessions **(b) Interactions present in traditional (PC) session**

Figure 70. Interaction analysis framework

The four types of interactions are shown in Figure 70. Coding of each interaction begins when either participant displays one of the following behaviours whilst simultaneously speaking with their counterpart:

The 'Verbal' interaction between the participants is a type of interaction that is not supported by any other means. That is to say that the participant only speaks but does not rely on any other medium to communicate.

The 'Digital' interaction category includes any kind of representation displayed on a screen, such as a presentation on a personal laptop, tablet or any information shared from phones. Verbal exchange can still be present in this category, but the participants rely on a screen to support their verbal exchange.

The 'Mixed' interaction is described as including a physical prototype (physical mock-up with a predefined shape, mostly 3D printed) on which digital elements like pictures, images, textures, or text are projected through the mean of an SAR system. It should be noted that that the act of manipulating the tablet interface (which controls the SAR/PC Screen system) is considered as a mixed interaction. Verbal exchange can still be present in this category, but the participants rely on the SAR model to support their verbal exchange.

Finally, 'Ephemeral' refers to interactions that include gesticulation used instead of, or while, speaking with the purpose of communication. In addition, gestures made in the air made in order to mimic form or explain an idea are included. The person making the gesture can depict or mimic an object (shape, volume, surface), a usage (function in a specific context) or a behaviour (deformation of an object, simulate flashing lights etc.).

The interactions were captured through an on-the-fly method, which aims to provide a quantitative description of the interactions made by the participants (designers and clients) during a co-creative design session.

The on-the-fly method involves two phases: first, two coders in the experiment room code the live session. One is charged with identifying the actor (designer or client) who has initiated the interaction and a second coder identifies which type of interaction (Verbal, Digital, Mixed, or Ephemeral) occurred. This is achieved through a software tool called ‘Observer’ and coders are trained in advance from a coding book which provides a set of coding rules (Ben-Guefrache *et al.*, 2018). The second phase is dedicated to the analysis of the data gathered during the session in order to obtain a quantitative representation of the interactions that occurred during the co-design session. This is done by analysing the percentages of interactions initiated by the designer and client as well as the percentages of the types of interactions. The use of a meta-analysis such as this is more appropriate for a higher-level analysis, such as the one presented here, due to the simplicity of the implementation. This approach offers considerable savings in manpower and time that would otherwise be needed to transcribe (and potentially translate) each session while still capturing the data necessary to evaluate the interactions occurring between the participants. Furthermore, the (non)verbal interactions can provide an approximation of the number of communication transactions that occur in the session.

6.2.4.2 *Efficiency of Design Activity*

The efficiency of the design activity was also assessed through a comparison of the time taken to complete the activity for each condition. The start of the session was taken when either participant began speaking to the other about the design task. The session was ended when the client determined the representation was sufficiently close in appearance to the Target Model. No specific margin of error was set for the participants to have to fall within, but the observer in the room checked for completeness by visually inspecting the final model generated by the participants. The times for each session were then presented as a box-and-whisker plot to illustrate the differences in the time taken between the two conditions. Due to the controlled nature of the study, the only barrier to successfully completing the activity was the ability of the client to successfully communicate the design of the Target Model to the designer and the designer's ability to interpret and query these instructions. It then stands to reason that, the time taken to complete the task would be indicative of the ease with which the participants were able to communicate their intentions. It should be noted that, as with most experiments involving human participants, it is not possible to control every variable. As noted in section 6.1.5, other factors will invariably influence the amount of time taken by the participants, such as their engagement with the task. However, the more controlled nature of the experiment, and the use of a larger sample size when compared to the experiments reported on in chapter 5, enables a more precise analysis of the impact of the communication on the outcome of the design sessions.

6.2.5 Results

Fourteen participants were recruited for the experiments for a total of seven sessions. Of the seven, four sessions were undertaken using the SAR model and three used the 3D digital representation (Table 28). In all but one session, participants managed to replicate the model provided to the observer's satisfaction.

Table 28. Participant and session breakdown

TYPE	NR	PARTICIPANT INFORMATION	NOTES
SAR	**1**	Male “client” and male “designer”	
	2	Male “client” and female “designer”	Incorrect asset selected
	3	Female “client” and male “designer”	
	4	Female “client” and male “designer”	Used pen and paper to support communication (omitted from interaction analysis)
PC	**1**	Female “client” and male “designer”	
	2	Male “client” and male “designer”	
	3	Male “client” and male “designer”	

Session SAR-02, which used the SAR representation, selected a similar, yet incorrect, version of one of the logos. The incorrect selection can be attributed to the low quality of the image projection on the top surface of the SAR representation. The resolution of the images projected was not sufficiently high, making the small text written on some of the assets hard to read. This was particularly true when the assets were scaled down to a small size. This made it difficult for the participants to read the text and identify that the logo they had selected, while the same shape, size, and colour, had different text. Nonetheless this data was included in the analysis, as the participants otherwise completed the task, however this failure to identify the correct asset was noted in Table 28.

It should also be noted that session SAR-04 was omitted from the interaction analysis due to a limitation in the coding methodology as the participants began to use pen and paper to support the design session, this is highlighted in Table 28. While this has no great impact on the efficiency results, the interaction analysis cannot differentiate between this type of interaction and the interaction with the physical model by the client. Lastly, it should be noted that some of the sessions suffered from interruptions due to technical issues. On occasion the platform froze or refused to acknowledge a command from the designer. The timing was paused during these interruptions.

6.2.5.1 Interaction Analysis

Figure 71 shows the percentage of interactions initiated by designers and clients within each session in order to evaluate and compare their participation. Clients perform a higher percentage of the interactions in all three SAR sessions (84-94.2%) with designers initiating between 5.8% and 16%. These results are to be expected because the scenario places a considerable emphasis on the client's ability to communicate the contents of the Target Model to the designer. On average clients in the PC sessions initiated 86.1% of interactions whereas, in the SAR sessions, the clients initiated 88.8% of interactions. This was not deemed to be a significant difference in the percentage of interaction across the two technologies.

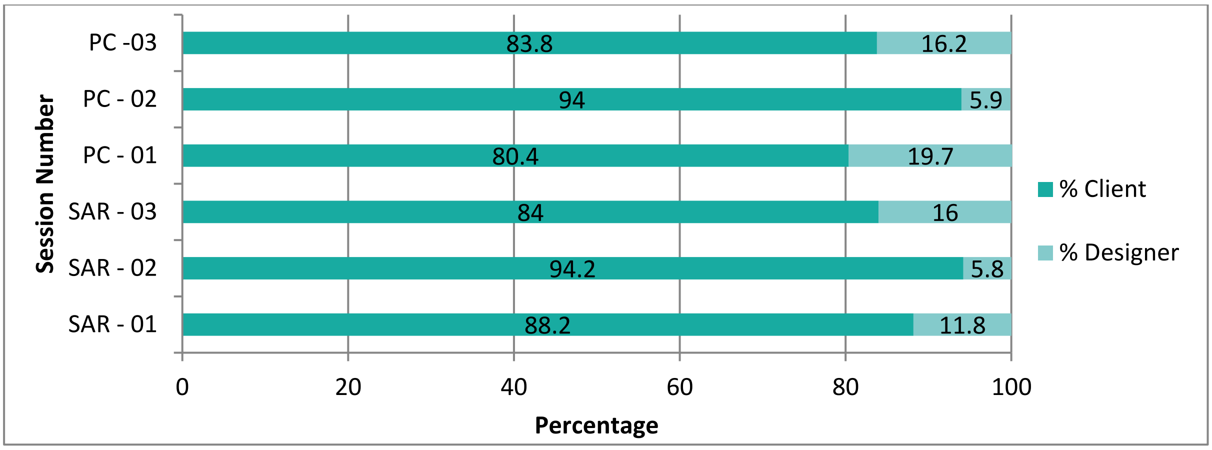


Figure 71. Interactions Initiated by Client or Designer for PC and SAR Sessions

The proportions of each interaction type for each session are calculated in Figure 72. In both the SAR and PC sessions, the majority of interaction took place using the shared model representation. For SAR sessions, an average of 61% interactions were mixed. Similarly, for the PC sessions, 55% of interactions were digital. It is worth noting that a higher percentage of mixed interactions (SAR sessions) occurred when compared to digital interactions (PC sessions). This is an encouraging result that demonstrates the salience of mixed interactions during these sessions. Furthermore, there is a surprisingly high use of verbal interactions. This accounted for up to 29% of the total interactions in the SAR sessions and 38% in the PC sessions. Gesturing in the air (ephemeral interactions) is the lowest observed interaction in both conditions and, as in other experiments (O'Hare *et al.*, 2018a), appears to be used as a way of actively expressing ideas and support during the discourse.

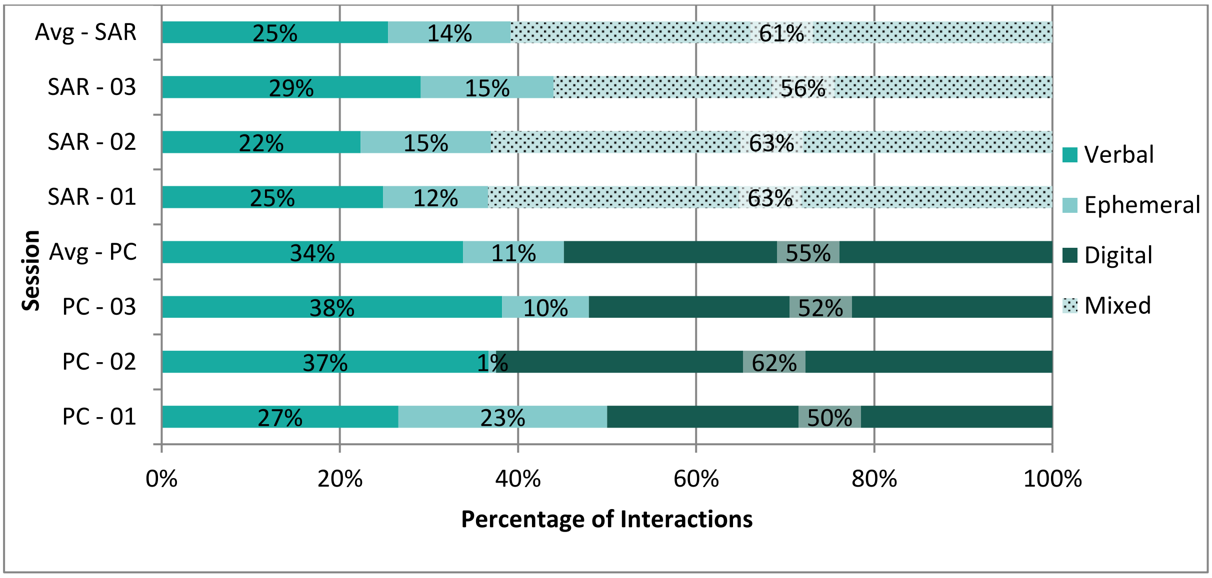


Figure 72. Proportion of each Interaction Type for SAR and PC

6.2.5.2 *Efficiency of Process*

Table 29 shows the time taken for each session to be completed as well as the experimental condition. Figure 73 consolidates this information into a box-and-whisker diagram.

Table 29. Time Taken and Number of Interactions for Each Session and Condition

	SAR				PC		
Session Number	**01**	**02**	**03**	**04**	**01**	**02**	**03**
Time Taken	9:52	23:41	21:12	16:39	26:38	17:15	28:03
Mean	17:51				23:59		
Median	18:56				26:38		
Standard Deviation (σ)	06:04				05:52		

As can be seen from Figure 73, the mean time taken for SAR sessions was lower than that of the PC sessions by 6 minutes and 8 seconds. Figure 73 also shows that there is relatively little overlap between the SAR and PC sessions suggesting a potentially significant difference. Increasing the sample size would confirm this. Nonetheless, it highlights that SAR reduces the time taken for the exercise.

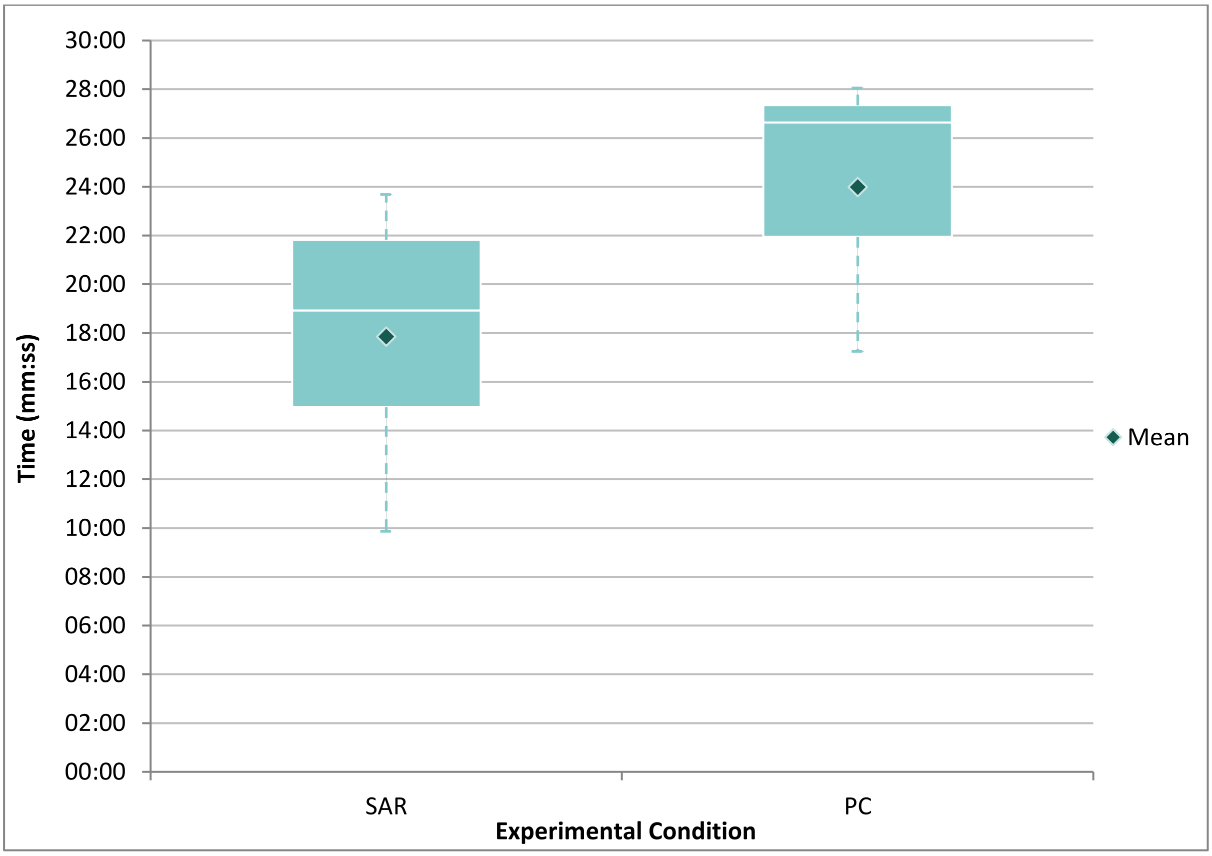


Figure 73. Comparison of Time Taken to Complete Task

6.2.6 Observations

As first mentioned in Table 13 and Table 15, section 6.2 attempted to address RQ-3: “How do specific SAR characteristics and features affect co-design sessions’ efficacy?” by means of achieving RO-4: “Analyse the impact of a sample of specific characteristics and features of the SAR platform on co-design sessions”. The specific characteristic analysed in this section was that of communication between design session participants, comparing the results from an SAR supported design session to those from a PC supported session.

The results of the efficiency of the process analysis, laid out in section 6.2.5.2, lends credence to the claim that SAR facilitates the communication of ideas between design session participants. The controlled nature of the task implies that, to successfully complete it, participants must be able to communicate efficiently. It then stands to reason that a reduction in the time taken to complete the task reflects an improvement in the ability of the participants to communicate with one another, as was the case for the SAR sessions.

One noteworthy result was that from session SAR-01, as shown in Table 29. This session took significantly less than any of the other sessions, both SAR and PC, being the only session to successfully complete the task in under ten minutes. It is not clear what led the participants of the session to complete the task so quickly. An informal, qualitative, analysis of the session showed that the participants quickly began to use the SAR platform to display the assets, iterating through them quickly to recognise the necessary ones. This provides additional support to the theory that SAR assists in communication.

The results from the interaction analysis, discussed in section 6.2.5.1, seem to suggest that SAR better supports communication between participants as the percentage of interactions involving the shared design representation is higher. As was noted in Figure 72, 29% of interactions in the SAR sessions, and 38% in the PC sessions, was verbal. This discrepancy may potentially be explained by the decreased support that the PC provides when compared to the SAR system, forcing the participants to rely on describing in greater detail their desired results.

This is further supported by the observation that the percentage of ephemeral interaction remains similar between PC and SAR sessions, but the amount of verbal interaction increases in the PC sessions. This seems to support the theory that the client does not compensate for the lack of support the PC screen offers in communicating their ideas to the designer by using hand gestures but rather, through verbal interactions. This suggests that SAR better supports the communication between the participants.

The study does have some limitations, such as the failure of SAR-02's participants to select all the correct assets, is in of itself not a major issue, as the asset they selected was very similar in appearance and had the correct scale and rotation. It is possible that this issue may have been caused by miscommunication between the participants. However, it betrays a more serious underlying problem: the user interface on the tablet and the resolution of the projector is deficient and causes the participants to make simple mistakes. This is substantiated by the findings highlighted in section 5.2.4.

Having the participants fill in a System Usability Score (SUS) after completing their session could shed more light into their confidence in using the platform. As the poor user interface was mentioned by many of the designers in the post-session discussions as a point of frustration, an SUS would allow for normalising the results across different sessions. Furthermore, the SPARK platform suffered stability issues, with the system crashing or hanging on occasion during the sessions, necessitating a restart. This adds to the frustration of the participants and also affects the quality of the research data produced as the participants are forced to stop the session, interrupting their workflow. Fortunately, the majority of these issues are a matter of implementing bug fixes to the technology in use and have not highlighted an underlying issue with the strategy of using SAR in design.

The experimental setup implemented did not implement and/or factor in all the controls needed. One element that was not captured was the variance between clients' existing communication skills, as this influences their ability to more effectively guide the designer.

The small sample size used in this study has precluded the ability to draw statistically significant conclusions from the data. However, this analysis provides some support to substantiate the hypothesis, and certainly suggest that more research is warranted. As such, it would certainly be advantageous to repeat the study with a larger sample size. Furthermore, future sessions should attempt to gauge the ability of the participants to communicate effectively and be able to adjust for this when evaluating the results. Additionally, as there were issues with SAR-02 and SAR-04, the methodology for giving the participants instructions on what is expected of them in the session should be updated to guarantee that they are fully aware that they must review their final work and that they should rely only on the tool provided to them. Some additional method of checking the accuracy with which they replicate the design may also be warranted. The surprising result that the number of ephemeral interactions does not increase between session types, while verbal interactions do, also suggests that closer attention should be paid to analysing the richness of communication between the participants to analyse not just the number or type of interaction but also the amount of data transmitted between the participants during each interaction.

One type of analysis that was not performed during this study was log analysis. With this method of analysis, the possibility of analysing the rework done by the designer to achieve the client's demands becomes available. Gopsill et al. (2016) suggests that the use of CAD logs can be a good source of large amounts of quantitative data that can be used to determine position, rotation, scale, etc. of art assets. The large data sample would enable more sophisticated analysis of the session efficiency from primary data and create the possibility to record the type of workflow chosen (e.g.: working on one asset at a time versus placing and adjusting multiple simultaneously).

Additionally, gaze analysis was not performed in this study as the interaction analysis was deemed sufficient. However Boa and Hicks (2016) provide a solid methodology for approaching gaze analysis that might prove to be a valuable addition to supplement the data obtained through the use of interaction analysis. Of particular interest would be to cross-reference the type of interaction occurring depending on where the participants focus their gaze.

In addition to techniques that could be used to expand the data collected, it is important to focus on the expansion of the scope of the study. The methodology implemented seems to have proven effective and it now becomes necessary to expand into exploring how the communication between participants is affected in design sessions with more variables and a more naturalistic design challenge.

6.2.7 Summary of Findings

This study set out to analyse whether SAR can support communication between design session participants more effectively when compared to more traditional design representations (i.e., 3D models on a screen). The study showed how the mode of interaction between participants is affected by the type of technology used.

The key findings identified as part of this study were:

KF-19. The study provides evidence that SAR supports the communication between participants of co-design sessions with sessions noticeably shorter than sessions run using a 3D digital representation on a PC screen.

KF-20. The percentage of interaction between participants with the shared design representation was also higher in the SAR scenarios than in the 3D model scenarios indicating that the SAR system better supports communication in the shared space.

KF-21. Clients in the PC sessions had to rely more on the use of verbal cues to guide the designer to compensate for this. This may also be linked to the amount of data that can be transmitted between participants during the interaction, with the SAR model supporting larger amounts of data transmitted.

KF-22. Lastly, this study shows that the methodology applied is a valid approach for the investigation of the influence SAR has on communication between design session participants.

6.3 IMPACT OF SCALE IN DESIGN SESSIONS SUPPORTED BY AN SAR PLATFORM

The nature of SAR utilizing a physical object for projection means that, although SAR allows participants to better place an object into context, some objects will need to be scaled due to the projection envelope and resolution of the digital projection technology used. This may limit the applicability of SAR should scale influence the ability of designers to perform their tasks. Thus, raises questions as to the effects scale may have on the capability and behaviour of participants to design.

To investigate this, a study was performed that sought to evaluate the impact of scale on the generation of concepts within a co-design review session. This research contributes to the development of SAR systems for design by showing how designers and their collaborators respond to a scenario where the physical model was scaled both up and down. Scaling up was performed to overcome the limitations of the resolution of the digital projection, while scaling down was done to fit within the projection envelope. In doing so, the experiments presented in this section seek to address RO-4: “Analyse the impact of a sample of specific characteristics and features of the SAR platform on co-design sessions” and, by extension, answer RQ-3: “How do specific SAR characteristics and features affect co-design sessions’ efficacy?”

Some of the behaviours explored in this research include: how the design assets (graphics) are used at smaller scales vs larger scales; and, whether there is a difference in the number of concepts created at the different scales. Highlighting these and any other differences in behaviour between conditions will provide the field with a better understanding of how scale influences designers.

The following sections identify the parameters that limit the size of the models and studies that have demonstrated SAR’s potential to support transdisciplinary design activities. This is followed by a description of the study that has been conducted to investigate the effect of scale on a collaborative design review activity. The results of the study are presented and followed by a discussion to the underlying theory that could be developed, as well as the inherent limitations of the results presented. The key findings are then mapped back to the research questions.

6.3.1 SAR Projection Envelope Limitations

SAR's use of projector(s) enforces a projection envelope that limits the maximum size of the augmented model (Figure 74). As originally discussed in sections 4.1 and 4.3.2, there are a number of parameters that limit this envelope size. These include the projector's technical specifications (such as, but not limited to: throw, aspect ratio, focal length, resolution, and luminosity) and distance between the model and projector (limited by the space available for the SAR setup). Varying this distance will increase the projection envelope's size yet decrease the luminosity and resolution of the projection on the model. Therefore, it is non-trivial to determine the maximum and minimum projection envelope to produce an effective design environment[1].

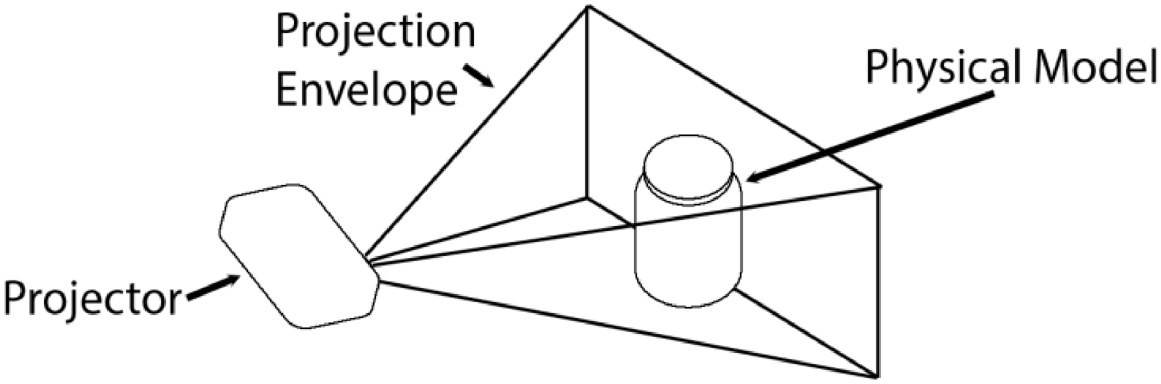


Figure 74. Projection SAR setup

Figure 75 shows an augmented prototype side by side with a regular object with the projector creating a “spotlight” effect on the augmented prototype. The black cloth denotes the projection envelope of this system. While an object this size easily fits within the projection envelope at a 1:1 scale, a larger object, such as a suitcase or a car, would struggle. Additionally, a smaller object, such as a watch would be at risk of suffering from decreased resolution.

Furthermore, environmental and technical limitations can play a role in determining the size of the projection envelope. The projector has a maximum brightness output, and the distance between the projector and the projection surface limits the amount of light reaching the projection surface, as dictated by the inverse square law. As such, should the SAR system need to be deployed in an environment with a high level of ambient light, it may be necessary to bring the projector closer to the physical model to guarantee a sufficient level of brightness for the SAR effect to be visible. However, this limits the size of the projection envelope, and thus the maximum size of the physical model in use.

---

[1] For a more detailed discussion on these factors, please see Section 4.3.2

Figure 75. A real life object (left) next to a physical model with a digital overlay (right) the tablet shows the interface used to control the SAR projection (Becattini et al., 2018)

Remaining within the projection envelope is thus imperative to SAR's function. When this area is exited the model is no longer augmented. Additionally, when a model is too small, the resolution of the projector may be insufficient to correctly display images and/or be too finicky for a designer to arrange and move graphics. As a result, scaling the model will be necessary if one wishes to use SAR, thus leading to question the potential impact that scale may have on a team's design behaviour in generating solutions for a given design problem.

6.3.2 Experimental Setup

The experiment evaluated the generation of concepts, analysing both the impact of scaling a model up or down. Two scenarios were envisioned: one where the physical prototype had to be scaled up, and one where it had to be scaled down. The scenarios consisted of two designers working to design different concepts for a watch (used to evaluate the impact of scaling up a physical model) or a race car (used to evaluate the impact of scaling down a physical model).

These particular design scenarios were chosen as they are representative of industry practices (O'Hare *et al.*, 2018b). In all the scenarios, the designers were provided with a:

- SAR platform to create, develop and evaluate designs on a physical model of a watch/race car
- Inspiration board to present the market and provide greater context to the activity
- Set of digital graphics that could be applied

The participants were all students from the Faculty of Engineering and Design at the University of Bath. All participants had experience in CAD and were either in their final year of MEng, were pursuing an MSc, or a doctoral degree. Ages ranged from 20 to 38 and the gender ratio was 44 males to 12 females.

The mobile SAR platform (miniSPARK), first discussed in section 4.3.4, was used to support these experiments. The experimental setup is shown in Figure 76. The SAR platform consisted of a projector mounted to an arm, which was subsequently attached to a base that held the physical model. The platform could track the model as it rotated; however, for the purpose of the experiment, the participants were not informed of this feature and thus did not attempt to do so. The distance from the projector to the model was kept constant across all conditions. Only the model was replaced when switching between the conditions. The digital overlay was set to project and fully cover the model with no "spillage"; thus projecting only onto the physical model.

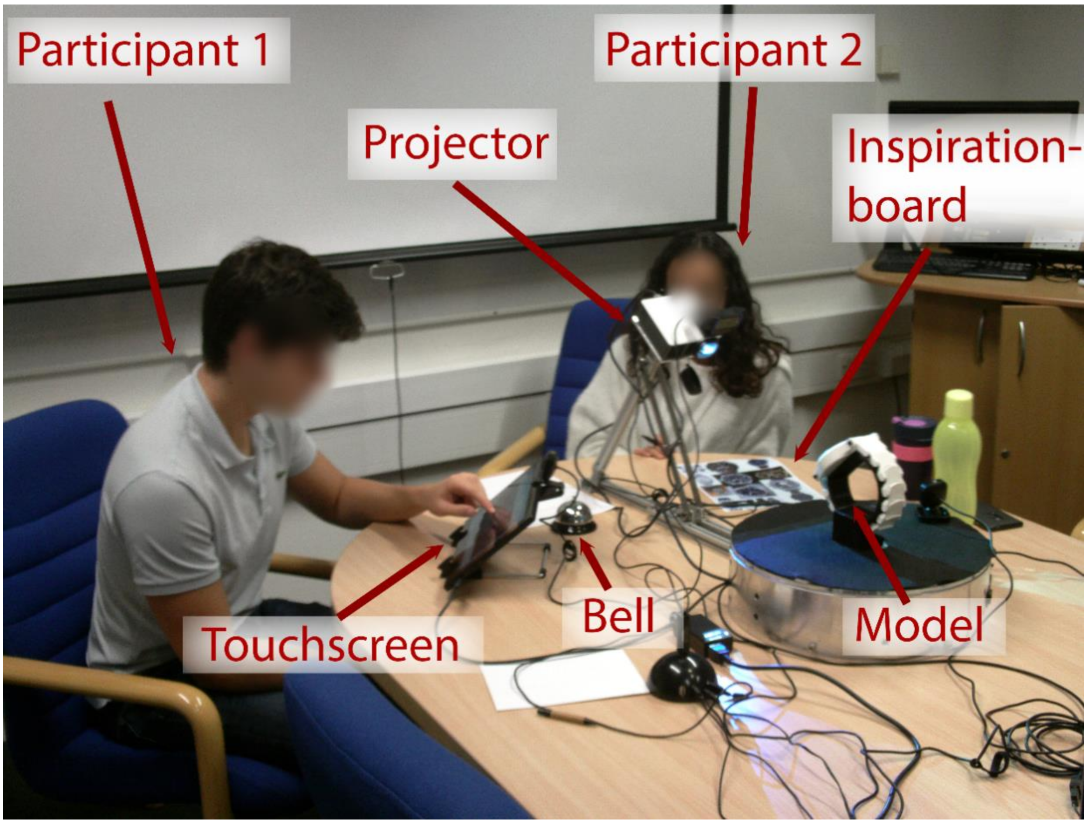
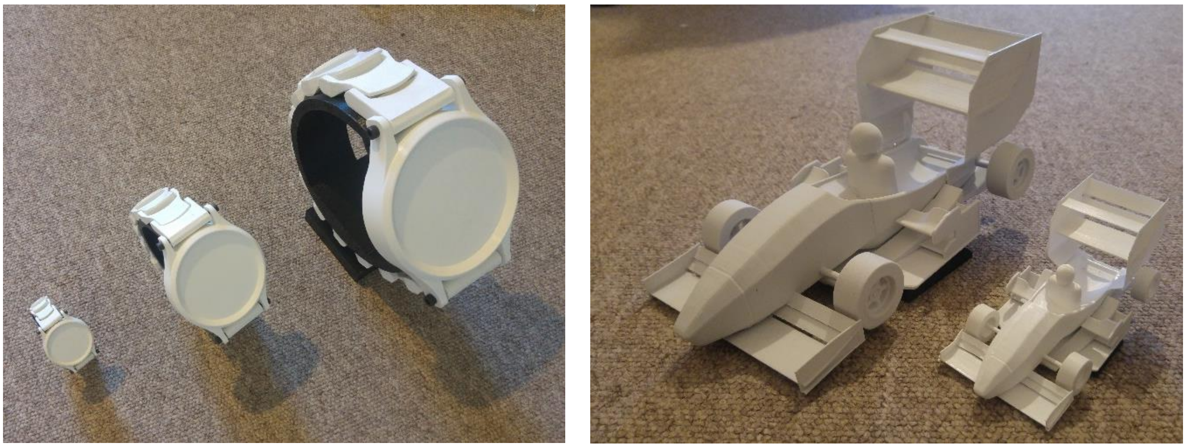


Figure 76. Experimental setup using 2:1 scale model. The model was replaced based on the scale condition being tested. Participants were only allowed to participate in one of the two conditions. These images were taken after completion of all experiments

Figure 77 shows the five physical models used across all the experimental conditions in both the scaling up and scaling down scenarios. Figure 77a shows the watch models used to evaluate the effect of scaling up, with model scales ranging from 1:1 to 3:1. Conversely, Figure 77b shows the two race car models used to investigate the effect of scaling down; the two scales used were 1:10 and 1:20.



(a) Models used in the scale up scenario. From left to right: the 1:1, 2:1, and 3:1 models

(b) Models used in the scale down scenario. 1:10 model (left) and the 1:20 model (right)

Figure 77. Models used across all experimental conditions.

The inspiration board (Appendix, C.1) used in the scaling up scenario consisted of seventeen watch images showing the watch face and strap. Only one image showed the watch being worn by a male hand. All watches were analogue with varying numbers of complications and hands, however no watch had less than one additional feature (calendar, date, chronometer, etc.). The prevalent colours for the watch face and casing were: blue, black, and gold. The watch straps showed both leather and metal straps of various colours, predominately brown and black. All images were sourced through Google Image Search.

The inspiration board used in the scaling down scenario (found in the appendix, section C.2) consisted of eleven images showing the previous iterations of the Team Bath Racing race car. Only one image showed the car without a driver seated inside it. All the race cars had been previously built and used by Team Bath Racing, between 2010 and 2019. The prevalent colours for the race cars' livery and body were: green, black, and white. All images were sourced from the Team Bath Racing website.

The graphics provided to the participants in the scaling up scenario are shown in Figure 78. Twenty graphics were provided in total, some assets (i.e., WeekCount, SecHand, MinLines, Logo, DayTime, DayCount, and Circle) are unique, where the others vary in line weight and text size.

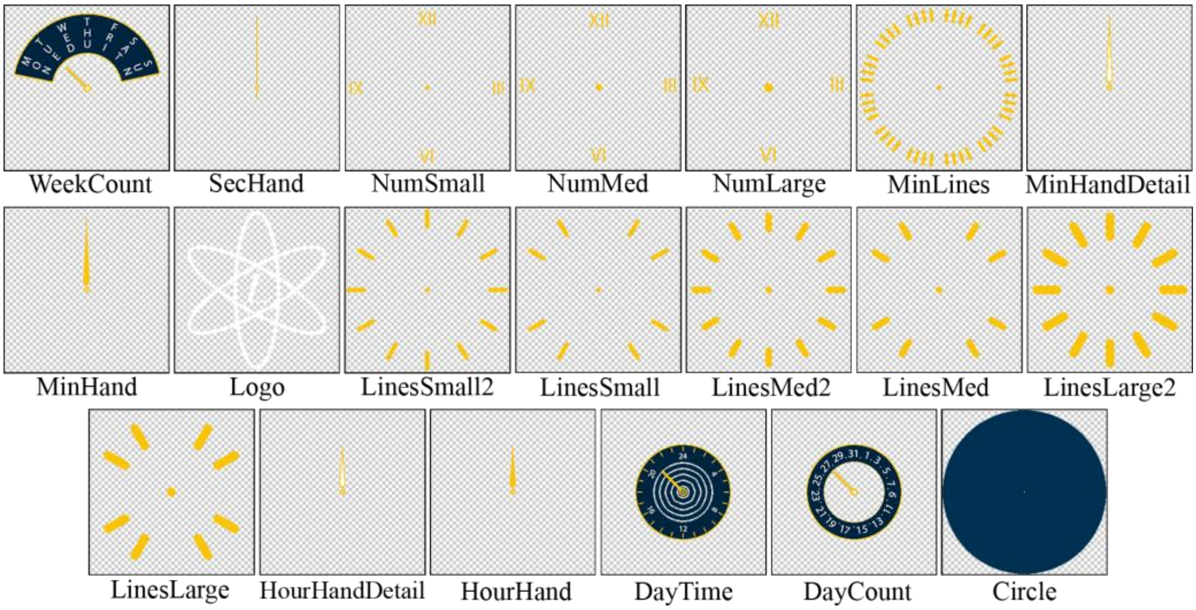


Figure 78. Graphics provided to participants for scale up scenario (watch design)

The graphics provided to the participants in the scaling down scenario are shown in Figure 79. Twenty-seven graphics were provided in total, some assets (i.e., SteelH, INA-FAG, LigthB, Castrol, ABC, 19, FreeF, Flux, EasyCom, Cosworth, G_CurveTri, G&YTri, Green, UniBath, and W_Stripe) are unique, where the others vary in orientation or colour.

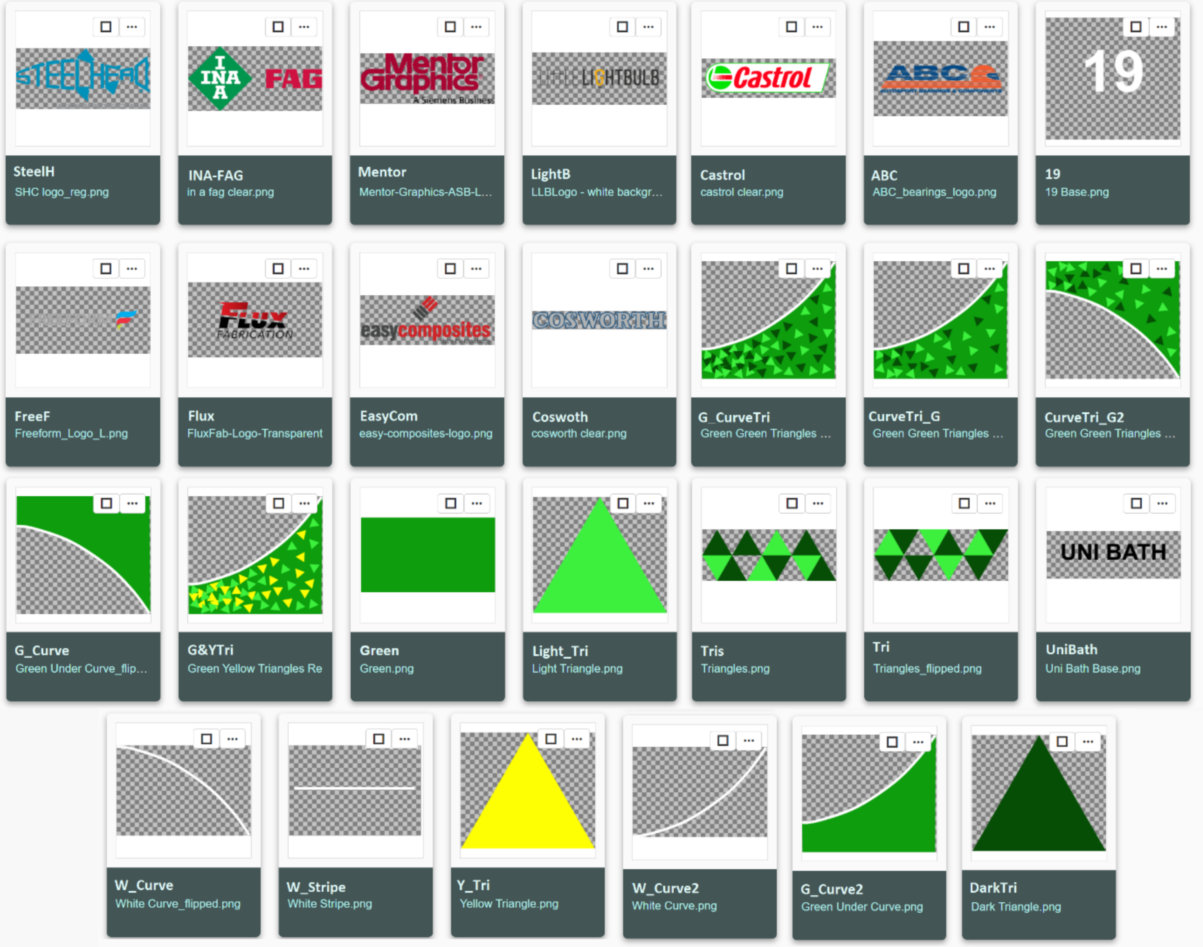


Figure 79. Graphics provided to participants for scale down scenario (race car design)

The SAR platform's tablet interface, shown in Figure 75, allows users to place, scale, rotate, and translate the assets on a digital representation of the watch. These changes are then displayed in real time on the physical model. It is worth emphasising that while all assets can be scaled to the desired dimensions by the participants, only some assets are available with thicker or thinner line weights and font sizes.

The experimental conditions were constant with the exception of the scale of the SAR model used for projection. In the scale up scenario, the first condition used a 1:1 scale model (external diameter 45mm), the second condition used a 2:1 (external diameter 90mm) model, and the third condition used a 3:1 model (external diameter 135mm) as shown in Figure 77a. In the scale down scenario, two experimental conditions were evaluated. The first condition made use of a 1:10 scale model, while the second condition used a 1:20 scale model, as shown in Figure 77b.

In the case of the scale up scenario, the hypothesis was that increasing the scale of the watch model would support a more comprehensive use of the design assets. A larger model would allow for greater exploration of the design assets, leading to more exploration. However, the larger scale may, at the same time, affect the design process due to the unnatural scale of the object. Similarly, for the scale down scenario, the hypothesis was that the larger model would lead to greater exploration. However, as both models were significantly reduced in scale when compared to the actual race car, the anticipated distortion due to the unnatural scale was expected to not play a significant role.

Each session was structured as follows: the participants (two in every session) were briefed on the task they needed to complete by means of a pre-prepared statement. A copy of this statement can be found in the appendix; in section G. In the case of the scale up scenario, the participants were told to imagine themselves as designers tasked to design a number of watches for a client in line with the inspiration board. In the scale down scenario, the client was replaced with Team Bath Racing. The participants were informed they had to redesign the livery for Team Bath Racing's new race car using the inspiration board to help guide their designs.

It was made clear that the participants could create as many designs as they felt necessary but each one would be reviewed by their "client". Each time they felt they had developed a concept they would later like to present to their client, they would have to hit a bell to record the design. It was also stressed that the "client" would be reviewing the proposed watches/liveries by means of the SAR system. Thus, in the event of any discrepancies between the view on the interface and the SAR system, the SAR system would take primacy.

With the scene set, one of the participants was given the interface that controlled the SAR system and allowed to familiarise themselves with it. During the familiarisation process the other participant was provided with a text description of all the assets that had been preloaded onto the system, allowing them to know what assets they could request without having to look at the other designer's interface. This was done to ensure the physical model of the watch/race car was used as much as possible as the 'shared design representation'. The session started once the participants felt confident in using the interface and finished whenever the participants felt they had developed as many concepts as they felt necessary, or when the one-hour mark was reached. The duration of each session was timed from start to whenever the participants claimed to have been satisfied with the number of concepts generated. The participants were randomly assigned to an experimental condition (conditions 1:1, 2:1, or 3:1 for the scale up experiments; conditions 1:10 or 1:20 for the scale down).

Once the session was complete the participants were asked to review the concepts they had saved, describing whether they still wished to take them to their "client" or not. This resulted in the "concepts taken forward" score. The method used was the novelty metric described in Dekoninck *et al.* (2018). Each participant was also asked to score the SAR platform using the Creativity Support Index (CSI) (Cherry and Latulipe, 2014) as recommended by O'Hare *et al.* (2018a) in order to establish the designers' scores for the usability of the SAR system itself. The CSI survey was developed to analyse ICT tools and is based on the well-established NASA Task Load Index (TLX) developed by Hart and Staveland (1988). Unlike the NASA TLX, the CSI survey better supports the evaluation of creative activities with a score between 0-100.

In addition to the data collected at the end of the session, a log of all the interactions with the SAR interface was captured enabling an analysis of the graphics used during the session. These assets can be tracked and catalogued to reveal which condition favoured which assets, as well as highlight which percentage of the available assets were used and placed onto the model. Lastly, the log file provides a chronological overview of all the additions and removals of assets that resulted in a concept.

6.3.3 Results

The analysis of scale and its effect on the concept generation behaviour of the participants was investigated from three perspectives:

- Concept Generation process: Total time taken to complete the exercise; Number of concepts generated
- Ease of Designing: CSI survey results
- Design Behaviour: Percentage of assets used; Individual asset usage across conditions; Asset use and Interaction

These provide a comprehensive understanding of the events in the design sessions, combining both qualitative and quantitative data to examine how the model's scale has influenced the design sessions. A total of twenty-eight sessions were completed. Six and four for the 1:20 and 1:10 scale conditions, respectively, in the scale down scenario and six sessions for each condition in the scale up scenario. This culminated in circa twenty session hours with video, audio, and log recording.

6.3.4 Concept Generation

Table 30, Table 31, and Table 32 show the time taken to complete each session in the scale up scenario. The data in the tables is further summarised as a boxplot in Figure 80 where it can be seen that the mean differed by more than 10 minutes and the distributions featured little overlap between the 2:1 condition and the other two conditions. Conversely, the 1:1 and the 3:1 conditions seemed highly comparable. This result suggests that an “optimum” scale may exist at which the time spent generating concepts is reduced. In addition, participants in conditions 2:1 and 3:1 seemed to be more consistent in the amount of time they required to complete the activity.

Table 30. Average time taken per concept and total session time for 1:1 scale

SESSION NR.	TOTAL TIME TAKEN (MM:SS)	NUMBER OF CONCEPTS GENERATED	AVG TIME PER CONCEPT (MM:SS)	CONCEPTS TAKEN FORWARD	PERCENTAGE TAKEN FORWARD
1	48:35	4	12:09	4	100
2	42:23	5	8:29	3	60
3	53:31	7	7:39	6	85.71
4	25:36	3	8:32	3	100
5	61:26	4	15:21	3	75
6	45:29	7	6:30	6	85.71
Average	**46:10**	**5**	**9:47**	**4.166667**	**84.40**

Table 31. Average time taken per concept and total session time for 2:1 scale

SESSION NR.	TOTAL TIME TAKEN (MM:SS)	NUMBER OF CONCEPTS GENERATED	AVG TIME PER CONCEPT (MM:SS)	CONCEPTS TAKEN FORWARD	PERCENTAGE TAKEN FORWARD
1	34:40	3	11:33	3	100
2	53:45	4	13:26	4	100
3	32:56	2	16:28	2	100
4	42:26	2	21:13	2	100
5	27:38	3	9:13	3	100
6	33:43	1	33:43	1	100
Average	**37:31**	**2.5**	**17:36**	**2.5**	**100**

Table 32. Average time taken per concept and total session time for 3:1 scale

SESSION NR.	TOTAL TIME TAKEN (MM:SS)	NUMBER OF CONCEPTS GENERATED	AVG TIME PER CONCEPT (MM:SS)	CONCEPTS TAKEN FORWARD	PERCENTAGE TAKEN FORWARD
1	31:36	12	2:38	7	58.33333
2	42:42	3	14:14	3	100
3	40:41	2	20:20	2	100
4	49:34	7	7:05	6	85.71429
5	47:09	6	7:51	5	83.33333
6	55:58	3	18:39	3	100
Average	**44:37**	**5.5**	**11:48**	**4.333333**	**87.89683**

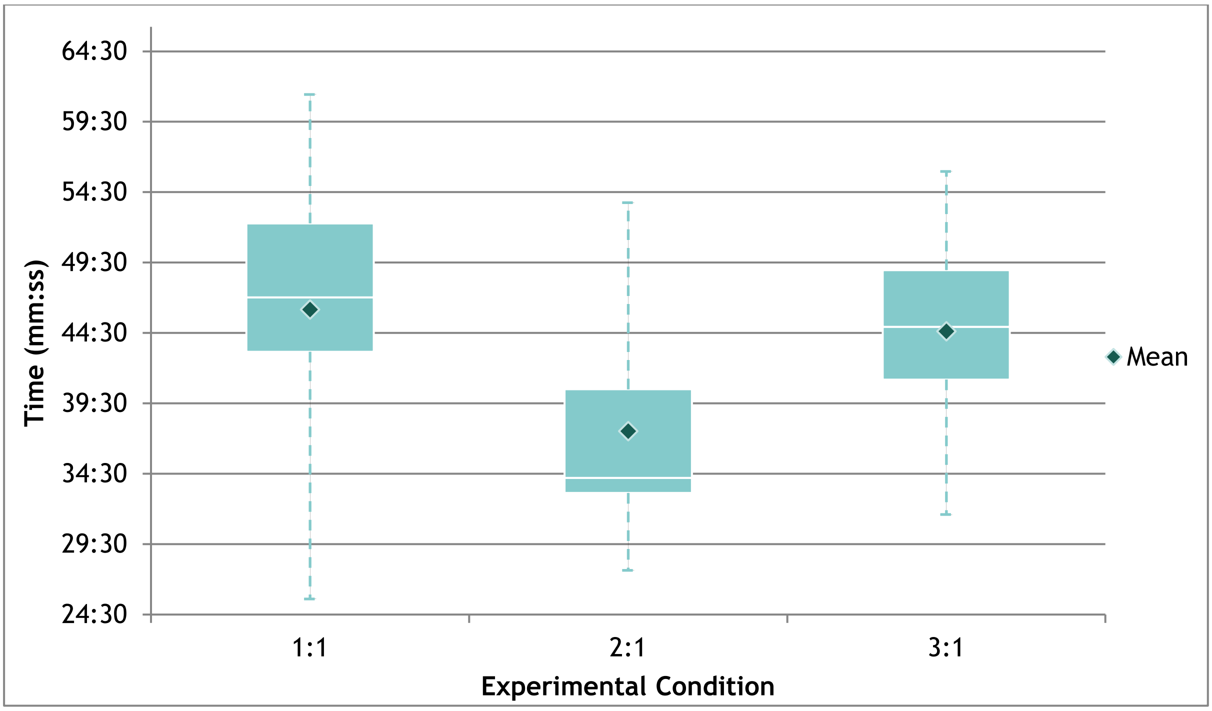


Figure 80. Comparison of time taken to complete the task for the scale up experiments

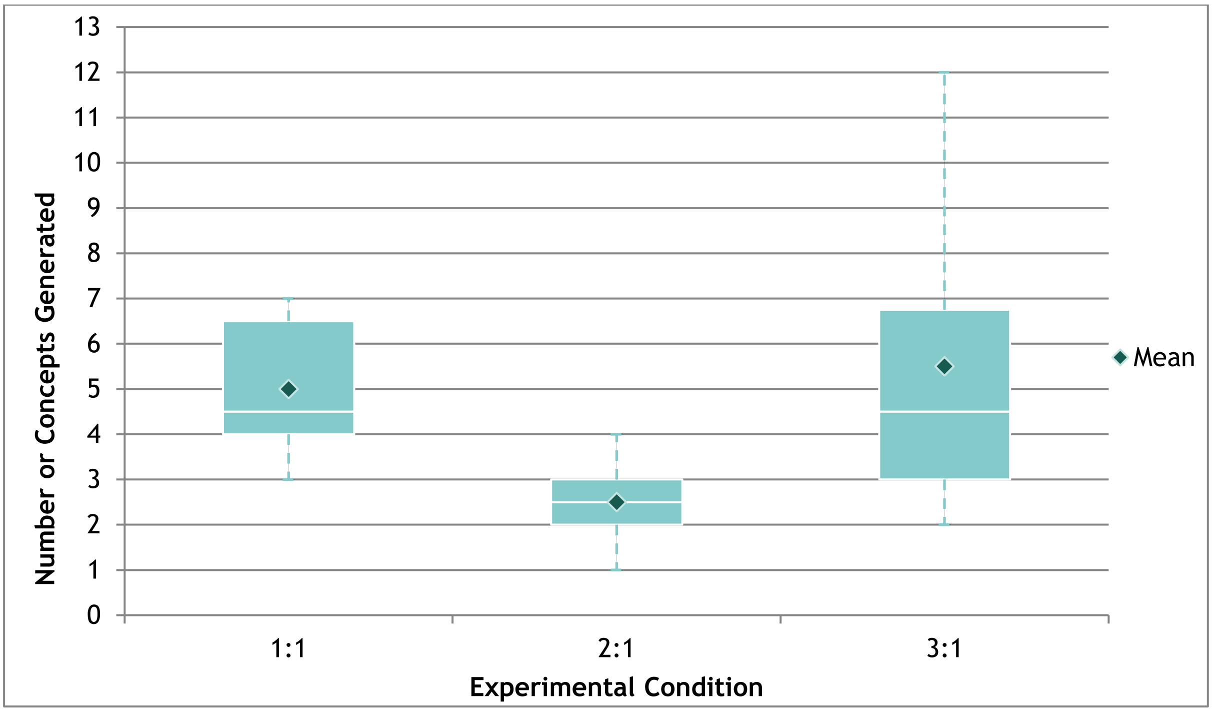


Figure 81. Number of ideas generated across experimental conditions for scale up experiments

It is also worth noting that, while the average time taken per session was higher for conditions 1:1 and 3:1, the time taken per concept was lower than in condition 2:1. In addition, it is important to note that, while both conditions 1:1 and 3:1 did generate more concepts (Figure 81) in more time (Figure 80), the participants were less satisfied with their concepts by the end of the session. Conditions 1:1 and 3:1 only took, respectively, 84.4% and 87.9% of concepts generated forward. This is in contrast to Condition 2:1 where, despite the lower number of concepts generated, 100% of concepts were taken forward.

Table 33 and Table 34 show the time taken to complete each session in the scale down scenario. The data in the tables is further summarised as a boxplot in Figure 82 where it can be seen that there is little variation between the mean time taken to complete the sessions in the 1:10 and the 1:20 conditions. However, the variance in the time taken appears somewhat lower in the 1:10 condition. This can be attributed to the smaller sample size used in the 1:10 condition. When compared to the results from the scale up scenario, it appears that the variance is much higher.

Table 33. Average time taken per concept and total session time for 1:20 scale

SESSION NR.	TOTAL TIME TAKEN (MM:SS)	NUMBER OF CONCEPTS GENERATED	AVG TIME PER CONCEPT (MM:SS)	CONCEPTS TAKEN FORWARD	PERCENTAGE TAKEN FORWARD
1	53:57	3	17:59	3	100
2	57:52	3	19:17	1	33.33333
3	41:29	3	13:50	3	100
4	23:42	1	23:42	1	100
5	20:35	3	6:52	2	66.66667
6	45:12	3	15:04	2	66.66667
Average	**40:28**	**2.666667**	**16:07**	**2**	**77.77778**

Table 34. Average time taken per concept and total session time for 1:10 scale

SESSION NR.	TOTAL TIME TAKEN (MM:SS)	NUMBER OF CONCEPTS GENERATED	AVG TIME PER CONCEPT (MM:SS)	CONCEPTS TAKEN FORWARD	PERCENTAGE TAKEN FORWARD
1	45:43	6	7:37	4	66.66667
2	33:10	3	11:03	2	66.66667
3	22:37	1	22:37	1	100
4	47:12	2	23:36	1	50
Average	**37:11**	**3**	**16:13**	**2**	**70.83333**

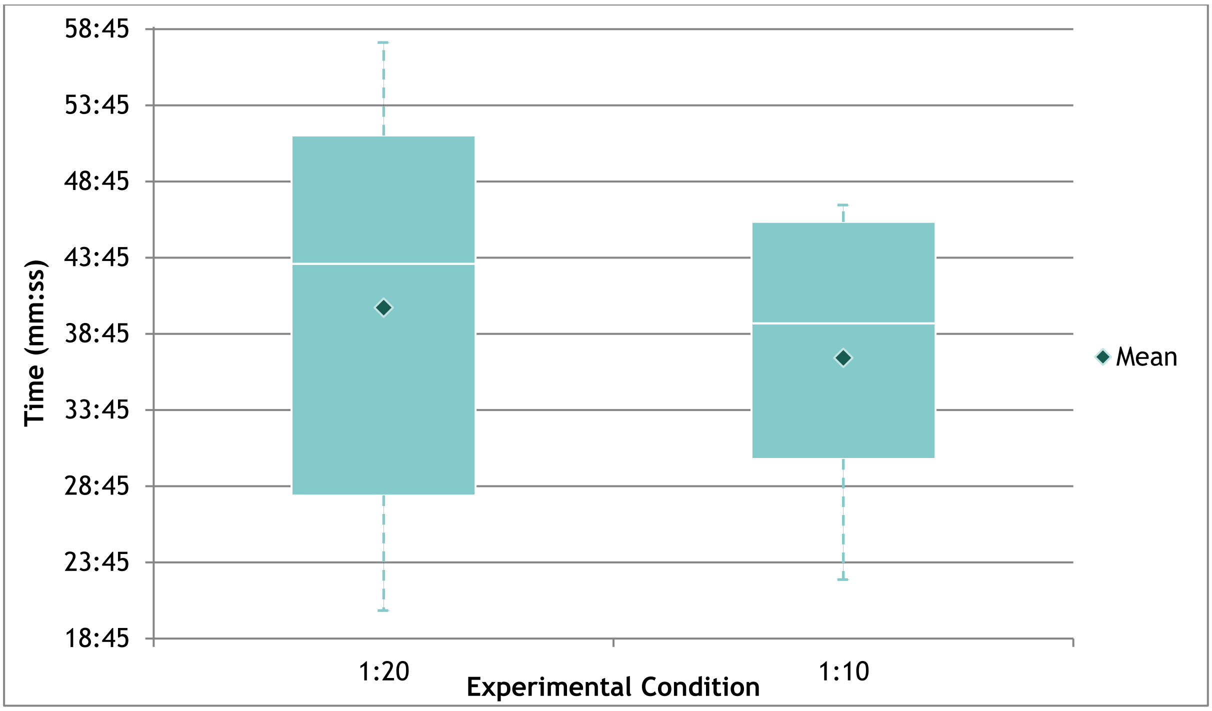


Figure 82. Comparison of time taken to complete the task for the scale down experiments

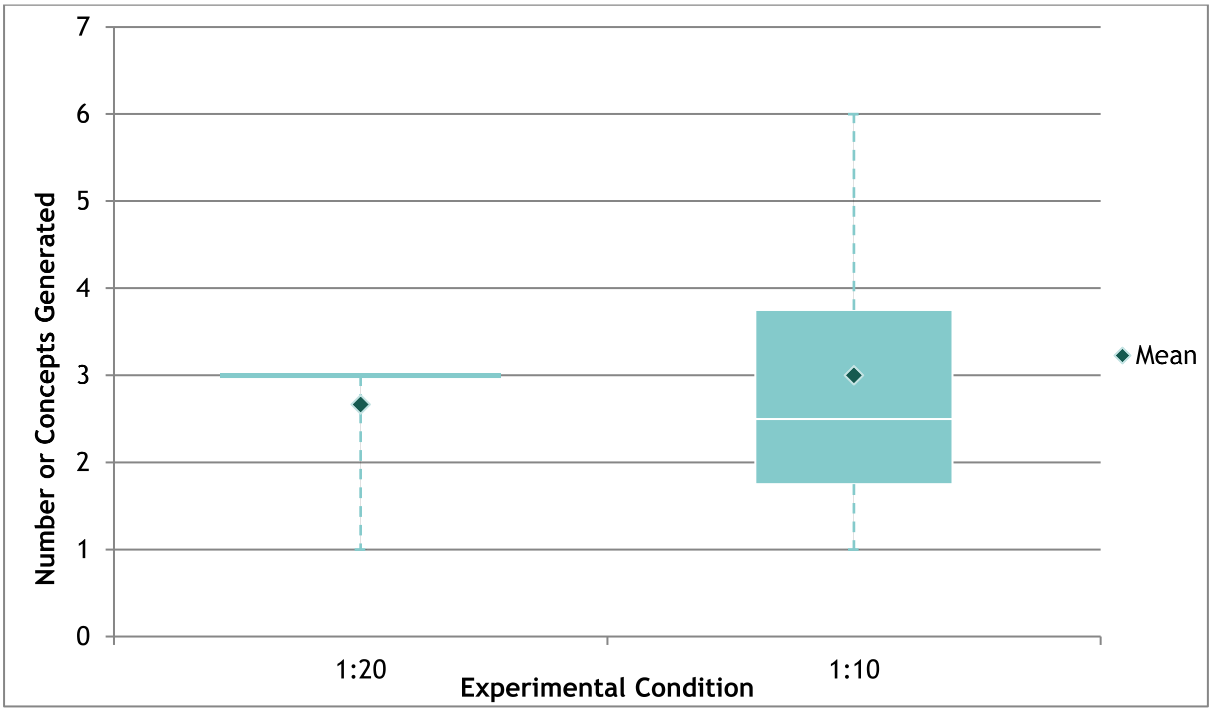


Figure 83. Number of ideas generated across experimental conditions for scale down experiments

The number of concepts generated in the 1:20 and 1:10 conditions highlight an unusual result. While the distribution of data for the 1:10 condition seems to be fully consistent with that from the scale up scenario, the results from the 1:20 condition have essentially no variance. All but one group of participants chose to generate three concepts. It is at present unclear what may have caused this behaviour, and why it was localised in the 1:20 condition. However, while the distribution may appear peculiar, the averages remain consistent with those seen in the scale up scenarios. It should also be noted that the average number of concepts take forward in the scale down scenario conditions is lower than that of the scale up conditions.

6.3.5 Ease of Designing

The results of the CSI survey are shown in Figure 84 (scale up scenario) and Figure 85 (scale down scenario) demonstrate that the participants did not feel any less comfortable using either model in either scenario. While this may initially seem somewhat counterintuitive, in particular for the scale up scenario where one would expect the larger scale model to be easier to view and use, the result may reflect that the 1:1 scale model is better at conveying the object due to its realistic proportions and thus renders the design process less abstract and mentally intensive. One other explanation is that the participants reflect more on the SAR system as a whole when reviewing their experience and thus, the user interface plays a larger role in their experience when compared to the models. Additionally, the increase in scale appears to reduce the variance across participants in the scale down scenario. The scale up scenario shows a decreased variance for the 2:1 condition, while the variance for the 1:1 and 3:1 conditions remains equal.

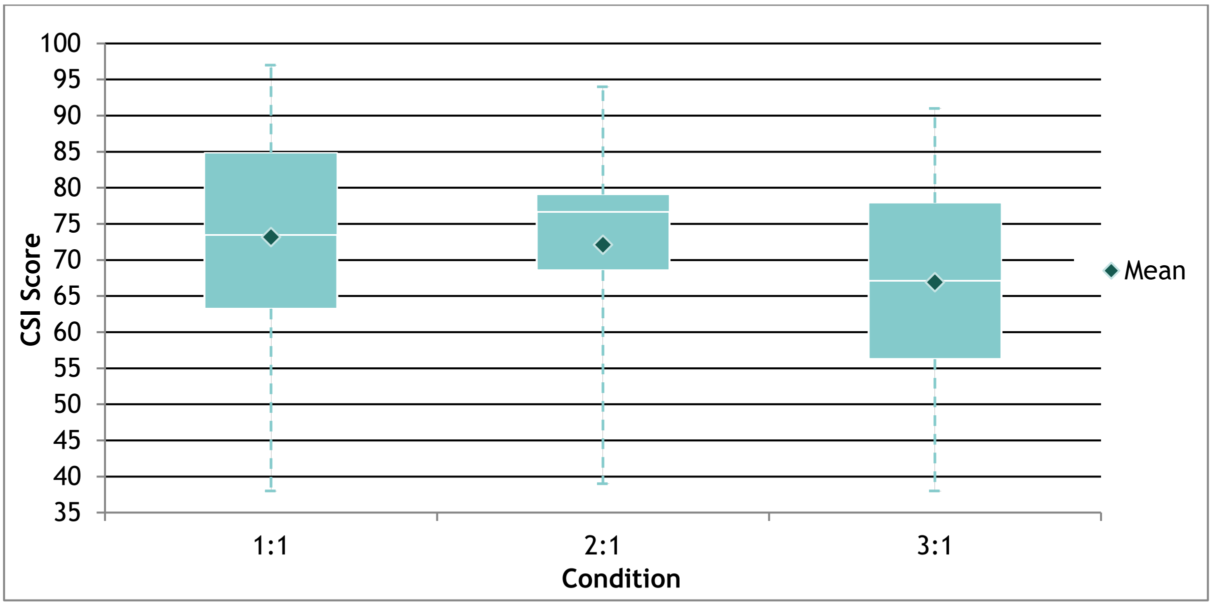


Figure 84. CSI results for scale up scenario

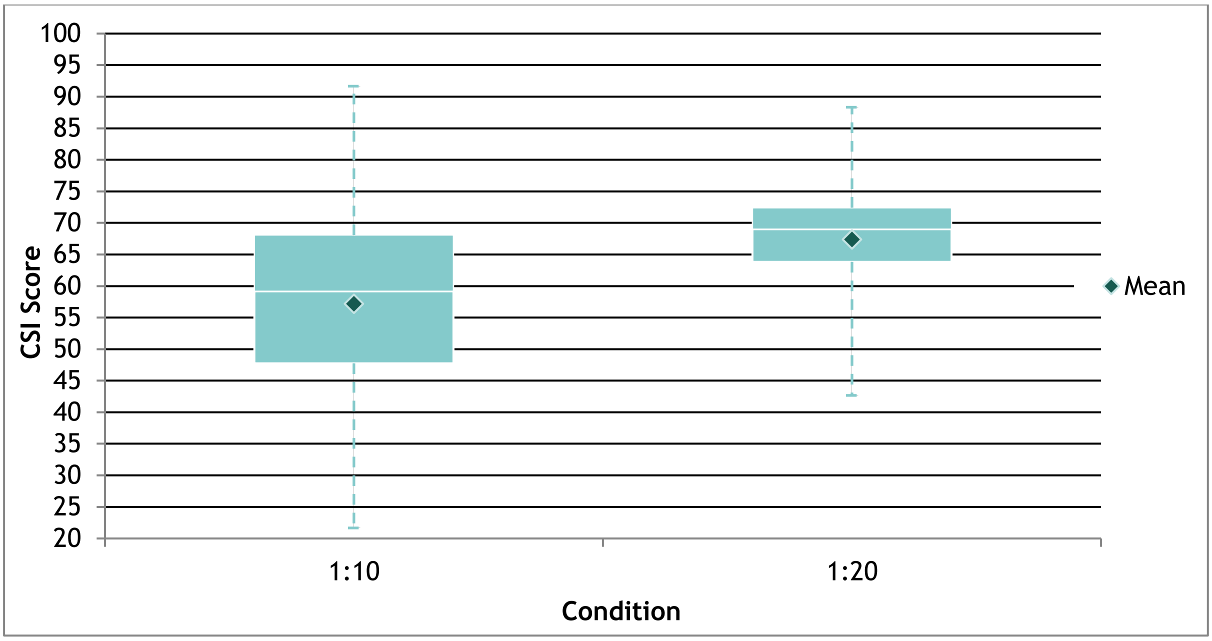
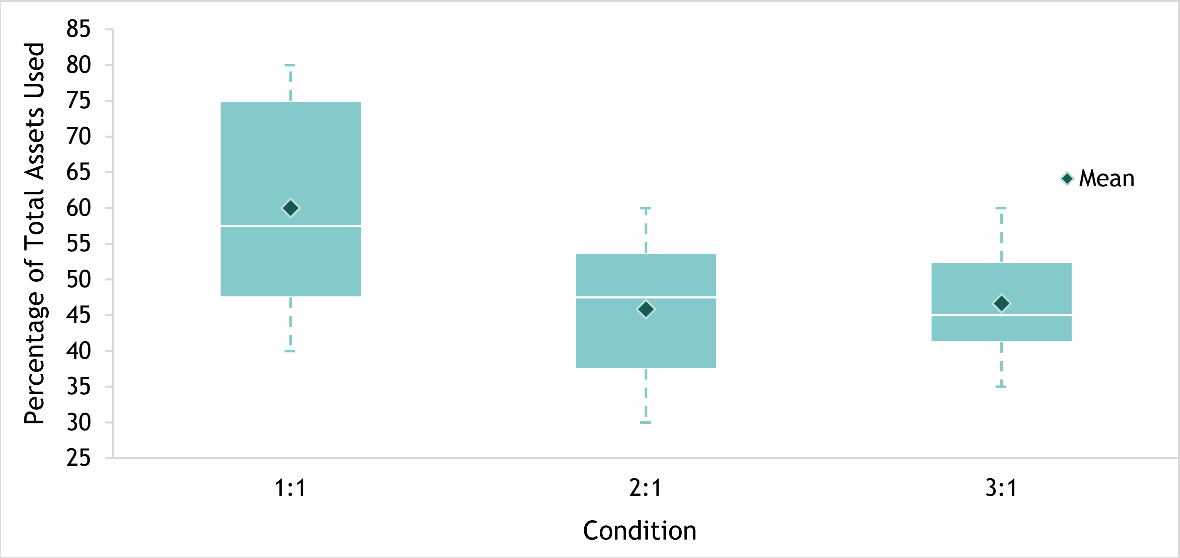


Figure 85. CSI results for scale down scenario

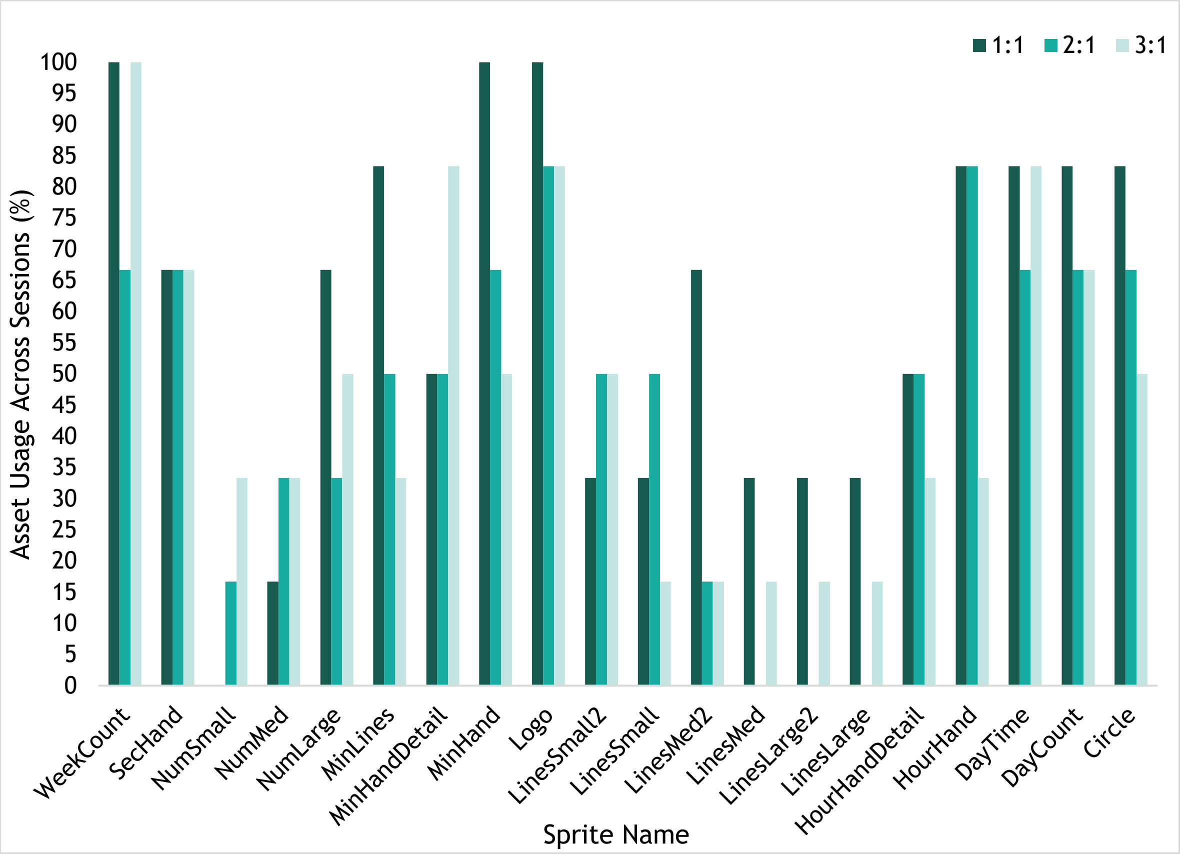
6.3.6 Design Behaviour

Figure 86 reveals the use of assets across all sessions for the scale up scenario. Figure 86a shows the usage of assets within each condition and reveals that condition 1:1 tended to cover more of the asset library than condition 2:1 although there is some variance and overlap present. This may imply that condition 1:1 featured more experimentation and iteration. This would help explain the higher concept generation and lower concepts taken forward observed in condition 1:1. The visual difficulty of using the 1:1 model may invite more experimentation and exploration but at the expense of more detailed design development and thus satisfaction with the finished results. Unlike with the results presented in Figure 80 and Figure 81, which presented the time taken and the number of concepts generated respectively, the results presented in Figure 86a show that the results from conditions 2:1 and 3:1 are comparable. This is in contrast to the aforementioned results from Figure 80 and Figure 81 where the results from conditions 1:1 and 3:1 had been comparable and condition 2:1 acted as the outlier. It may be that the small size of the physical model used for the condition 1:1 distorts the participants' ability to effectively view the assets, forcing them to explore them to explore the asset library better in order to find more suitable assets.

Figure 86b shows the assets evaluated by the designers in each condition; with the x-axis indicating the percentage of sessions that used that asset for a particular condition. It can be seen that almost all assets were used by at least some of the participants in condition 1:1. Conversely, there were some assets that were never trialled by any of the participants in condition 2:1, such as LinesMed, LinesLarge2, LinesLarge. All assets were used in at least one session in condition 3:1. However, while the use of the assets was greater, the percentage of sessions that made use of any one asset was generally lower for condition 3:1, in particular when compared to condition 1:1. This result, combined with that presented in Figure 86a, seem to suggest that no one session in the 3:1 condition explored more than ca. 47% of the assets at any one time. However, the assets chosen by the participants varied from session to session. This can be contrasted in particular to the 2:1 condition, which similarly explored ca. 46% of the asset library at any one time but appears to have been much more consistent in the type of assets omitted. This may indicate that the scale of the physical prototype does indeed play a role in influencing the designers' choice of assets and thus the design outcome.



(a) Use of assets available



(b) Asset use per condition

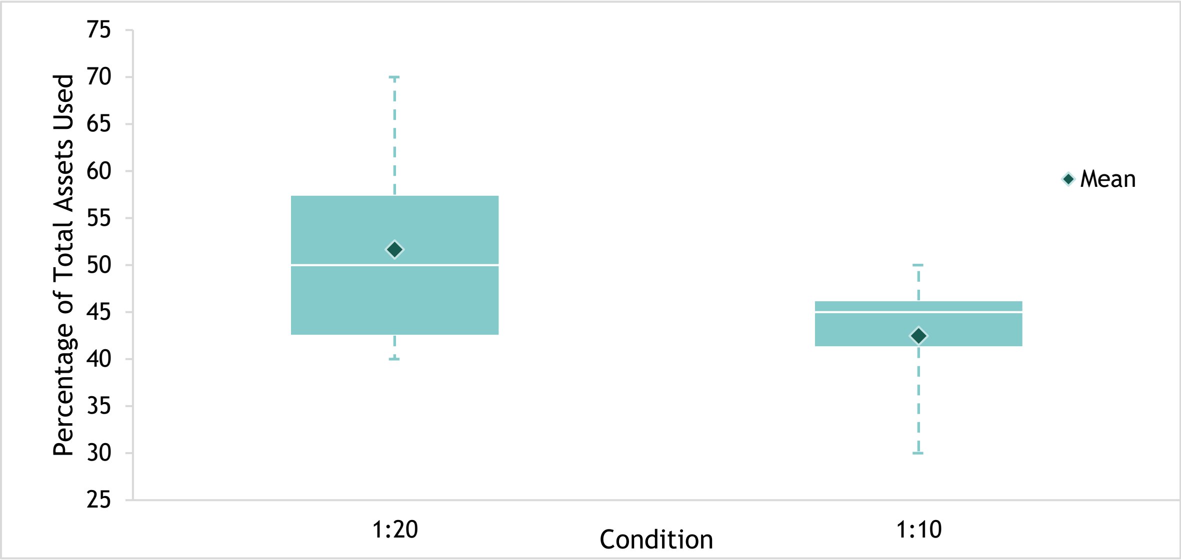
Figure 86. Asset use across the scale up (1:1, 2:1, 3:1) conditions

Figure 87 reveals the use of assets across all sessions for the scale down scenario. Figure 87a shows the usage of assets within each condition and reveals that there is some overlap between the percentage of assets used in both conditions 1:20 and 1:10. The one distinction that can be made is that the 1:10 condition has a smaller variation in the data, when compared to condition 1:20. Unlike with the results presented in Figure 86, where a clearer distinction was visible between two conditions, it is harder to make claims regarding the differences in the percentage of assets used in the 1:10 and 1:20 conditions. This appears to be consistent with the findings highlighted in Figure 82 and Figure 83, which showed the time taken and the number of ideas generated respectively, as neither of those figures highlighted any differences between the conditions 1:10 and 1:20.

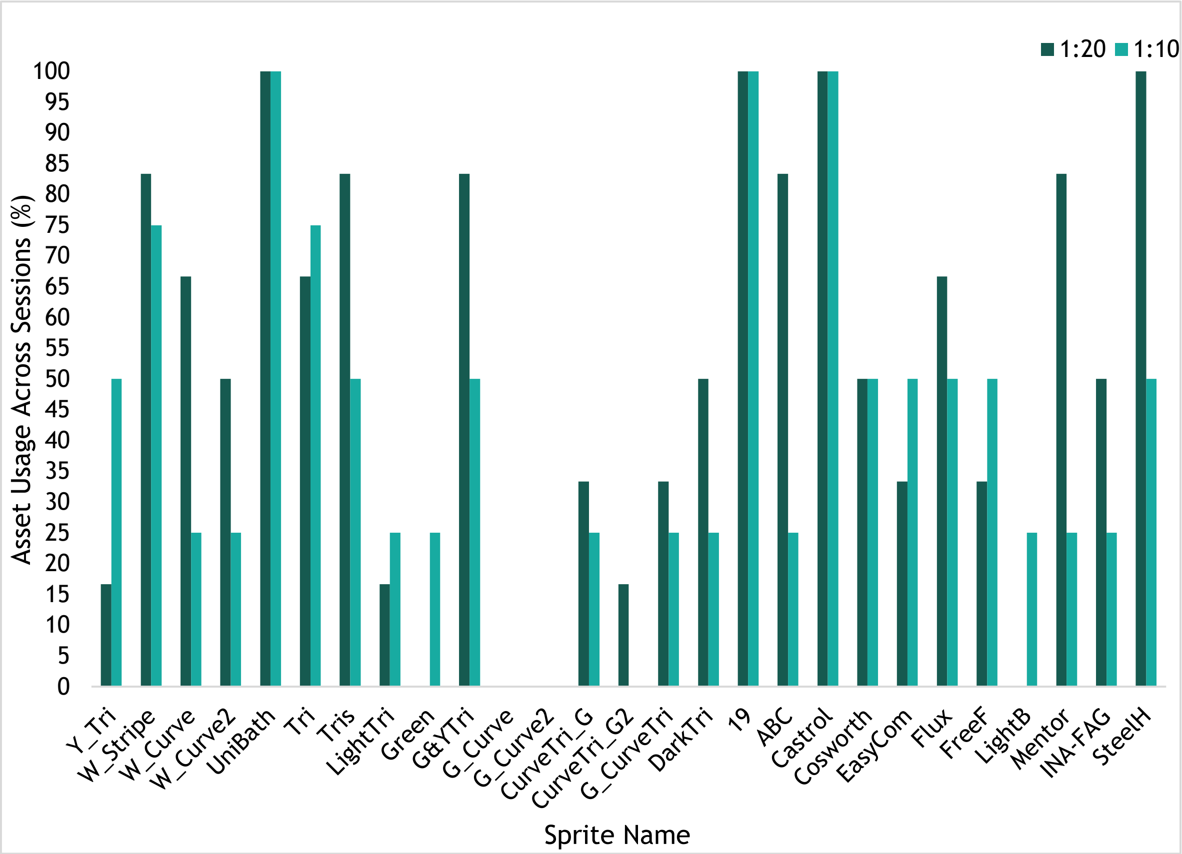
Figure 87b shows the assets evaluated by the designers in each condition; with the x-axis indicating the percentage of sessions that used that asset for a particular condition. Unlike with the assets used for the scale up scenario, and presented in Figure 86b, some of the assets used in the scale down scenario went unused in either condition (G_Curve, G_Curve2). This contrasts with the findings highlighted in Figure 86b, which show that any one asset was used in at least two of the three scale up scenario conditions. This may be the result of the larger asset pool available to the participants, a limitation of the technology that made the asset unappealing, or the asset being considered somehow inappropriate for the task.

The nature of the task also resulted in some assets being chosen in all conditions. Assets UniBath, 19, and Castrol were used in all sessions across all conditions. This is likely caused by the priming material as well as the preconceptions of the participants. They include the UniBath asset as the race car is for Team Bath Racing, and the 19 as race cars are expected to have a number. The consistent choice of the Castrol asset may be caused by the recognisability of the brand, the priming material, or the position of the asset in the asset library.

Figure 87b also highlights that sessions in the condition 1:20 used assets with a higher frequency compared to the sessions in condition 1:10. In addition, the difference in the percentage of the asset used across sessions is very noticeable for some assets. In particular, the 1:20 condition appeared to make more consistent use of assets across the sessions. This difference substantiates the findings presented in Figure 86, which had also highlighted that the size and scale of the physical prototype could play a role in the designers’ choice of assets and ultimately on the design session outcome.



(a) Use of assets available

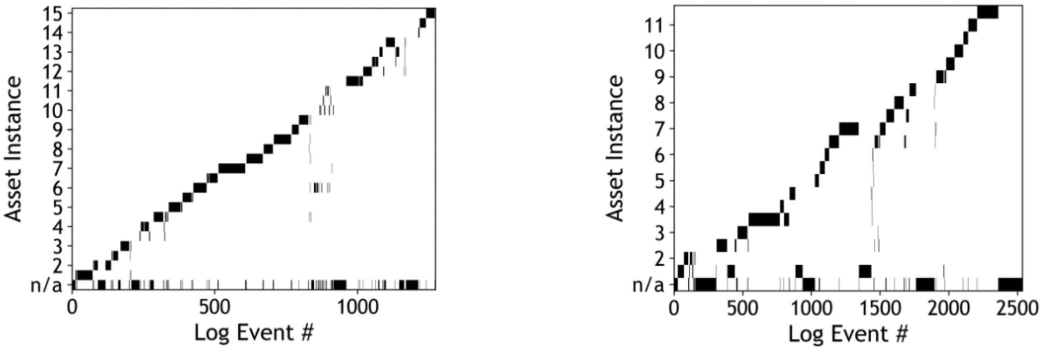


(b) Asset use per condition

Figure 87. Asset use across the scale down (1:20, 1:10) conditions

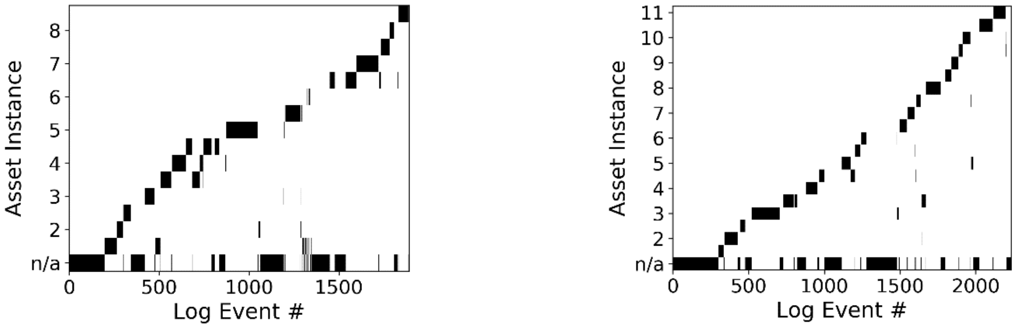
Figure 88 shows a sample of the interaction behaviour for each condition; the full set of graphs for each session is shown in the appendix in section D. The sequence features the events captured, mapped to the y-axis of the graphs and increasing by one with each event. These are plotted against the instances of assets from the asset library on the x-axis (N.B. designers could have multiple instances of the same asset within the scene. For example, two hour hands). Asset labelled “n/a” refer to events that do not map to a specific asset and are events that effect the entire scene (e.g., reset scene, UI button presses, zooming in/out).

The striking result from visualising the designers’ behaviour on the assets is the waterfall effect where designers will focus on a particular asset at a time but, once finished, they rarely revisit or iterate the position of the asset again. What one is seeing here is the creation of concepts on an asset-by-asset basis with the assets in the scene forming additional constraints for the next asset when it is placed. The designers will place a new asset and adjust until they are happy that it fits with the current set of assets or remove it from the scene. It is interesting to see that the designers rarely challenge the previous assets’ placement. This phenomenon is exhibited independent of the scale used by SAR. The variance between the axes of the different (random) samples does also show how much variance there was between teams in terms of the number of assets manipulated and for how long.



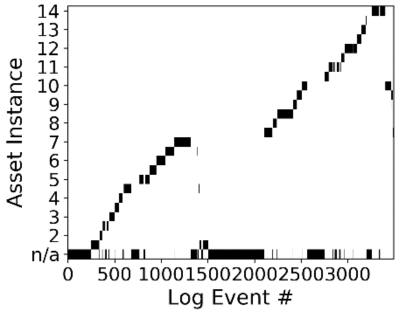
(a) 1:1 Condition

(b) 2:1 Condition



(c) 3:1 Condition

(d) 1:10 Condition



(e) 1:20 Condition

Figure 88. Asset use and interaction during sessions for a random sample of all conditions

6.3.7 Observations

The concept generation analysis (Figure 80 and Figure 81 for the scale up scenario, Figure 82 and Figure 83 for the scale down scenario) shows that the time taken to complete the task was not impacted in the scale down scenario. In the scale up scenario, the 2:1 condition showed reduction in the time needed to complete the session, with a mean time to session completion of 37 minutes and 31 seconds. However, the time required for both the 1:1 and 3:1 conditions was comparable, with mean session times of 46m:10s and 44m:37s respectively. This may be indicative of an “optimum” scale for the physical prototype at which design speed can be maximised.

What is noteworthy, however, is the number of concepts generated and the number of concepts taken forward as highlighted in Table 30, Table 31, and Table 32. Despite taking comparable amounts of time to complete the task, participants in condition 2:1 generated circa half as many concepts as participants in condition 1:1 and condition 3:1. Indeed, participants in the 2:1 condition averaged 2.5 concepts per session. This is in contrast to participants in the 1:1 and 3:1 conditions who, on average, generated 5 and 5.5 concepts per session respectively. Additionally, the condition 2:1 participants chose to take all their generated concepts forward. In contrast, the participants in condition 1:1 took only 84.40% of concepts forward and participants in condition 3:1 chose to take 87.9% of concepts forward. This, coupled with the time taken and number of concepts generated, implies that participants found the 2:1 scale model supported a more precise and deliberate approach, as they spent more time per concept.

It should be noted that this experiment focused specifically on the development of new concepts. The experimental setup specifically encouraged the development of novel concepts. In the design sessions explored in chapter 5, which were less controlled and more closely matched the reality of co-design sessions, filtering effectiveness was considered to be a positive parameter. In those sessions, a high level of filtering, leading to a larger number of discarded concepts, was considered a positive outcome to a design session as it was seen as indicative of a narrowing of the scope towards a final design.

However, as the participants in the experiment reported on in this section were specifically instructed to develop multiple concepts for the approval of a “customer”, the application of a filtering efficiency metric would prove misleading. Furthermore, the abstract nature of the task, when compared to the tasks undertaken by the professional designers in the experiments reported on in chapter 5, leads to the scope of the design session being much broader. It would stand to reason that the professional designers would have a considerably better-defined scope for the outcome of each design session, as well as the overall design process, than students given a one-off design challenge. As such, the high percentage of concepts being taken forward in these studies can be seen as positive, implying high levels of satisfaction on behalf of the participants with the ideas generated.

Further supporting this argument are the findings displayed in Figure 86b, which show that the participants in conditions 1:1 and 3:1 explored a wider variety of assets. The asset usage across sessions shows that, out of twenty available assets, all the assets were used at least once by any one group in the 3:1 condition. Similarly, participants in the 1:1 condition made use of almost all the available assets, with nineteen assets used at least once by any one group (the singular omitted asset was NumSmall). In contrast, participants in the 2:1 condition used only seventeen of the available assets. LinesMed, LinesLarge2, and LinesLarge were the assets omitted. This, combined with the preceding observations, may indicate that an "optimum" scale could exist for specific outcomes. Certain physical prototype scales may favour a more deliberate approach, while other scales may allow for more explorative design.

This is what may have caused participants in condition 2:1 to be slower in their concept generation; by conducting a more detailed design, shown by the lower use of assets as highlighted in Figure 86a and Figure 86b, they were subsequently more satisfied. This was highlighted by Table 31, which shows that participants in condition 2:1 took all designs generated forward across all sessions. Future analysis should thus attempt to capture more information into participant behaviour by means of either post-session questionnaires or more detailed analysis of the session video and audio recordings.

While the results show a link between model size and asset usage, increasing the number of sessions may be beneficial to highlight these differences further. It is also the case that an increase in the size of the digital asset library would diminish the designers' ability to explore it fully within the session time, therefore preventing it from influencing the emerging distribution. Accepting these limitations, condition 1:1 showed a higher frequency of asset usage across sessions. In the majority of cases, condition 1:1 participants used the same asset across more sessions when compared to condition 2:1.

Future work could build on this to investigate the types of assets present in the asset library and whether scale is more appropriate for certain types of design. It may be that modifying the experiment to have the assets exhibit greater variance, such as in resolution, contrast, line thickness, size, detail, etc. may aid in better understanding which types of assets are preferred at which scale and thus better gauge the influence of these characteristics on the participants. Additionally, this may yield a more detailed understanding of what drives the asset selection process and how the scale of the physical model may influence this.

The behaviour of the participants in the scale down scenario appears to differ from that of the participants in the scale up scenario, where asset usage is concerned. Figure 87a did not highlight a difference in the percentage of assets used between the 1:20 and the 1:10 conditions. However, the asset usage across sessions, shown in Figure 87b, differed from the asset usage in the scale up scenario and highlighted some interesting behaviour. Some of the assets provided for the participants in the scale down scenario, namely G_Curve and G_Curve2, went completely unused by participants in either condition; this is in contrast to the scale up scenario where all assets were used in at least two of the three conditions.

Furthermore, the three assets, namely UniBath, 19, and Castrol, were used in both conditions across all sessions. This is likely caused by the conditions of the study. As the participants were tasked with designing a race car for Team Bath Racing, it is unsurprising that the UniBath asset was considered in all sessions. The use of the 19 asset was likely also influenced by the inspiration board where a racing style number features prominently on nine of the eleven images. The widespread use of the Castrol asset is harder to explain but may have been caused by the greater recognisability of the brand, when compared to the other assets.

Finally, it is interesting to note that Figure 87b highlights the difference in the asset selection frequency between conditions 1:20 and 1:10. Participants in condition 1:20 were more consistent in their application of assets. Asset usage across sessions was higher in the 1:20 condition for fourteen assets. In contrast, in the 1:10 condition, the asset usage across sessions was only higher for seven assets. This, combined with the results highlighted in the scale up scenario, may indicate that there is indeed a link in how the physical prototype size and scale impact design outcomes.

The results of the CSI surveys provide some interesting insights into how the participants enjoyed and felt supported in their task by the SAR platform. The results of the surveys show no significant difference between how the participants experienced the SAR platform across all conditions and scenarios. This indicates that the 2:1 model performs to the same standard, from the users' perspective of the SAR system as a whole, as the 1:1 or 3:1 model and thus offers no detraction. Comparable results were found for the scale down scenario. One alternative theory is that while the users are influenced by the scale of the projection model, this is balanced out in some other way. For example, the 2:1 model may be easier to view but may be causing the participants to put more effort into bridging the difference between the scaled model and the real-life product they are creating. Future work should focus on the analysis of more scales and the use of a questionnaire which clearly separates feedback on the shared design representation (the model) and the user interface on the tablet.

Finally, looking into the analysis of the interaction behaviour revealed a "waterfall" pattern across all sessions in both the scale up and scale down scenarios. This analysis shows that the typical behaviour for the design team would be to place and iterate an asset's position, which they then determined to keep or reject. A concept is then built on an asset-by-asset basis where the design team is evaluating the application of a new asset relative to the current layout of assets on the model. Once an initial concept is created, very few changes to pre-existing assets are made to generate a new one. The addition of assets to a concept act as a method to constrain the design to a point where no new assets will aid in the design team's objective of making a valid concept and at this point, the team determines a concept has been reached. Being able to see when a new concept is generated in the log itself and mapping this within the graph would improve the understanding of how the shared design space is populated. This would provide a better understanding of the steps preceding the generation of a concept and thus perhaps illustrate which factors influence concept generation. It must, however, be noted that the results presented here are based on a study conducted using students as participants, rather than industry professionals, and that the tasks provided are fairly artificial. As such, it may be that these results are caused by the nature of the task or the participants' predisposition to the task, rather than revealing an inherent property of the design space. Additional studies, using professionals and in less constrained circumstances, would be required to investigate the validity of this hypothesis.

While very few changes to existing assets were observed in the logs, this does not necessarily mean that the design teams did not discuss or challenge the position of existing assets. Further processing of the audio transcripts could provide an insight into this, such as the apparent barriers that designers encountered when changing existing assets during a concept generation process.

6.3.8 Summary of Findings

The preceding section presented a study aimed at identifying the potential influence of scale on the design process when using an SAR system. Further work could focus on exploring how the participants feel about their concepts once these are returned to their proper scale. This could prove valuable as an additional metric to evaluate the impact of scale from the designer's own perspective, something these experiments did not consider in depth.

The key findings identified as part of this study were:

KF-23. The physical model's scale plays a role in how design session participants accomplish their task, influencing how they utilise the assets provided.
KF-24. The scale of the model does not play a major role in how the designers experience the SAR platform itself and their experience whilst designing.
KF-25. The scale of the model appears to have played a role in the designers' satisfaction with the concepts generated.
KF-26. There may be an "optimum" scale at which specific activities, such as exploration or concept generation, are more supported.
KF-27. Lastly, the study identified a consistent "waterfall" asset-by-asset layering design behaviour when generating concepts with SAR.

6.4 OBSERVATIONS ON THE SET OF STUDIES

The experiments detailed in this chapter were, as discussed in section 3.4.4, Table 12, aimed at addressing RO-4: "Analyse the impact of a sample of specific characteristics and features of the SAR platform on co-design sessions". The experiments achieved this objective through the analysis of the interface, the communication between participants, and the physical model scale. In focusing the experiments on these three items, it was possible to explore the impact of specific characteristics (model scale and interface) and features (communication).

Table 35 shows the research outcomes for each of the experiments mapped back to the research questions first described in section 3.4.2.

Table 35. Research Outcomes Compared to Research Questions

RESEARCH QUESTIONS	**RESEARCH ACTIVITIES**		
	6.1 Interface Comparison Study	6.2 Impact of SAR on Communication between Design Session Participants	6.3 Impact of Scale in Design Sessions Supported by an SAR Platform
RQ-1: "How can co-design sessions' efficacy be measured?"	• The repetitiveness of the task can cause participants to feel uninvested and uninterested. • The experiments showed that providing a training scenario to the participants to familiarise them with the interface and task did not create a learning effect. • The results from both the CSI and SUS surveys were closely correlated.	• The study shows that the methodology applied is valid for analysing communication.	
RQ-2: "How does an SAR system affect co-design sessions' efficacy?"	• Participant analysis was inconclusive, but time taken was overall shorter at UBATH; potentially indicating a link between SAR platform type and time taken.	• Time taken to complete the session was shorter in SAR supported sessions over PC screen sessions.	• A "waterfall", asset-by-asset, layering was observed when participants used SAR to design.
RQ-3: "How do specific SAR characteristics and features affect co-design sessions' efficacy?"	• Accuracy of position, rotation, and scale was comparable between conditions relying on the use of a touch interface and a mouse interface. • The touch interface conditions showed comparable levels of accuracy. • Participants rated the usability of a PC/Mouse interface highest of all. • Of the touch interfaces, the UV Map was rated highest.	• Interaction was higher between participants in the SAR scenario. • Participants in the PC sessions relied more on verbal cues; this may indicate that the information throughput for SAR supported design sessions is higher.	• There appears to be a link between physical model scale and how the participants use the assets provided. • The scale of the model does not play a major role in how the participants experience the design activity or the platform. •The scale of the model seems to influence the designers' level of satisfaction with their output. • There may be an "optimum" scale at which specific effects are more pronounced.

7 STUDIES WITH INDUSTRY

As part of the SPARK project's development of an SAR platform, industry feedback was sought to shed additional insight into SAR's potential role as a tool to support design and co-design. This feedback was used to guide the SPARK platform's development to better meet the needs and wishes of potential investors and users. Furthermore, the feedback gathered enabled a better understanding of how end-users desired to implement an SAR platform into their existing design process.

Chapter 7 presents a secondary analysis of the data gathered as part of the SPARK project's collection of industry feedback. This data was gathered as a collective effort by the SPARK project partners. The secondary analysis focuses on using the data to answer RQ-4: "What are the industry requirements for an SAR system to support co-design?". This is achieved, in part, through the fulfilment of RO-5: "Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform". While the setup and running of the studies was a collaborative SPARK effort, the secondary analysis was performed solely by the author.

Chapter 7 begins by describing general feedback obtained from potential interested parties who agreed to provide feedback on the SPARK platform at trade fairs. These parties were encountered at trade fairs attended by members of the SPARK Consortium and chose to provide feedback and information regarding their potential interest while viewing the SPARK project stand. This general feedback is further expanded upon in section 7.2, where members of industry who had a chance to design using the SAR platform provide more detailed feedback on their expectations for an SAR platform, how they would wish to integrate SAR within their existing design process, and their overall impressions of the SPARK platform. Finally, this chapter concludes by providing a more detailed analysis of how SAR influences industrial design processes and the value-add that SAR could generate for industry by analysing the longitudinal impact of SAR on a design project.

7.1 INDUSTRY FEEDBACK FROM TRADE FAIRS

The data presented in this section was originally published as part of the report "D5.4 Show-cases for Increasing the Awareness of SPARK Platform" by Bellucci *et al.* (2018b). This document reports on the planning, preparation, and execution of events to showcase the SPARK platform.

To engage industry, ten trade fairs were attended. The intention was to showcase the technology and network with interested parties. In networking with these interested parties, it was hoped to gain a general understanding of the expectations potential end-users had for novel SAR and AR technologies. Furthermore, the feedback was expected to provide greater insight into how an SAR platform would be implemented into the design process as well as what the overall industry expectations were. Table 36 shows the events, the target audience, and the approximate attendance of the ten trade shows attended the by SPARK consortium. Communication with attendees was done primarily in English but, where feasible, additional information and explanations were provided in the attendees preferred language.

Table 36. Trade Fairs Attended by SPARK Consortium Members (Bellucci et al., 2018b)

TRADE FAIR NAME	TARGET GROUP	APPROXIMATE PARTICIPANT NUMBERS
MILANO DESIGN WEEK	Designers, all	4270
BARCELONA DESIGN WEEK	Designers, all	1800
DEVELOP 3D LIVE, WARWICK	Designers, Design tech. managers	1800
EU DIGITAL ASSEMBLY, SOFIA	EU H2020 community	130
DESIGN COMPUTING COGNITION	AI experts, researchers in design field	130
SUPERNOVA, ANTWERP	General public	5000
EMPACK, BRUSSELS	Packaging designers	100
PROTOTYPING '18, KORTRIJK	Designers	3500
ICT, VIENNA	Digital EU	37
ARGONAUTS VISIT AT POLIMI (MI)	Professionals	15

Based on the attendance numbers presented in Table 36 the estimated reach is of circa 16800 individuals. As a result of attending the fairs, over 700 individuals were able to interact with the SAR platform(s), with approximately 210 and 82 participants interacting with the SAR platform at the Milano Design Week and Barcelona Design Week respectively.

7.1.1 Methodology

When visitors approached the SPARK exhibit, they were given background information on the SPARK project, a tour of the SPARK platform, and provided with an opportunity to interact with the platform. Any questions asked by the visitors were answered during their stay at the stand. Finally, visitors were asked to complete the survey. This survey was provided in English. Visitors were informed how the data from the survey would be used and informed that they did not have to complete the survey if they did not wish to do so.

Figure 89 illustrates the stands present at some of the fairs attended as well as showing interaction with the participants. Not all those who visited the SPARK exhibits completed the survey. Completion of the survey was not mandatory nor was it a condition to being able to interact with the SAR platform.

Figure 89. Photos taken at some of the trade fairs attended. Clockwise from the top left: EMPACK, Prototyping ’18, Barcelona Design Week, Milano Design Week (Bellucci et al., 2018b)

The survey consisted of six questions. The decision was made to keep the number of questions short in order to ensure completion by participants and maximise their engagement. Quantitative feedback was collected by means of four questions, listed in appendix H. Three of these questions collected information concerning the participant’s company and the frequency with which they engaged in collaborative design sessions. The final question focused on recording participants’ opinions regarding the challenges faced by their organization on a five-point Likert scale. Qualitative feedback was collected by means of two open-ended text boxes on the questionnaire but also through simple dialogue with attendees who used the SPARK platform.

The first quantitative feedback question was “How often does your company currently hold co-creative design sessions (or product development review meetings) with internal stakeholders, customers, or end-users?” Respondents could choose one of five options, from “never” to “one or more times per week”.

In addition to asking about the frequency of co-creative design sessions, the survey asked respondents to identify the importance of specific challenges their organization might face. These challenges were listed as: “Obtaining actionable feedback from stakeholders”, “Generating novel ideas”, “Reducing the cost of creating prototypes”, “Reducing unnecessary iteration in the design process”, “Reducing time to market”, and “Overcoming barriers to communication with stakeholders”. For each one of these six challenges, respondents were asked to score the level of importance on a scale of one to five. The challenges were derived based on the SAR platform’s intended goals for supporting design sessions. The SPARK platform’s design was specifically aimed at supporting the six challenges listed and, based on the responses of the participants, it was possible to assess whether these challenges were in line with the expectations of the members of industry.

Table 37 maps the challenges to the intended goals and objectives of the SPARK platform. If the feedback obtained from the members of industry had indicated that specific challenges were of little or no concern to them it would have become necessary to address the goals of the SPARK platform or of any future platforms that was developed to support co-design through SAR. The challenges and the goals and objectives that these were aimed at addressing were derived from the original guiding principles that led to the founding of the SPARK Consortium (SPARK Consortium, 2015).

Table 37. Mapping of challenges to the SPARK platform's intended goals and objectives

CHALLENGE	GOALS AND OBJECTIVES
OBTAINING ACTIONABLE FEEDBACK FROM STAKEHOLDERS	By increasing the ease of communication through the reduction of barriers and the promotion of intermediary objects, the SPARK platform aims to make it easier for the participants in design sessions to raise concerns, offer constructive criticism, and provide encouragement.
GENERATING NOVEL IDEAS	By allowing faster and more immediate feedback from design session participants, it is possible to better explore the design space potentially enabling more creative design solutions to be developed.
REDUCING THE COST OF CREATING PROTOTYPES	The use of SAR is expected to aid the prototyping process, as a simple blank physical prototype needs to be created. Iterations and variants that only need to change the layout, colour, or finish of the product do not need to be built, resulting in lower overall costs for prototyping.
REDUCING UNNECESSARY ITERATION IN THE DESIGN PROCESS	As SPARK is intended for live co-design sessions, it is expected that the immediacy of the feedback received can help reduce the likelihood of pursuing avenues that do not mesh with the clients' vision for the product. In doing so the design process can be kept more focused, reducing unnecessary iterations caused by chasing tangents.
REDUCING TIME TO MARKET	The reduction in unnecessary iteration, the provision of more direct and actionable feedback, and the ability to overcome barriers to communication should allow the SPARK to reduce the time to market.
OVERCOMING BARRIERS TO COMMUNICATION WITH STAKEHOLDERS	By providing participants to co-design sessions with a shared design representation that can be edited in real time, the expectation is that SPARK will allow participants to more effectively express their desires and opinions. By acting as an intermediary object, the SPARK platform is expected to allow participants, who may not have the same cultural, technical, or social background to explain themselves more effectively.

Subsequently, participants were asked to identify the type of organization they worked for. Four pre-selected options were provided as well as an "other" option, allowing respondents to provide their own descriptor. Lastly, the participants were asked to identify how many people were employed by their organization. Four options were provided, ranging from “less than 10 people” to “more than 250 people”.

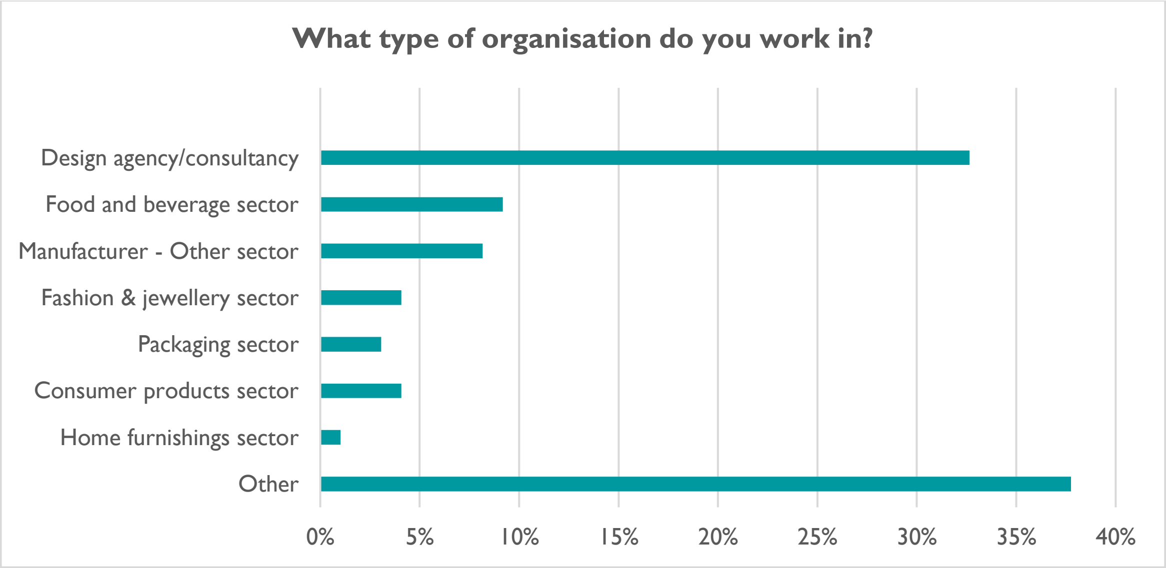
The qualitative feedback questions asked were: “In your own words, what are the things that you like most about this new technology?” and “In your own words, what are the things that you would most like to improve in this new technology?”. Respondents were provided with a text box where they could provide any and all feedback they wished.

As previously mentioned the data presented in this section represents a refactoring of the data originally presented in “D5.4 Show-cases for Increasing the Awareness of SPARK Platform” by Bellucci *et al.* (2018b). The secondary analysis of the quantitative data collected across the trade fairs was aggregated and subsequently tabulated. Based on the responses it was possible to gather data on the types of organizations that may have interest in an SAR platform, the frequency with which they engaged in co-design sessions, and the size of the organizations. This demographic information aids in creating a profile for a prospective end-user for an SAR platform. The information relating to the challenges faced enables a better understanding of how an SAR platform should be tailored to suit the needs of industry. This relates directly to RO-5: “Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform”.

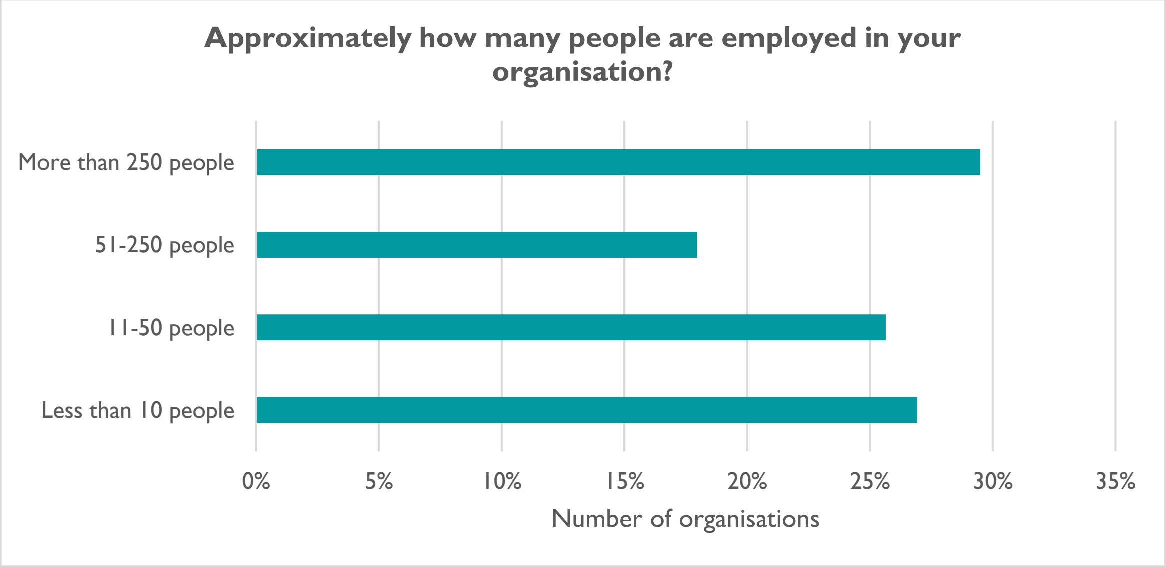
7.1.2 Survey Feedback

Figure 90 aids in breaking down the respondents’ demographics. A total of 108 individuals completed the survey. Figure 90(a) shows the types of industry survey respondents associated with. A large proportion of respondents did not associate with any of the listed categories, choosing to mark “other” as their response. The largest defined category respondents associated with was design agencies and consultancies. Members of other types of industries trailed in all other categories, with the next largest number of respondents coming from the food and beverage sector and the manufacturing sector.

Figure 90(b) shows that a fairly even distribution in the size of the respondent’s organizations. The largest number of respondents came from organizations that employed more than 250 employees, which is the maximum threshold set by the European Union for small to medium enterprises (European Commission, 2020). However, the rest of respondents came from organizations that would qualify as small to medium enterprises.



(a)



(b)

Figure 90. Respondent demographics: Figure 90(a) shows the industries respondents were most associated with. Figure 90(b) shows the size of the respondents' organizations (Bellucci et al., 2018b)

Figure 91 shows the frequency with which the respondents participate in co-creative design sessions. The graph shows that circa 65% respondents participate in a co-creative design session at least once per month. Of these respondents, almost 40% engaged in co-creative design sessions one or more times per week.

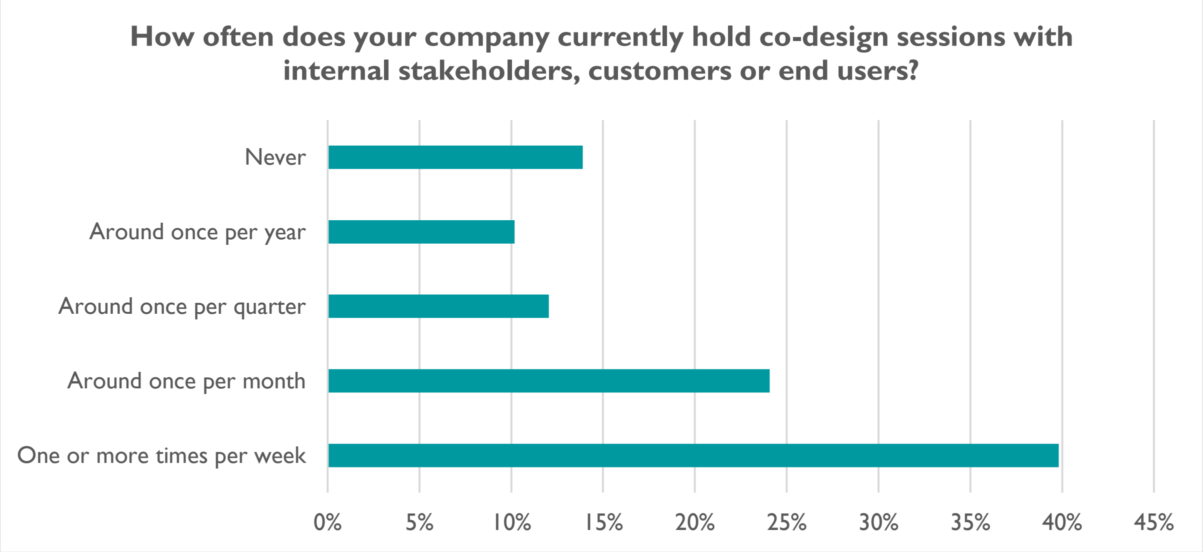


Figure 91. Frequency of respondent’s participation in co-creative design sessions by percentage (Bellucci et al., 2018b)

Figure 92 lists the results for question two of the survey. The results show at least half of the respondents found all challenges important or very important. Respondents were particularly concerned with reducing the time to market and being able to communicate effectively with stakeholders, wanting to overcome communication barriers a well as obtaining actionable feedback. It is interesting to note that respondents were least concerned about reducing prototyping costs, yet most concerned about reducing time to market.

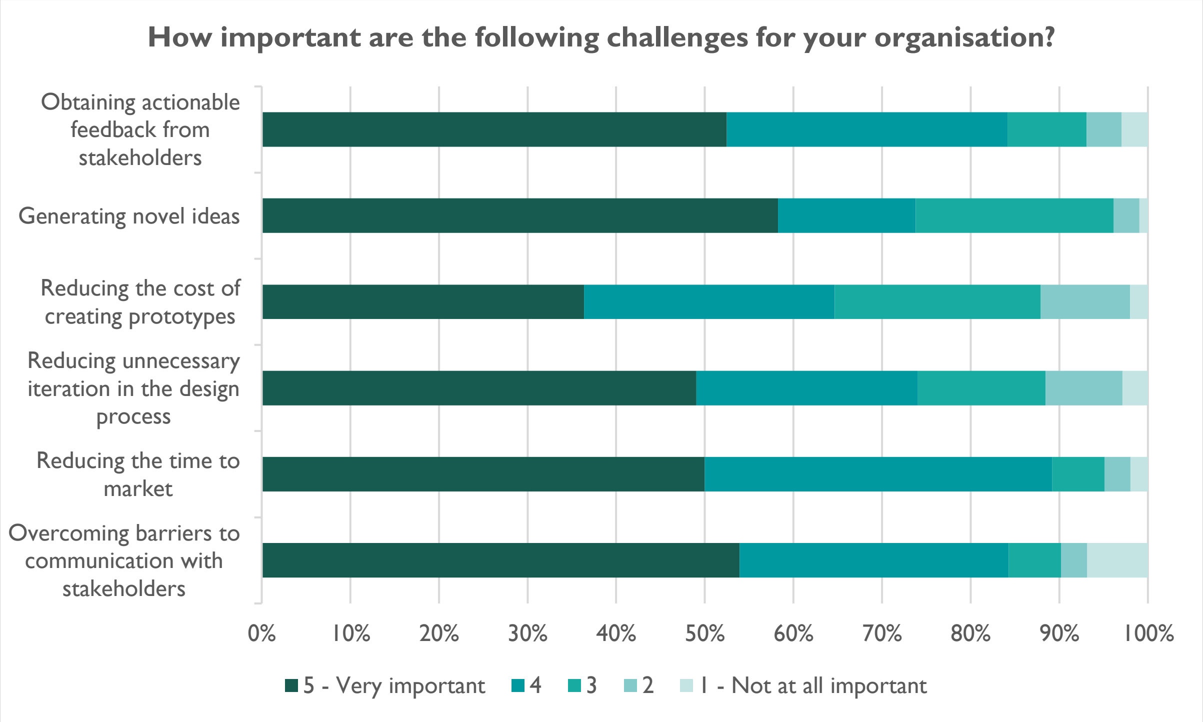


Figure 92. Respondent’s perception of the importance of specific challenges to their organization (Bellucci et al., 2018b)

7.1.3 Summary of Findings

The surveys showed that the respondents came from a wide variety of backgrounds, with the single largest category being design agencies or consultancies, as shown in Figure 90. The size of these organizations is distributed evenly amongst the categories the respondents were asked to select, but the majority fall within the definition of small to medium enterprises.

The trade fairs and events attended allowed for a considerable amount of data collection in the form of informal feedback as well as formal responses to a survey. The data collected in the survey in particular shows that members of industry may be interested in an SAR platform aimed at supporting their design process. Should an SAR platform be developed to address the challenges highlighted by Figure 92, it is possible that members of industry would be eager to adopt it. Indeed, as was shown by the responses collected in Figure 92, a majority of respondents rated each of the six challenges as important or very important to their organization.

The data collected showed that the respondents frequently engaged in co-design sessions with stakeholders or customers, as highlighted by Figure 91. As mentioned, the majority of respondents stated they valued or highly valued all the challenges they were asked to score. These challenges were based on the target functionalities an SAR platform could best target and improve as originally described in Table 37.

In summary, the evidence gathered shows that there is an interest from those in industry, as evidenced by Figure 91, who do engage in regular co-design sessions. Having highlighted that there are organizations that regularly engage in co-design sessions, and could be potentially interested in tools to support these sessions, it was possible to assess the types of challenges these companies faced. Figure 92 shows the responses of the participants assessing the challenges identified as potentially solvable through the use of the SPARK platform as highlighted in Table 37. An SAR platform tailored to tackle these needs would then be likely to gain traction as the majority of participants identified all the challenges as important or very important to their organization.

Thus, the key findings were:

KF-28. A majority of respondents rated each of the six challenges as important or very important to their organization

KF-29. Respondents frequently engaged in co-design sessions with stakeholders or customers

7.2 SAR PLATFORM PILOTS WITH INDUSTRY MEMBERS

This section further addresses RO-5: “Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform”. By building on the feedback obtained in section 7.1, this section attempts to provide additional insight to answer RQ-4: “What are the industry requirements for an SAR system to support co-design?”. The more targeted and in-depth industry feedback collected here is aimed at providing additional insight into the features and characteristics of SAR that members of industry consider crucial to facilitate and promote adoption of SAR technology as a tool to support co-design.

The feedback gathered from trade fairs was used to better target the needs of potential users in industry. On the basis of the feedback received the SPARK platform was improved to better address these needs. However, the feedback gathered from the trade fairs, in particular any informal or qualitative feedback, was fairly generic. The data collection from the trade fairs was intended to broadly capture and understand the general opinion of prospective future users of the platform, and could not, in large part due to the informal and brief nature of the interaction, capture feedback on more specific features and characteristics of the SAR platform. The information reported on in this section was originally published as part of the report "D5.3 Demonstration with Other Creative Industries And With Customers" (Majoral *et al.*, 2018).

To address this, a number of companies and organizations who had shown interest in the development of the SPARK platform, and who had already indicated they would be interested in acquiring the system for their own uses, were invited to participate in a number of pilot design sessions.

The design sessions allowed these companies to familiarise themselves with an SAR platform, provide feedback, and discuss how they would wish to implement such a platform into their workflow. The following sections discuss the companies who participated in these pilot design sessions, the work performed during the sessions, and the feedback received. It should be noted that these pilot design sessions were conducted only partially in English to better accommodate the needs of the participants. Sessions were led in Italian, Catalan/Spanish, and Dutch for the Artefice, Stimulo, and AMS sessions respectively.

7.2.1 Pilot Design Sessions

Table 38 shows the list of companies that participated in the pilot design sessions. The types of sessions the companies participated in have been categorised into either product design or packaging design sessions. In the former type of design session, the distinction between the two is that product design is focused more on the utility, function, and experience of the product. In contrast, packaging design focuses mostly on the message conveyed by the styling, the information present, and the general layout of the packaging. At the companies' request some of the companies or products listed in Table 38 have been anonymised.

Members from the SPARK Consortium were present at each of the design sessions and aided in the operation of the SPARK platform as well as mediated the discussion. These design sessions took place at three different locations: Stimulo, Artefice and Antwerp Management School (AMS).

Table 38. List of Participating Companies, Type of Session they Engaged in, Product, Objectives Set for the Session, and the Role of the Representatives Attending the Session (Majoral et al., 2018)

COMPANY	SESSION TYPE	PRODUCT(S)	SESSION OBJECTIVE(S)	COMPANY REPRESENTATIVES	LOCATION
COLRUYT GROUP	Packaging	• Chocolate Pudding • Chocolate Bar • Spirit Bottle	Evaluate new graphical concepts for the packaging of the product	• Product Manager • Retail Designer • Graphic Designers (x2) • Product Designer	AMS
FOOD INC.	Packaging	• New Product Packaging	Present rework done after a previous presentation	• CEO • Commercial Director • Marketing Consultant • Brand Manager	Artefice
ZOBELE	Packaging	• Fragrance	Check the feasibility of using SAR for showcasing products and variants at trade shows	• Design Manager • Product Designer	Stimulo
WAVECONTROL	Product	• ONIRIS Real-Time Handheld Analyzer Device	Perform user tests of the interaction aspects of the large touchscreen used in the product	• Director • Chief Engineer • Creative Director • Industrial Designer	Stimulo
SAMSONITE NV	Product	• Cosmolite Luggage Suitcase • Neopulse Luggage Suitcase	Check the utility of SAR as a tool for use during design review sessions	• Design Director (Europe) • Design Manager	AMS

7.2.1.1 Companies and Session objectives

The following section provides some additional information on the five companies that participated in the pilot design sessions. In addition to a brief background to aid in contextualizing the companies' nature, this section provides a summary of the objectives for each session. It should be noted that some companies worked on multiple products in their design session. It should also be noted that, as the session outputs were considered to be confidential, no images of these can be shared here.

Colruyt Group is a Belgian retail corporation with a yearly revenue of ca. 9.5 billon euro. The group deals predominately in supermarket retail but has interests in clothing and toy shops as well. Colryut Group is also one of the franchisees licenced to run SPAR brand supermarkets in Belgium. The group has interests globally but is particularly present in the Benelux area, as well as France (Colruyt Group, 2020).

Five participants from Colruyt Group attended the design session: a Product Manager, a Retail Designer, two Graphic Designers, and a Product Designer. The goal for the session was to work on the packaging of three products: a box of chocolate pudding, a chocolate bar, and a bottle of spirits. The intended outcome was to evaluate new graphical layout concepts for the packaging of each product.

Food Inc. (name changed to preserve confidentiality) is a food manufacturer specializing in high quality foods and condiments. This company had a pre-existing relationship with Artefice (one of the SPARK consortium industry members) and had relied on Artefice to redesign the packaging on other products they offer.

The session was attended by four members of Food Inc.: the CEO, a Commercial Director, a Marketing Consultant, and a Brand Manager. The goal of the session was to showcase some of the concepts generated during a prior design session (which had been conducted without the use of the SPARK platform) to obtain feedback and decide on a direction to take.

Zobele Group is a manufacturer and supplier of air fresheners and pesticides. Their products are separated into four main categories: “Air Care”, “Home Care”, “Personal Care”, and “Pest Control”. The first three categories include, but are not limited to, air fresheners, deodorants, and fabric softeners. The last category refers to aerosolised pesticides and similar products (Zobele Group, 2014).

Two participants attended the session: a Design Manager and a Product Designer. The purpose of the design session was to evaluate the functionalities of an SAR platform as a tool to use during international trade shows. The participants wished to see how they could use the platform to display various layouts and variations of packaging formats. The feasibility and the satisfactory output of such a display method would influence their decision to pursue such an SAR setup for this purpose. Zobele was particularly interested in a portable version of the SPARK platform for this purpose.

Wavecontrol is an engineering company that produces protective equipment for working in contact with electromagnetic (EM) fields. Their devices are used to detect, monitor, and register EM fields as well as provide workers with advanced warnings of danger as well as measuring levels of exposure to EM radiation (Wavecontrol, 2016).

The objective of the session was to gain better understanding about the integration of touchscreens in a new line of products the company was developing as part of the ONIRIS project they were working on. The session was attended by four participants from Wavecontrol. These were: a Director, a Chief Engineer, a Creative Director, and an Industrial Designer.

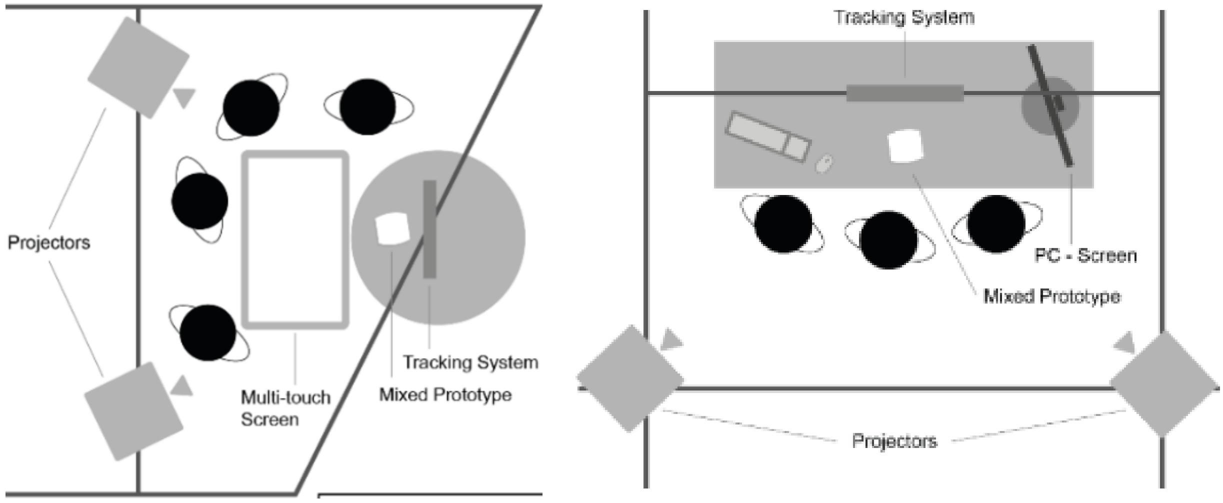
Samsonite is a luggage and suitcase manufacturer and retailer. Their product range spans from small laptop bags and backpacks to large suitcases intended for travel. The products are offered in a wide range of materials, including aluminium, canvas, and plastic.

Two participants joined the session, one Design Director (for Europe) and one Design Manager. The goal of the session was to evaluate different patterns, designs, and textures for two suitcases Samsonite was developing: the Cosmolite and the Neopulse. In doing so the participants hoped to evaluate the utility of the SPARK platform; understanding whether the system could be integrated within their workflow of their corporate design review sessions.

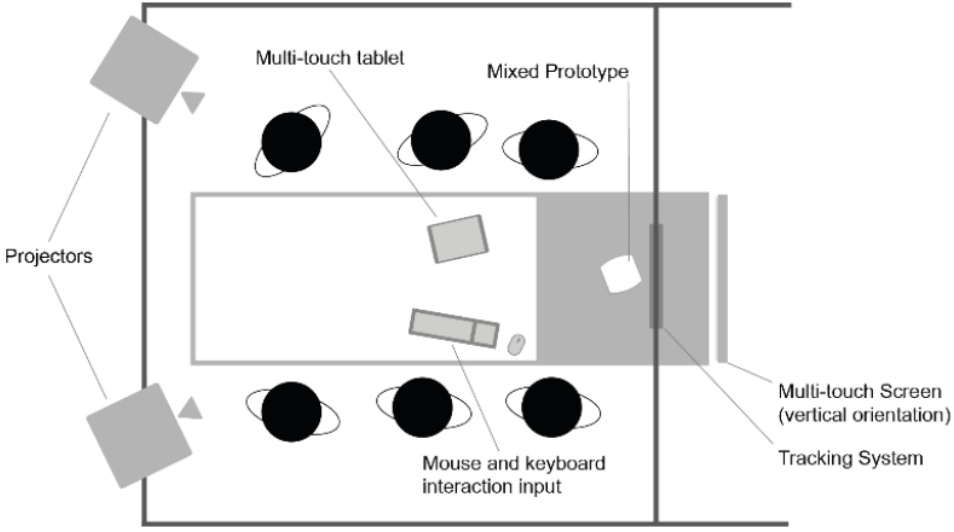
7.2.1.2 *Experimental Setup*

As mentioned in Table 38 the sessions were hosted at three different locations by three members of the SPARK Consortium. The members who hosted the sessions were: Artefice, Stimulo, and AMS. While the SAR platform setup was very similar across all locations, there were some slight variations between each setup as shown in Figure 93.

It is important to note that the SAR platforms deployed at each location made use of the same SPARK software (Section 4.2.2). Furthermore, the setups all made use of two projectors, had an infrared object tracking, and had a touchscreen interface (Section 4.2). As can be seen from Figure 93(a) the setup at Artefice included a large (40-inch) multi-touch screen. This was used as the user interface in the sessions conducted at Artefice. In the sessions at the other locations a regular 10-inch tablet was used as an input device.



(a) Artefice **(b) Stimulo**



(c) AMS

Figure 93. SAR Platform Setup (Majoral *et al.*, 2018)

7.2.1.3 *Data Collection Methodology*

The design sessions were structured in the same manner across all locations, with minor adjustments made to suit the needs of the participants and products they were designing, as listed in Table 38. Participants were presented with some general information on the SPARK project and the SPARK platform, its intended goals and how the system would be used to support the session. The introduction took approximately half an hour. The session started shortly after with the hosting partner acting as a facilitator operating the platform whilst the participants worked on the design task using the assets and models they had provided. The design part of the session took approximately one hour. Once complete, the participants were given half an hour to complete some surveys and were asked to participate in a semi-structured interview aimed at capturing their opinions on the SAR system and its implementation.

Three main strategies were used to capture the events of the sessions. These were:

- Interviews allowed participants to express their views and express any feedback they had.
- A survey to gather information on the participants' opinions where these could not be captured by the interviews (due to, for example, group size) and to guarantee consistency in the data collected across the sessions.
- SAR logs to collect quantitative data on user behaviour and interaction with the SAR platform.

The semi-structured interviews permitted the capture of the participant's opinions in a more organic manner enabling a discussion to emerge where the participants could more easily air their opinions, express their perplexities, and give more nuanced feedback on their experience. The survey enabled more standardised and quantifiable information to be captured. Furthermore, the data collected in the survey allowed all the participants to fully express their views on some specific elements where, otherwise, the size of some of the participant groups might have restricted their ability to all fully contribute to the discussion. The survey consisted of seven questions in total; four aimed at the capture of quantitative data and three open ended questions to allow participants to express any views or opinions they felt were not covered in the interview. The four quantitative questions focused: on the frequency of co-creative sessions, the types of challenges faced by the participants' organization and their importance, the participants' opinions on the features of the SPARK system, and how the participants believed the SPARK system could support their design process. The full text of the survey can be found in appendix section I

The logging system present in the SPARK platform collected information on the types of inputs made during the design session (Section 4.2.2). Using this it was possible to understand how assets were placed, scaled, and rotated. In addition, the background colour of the prototype could be tracked. Table 39 provides a breakdown of the data collected by the logging system and its value for quantifying the participants' behaviour. The analysis presented here is similar to that reported on in section 6.1.2.1 and 6.3.6. In both these sections the log file was used to provide insight into the use of the interface as well as provide information on how the participants interacted with the overall SAR system. In particular, as described in section 6.3.6, the log files shed insight into the asset placement behaviour of participants. In the context of the study presented here, the log files allow for a better understanding of how participants interact with the SAR platform outside of the context of a rigidly controlled study, in a setting that more closely resembles that found in a real-life design session.

Table 39. Log data gathered using SAR platform and significance (Majoral et al., 2018)

PARAMETER	METHOD OF EVALUATION	MEASURES	INSIGHT PROVIDED BY METRIC
#_FUNCT	Total rows in the log	Number of functions initiated or continued in the log file	Total amount of activities within the sessions
#_FUNCT_EFF	Switching between different functions or repeated call of a function after 3 seconds	Number of effective functions	Amount of changes made to the design/mixed prototype during session
#_ASSET	Total amount of previously prepared digital contents used within the sessions	Number of assets used within the sessions	Exploration of alternatives for the mixed prototype used for the evaluation
#_VARIANT	Total amount of functions involving #_asset and changes to background colour of the prototype	Number of solution variants potentially explored	Number of variations tested with a single mixed prototype
#_TIME	Difference between start and end time (seconds)	Duration	The time required to run co-creative session
#_EFFECTIVENESS	Ratio #Time / #_Variant	Session Effectiveness	The time required to switch from a variant to a next one.
#_GUI_EFF	Ratio #Funct_eff / #Funct	GUI efficiency	The efficiency of the SPARK UI in terms of number of changes made to the design with respect to the number of activities requested to the interacting user(s)

7.2.2 Design Session Summaries

The following section provides a brief summary of each of the packaging designs sessions as well as the feedback obtained from each individual session's structured interview. Major points of feedback or learnings are denoted with a number e.g.: #1. This allows for easier tracking of learnings throughout this section. Some numbers may be repeated throughout different sessions. This is to indicate that the same issue has re-occurred in a different session. The section also provides a comparison of the data collected using the surveys and the log file.

7.2.2.1 Colruyt

The Colruyt session focused on three main products. These were a chocolate pudding, a bottle of liquor, and a chocolate bar. The session lasted over an hour to consider all the products presented. For all the products evaluated, the SAR platform was used to replicate the existing packaging onto the physical model. This projection was then modified in accordance with the wishes of the participants.

As mentioned in section 7.2.1.3, a semi-structured interview was conducted after the session. In addition to the interview, the participants were asked to complete a survey. The feedback from the interview showed a generally positive attitude towards the use of SAR systems as well as the SPARK platform itself. In particular, the participants praised the "speed and ease with which they could switch between artefacts in one SAR session" (Majoral *et al.*, 2018) (#1).

However, the participants felt that there was an abundance of technical setbacks and issues during the session (#2). These included software crashes, which led to the GUI needing to be rebooted, as well as the instability of the tracking system which caused the images being projected onto the physical prototype to "flicker". The participants were of the opinion that, prior to adopting any such technology, they would need the system to guarantee more robustness. In addition, they commented that the system's latency and rendering quality should be improved, as they considered them to be average.

One major point made by Colruyt's participants regarding the implementation of an SAR system within their workflow concerned the integration of any SAR software with their existing software suite. The participants mentioned that their workflow relied heavily on the use of design tools such as Adobe's Photoshop and Illustrator. Were they to adopt an SAR platform and attempt to integrate it into their workflow, they would not want it to impair or conflict with their current approach to design but rather wished for it to integrate seamlessly into their current process (#3). Currently, Colruyt relies on these pieces of software to design and prototype all their products, having to export graphics to be integrated into the SAR supported design session, and not being able to take the output of the design session and place it back into either software, was seen as a considerable roadblock as it risked increasing the designers' workload.

7.2.2.2 Food Inc.

Food Inc.'s session was part of an ongoing project between them and Artefice. Food Inc. had previously requested Artefice's help in designing the packaging for a new product of theirs and the two organizations had already completed a session, without the use of any SAR systems, prior to the session discussed here. This prior session had already narrowed down the scope from four potential concepts presented by Artefice for Food Inc.'s approval to two. For the SAR supported session Food Inc. had requested that Artefice adjust some elements of the two concepts selected, incorporating some of the elements present in the other, discarded, concepts. Artefice responded to this request by presenting three revised concepts for Food Inc's consideration. Each concept was then individually presented and evaluated during the SAR supported design session. Unfortunately, during the session itself a technical issue arose. The SAR projection, bright by its very nature, masked some of the subtler texturing present in the concepts being shown. As such, Artefice was forced to rely on the use of printed boards to aid in conveying the concepts to the participants from Food Inc.

The participants from Food Inc. were, technical issues notwithstanding, satisfied with some of the new possibilities that using an SAR system afforded them. Overall, they were pleased with the ability to quickly make changes to the concepts they were presented with and integrate elements from different concepts together (#4). Furthermore, the participants felt the system aided them in being able to co-create and get better feedback on modifications made in real time. One particular application that the participants felt made the most of the SAR technology's potential was the ability to compare their product, as it was being designed, to competitor products in real time (#5). This application, they felt, provided an advantage not only during design sessions, but also in “shelf tests” where the product is placed near its competitors to gauge its ability to stand out from the crowd.

As mentioned, the implementation of the SAR system within the design session was not without difficulties. The brightness of the projected image masked some of the finer details that the designers from Artefice wished to showcase, forcing them to rely on more traditional printed mediums to communicate their concepts’ subtleties. Furthermore, the participants felt the resolution, colour rendering, and latency of the system were considerable drawbacks (#6).

Participants provided some suggestions for improvements to the SAR system. In addition to improvements to rendering and latency, the participants suggested that the user interface be improved, as the tablet was deemed cumbersome. The participants, furthermore, suggested that the projector arrangement be modified. The SPARK system, as set up in the session, consisted of two projectors placed at a short distance from one another. This makes the viewing angle 120°, calculated from the front of the physical prototype where the projectors reach. This means that the rear of the prototype is not illuminated by the projectors. The substantial number of participants, combined with the small room where the session took place, probably influenced this feedback as participants had to crowd on one side of the room to view the model. Despite the frustrations, one positive event was that Food Inc. requested that their next session, to follow up on the decisions made, also make use of the SPARK platform. This shows that, despite the negative aspects of the system, the participants associate a benefit to the use of an SAR platform.

7.2.2.3 Zobele

Zobele’s session had two objectives. The first was to evaluate the effectiveness of using an SAR system as part of their design and design review process. Currently, Zobele’s processes rely on email to communicate and distribute concepts as well as receive feedback. They believe this process is cumbersome and hinders collaborative design. The second was their wish to explore the use of SAR as a device for commercial presentations. Zobele attends multiple design and trade fairs and wished to explore the potential value that SAR’s capabilities could add to these events.

The session focused on the redesign of the packaging for a fragrance diffuser. The current packaging was projected onto the physical model using the SPARK platform and, during the course of the session, the participants made edits to this packaging and evaluated new concepts for the packaging that had been developed by Stimulo, this process, as well as the output, are shown in Figure 94.

Figure 94. Photos from the Zobele session. The photo on the left shows the session in progress with the participants discussing the product and the physical prototype in the background. The image on the right shows a close up of the physical prototype of the diffuser.

Once the session concluded, feedback from the Zobele participants was collected. The participants remarked that the use of SAR compared well with some of the other technologies that they are considering using to support their collaborative design process. The design manager remarked that Zobele had recently been looking for new technologies aimed at better supporting their co-creative process. Two of the technologies that had been considered by them were SKETCHFAB and AUGMENT. Both these tools are aimed at users who wish to build and edit models for augmented and virtual reality with an eye to product visualization. In particular, Zobele was interested in identifying a software or hardware solution that would support designers in visualizing logos, product information, or other marketing material on the packaging during the design process as they claimed this was one of the major bottlenecks encountered during their design process.

Four major points were raised. These were:

- Unlike some immersive VR setups, such as CAVE, SAR does not require a dedicated room; any standard meeting room can be repurposed to host SAR supported sessions. The room can also continue to be used as a meeting room when the SAR system is not in use (#7).
- Secondly, the ability to track the prototype as it is moved is especially useful as it increases the immersion and thus improves the co-creative experience (#8).
- Relatedly, this allows for the packaging to be designed around the actual product, which allows the session participants to feel the real weight of the object and thus also increasing the immersivity (#5).
- Lastly, the participants remarked that, compared to other solutions they had previously evaluated, the projected cost of a SPARK setup was advantageous if the implementation of the system could be kept within the ten to fifteen thousand Euro range (#9).

Despite the positive feedback, the participants remarked that the resolution of the projection left much to be desired, this issue was exacerbated by the small size of the packaging onto which the projection was made (#6). Indeed, due to calibration issues, only one projector was used during this session in an attempt to mitigate the problem. The use of multiple projectors, while useful for improving the SAR illusion by covering more of the physical prototype in the projection, causes issues where the two separate projected images meet causing poor resolution, blurring, and a lack of synchronization between the images projected. The participants also remarked that they were disappointed with the difficulties encountered in switching prototypes during the session. Due to technical issues with the information system, on which the SPARK platform relies, it was not possible to load the second prototype that Zobele wished to work with (#2). Finally, the participants remarked that the colour accuracy of the system was not always close to the desired outcome (#10), with a noticeable difference in the colours shown in the GUI and those projected onto the physical model.

Some of the improvements suggested by the participants related to the ability to export files from the SAR platform in a format compatible with Adobe Illustrator (#3). In addition, the participants noted that there was no tool in the GUI to force the alignment of the graphics placed onto the packaging. As such the accurate placement of graphics, and their alignment with one another, fell entirely within the responsibilities of the GUI operator. Furthermore, the participants remarked on the lack of primitive shapes (triangles, circles, squares, etc.) available for use during a session. The lack of primitive shapes combined with the inability to import assets during the session from repositories complicated and slowed down the initial setup and design (#11). This, in the eyes of the Zobele participants, was a drawback of the system.

Two final remarks were made: the first was about the ability of the SPARK system to support remote sessions to allow collaboration across different locations, as Zobele has multiple offices spread around the world (#12). The second remark made was regarding a “presentation mode” where multiple asset configurations could be looped automatically over the physical prototype to display different variations and concepts at trade fairs without input from a designer (#13).

Despite the issues, the overall feedback was positive. The participants remarked that the system could help them reduce prototyping costs and times by avoiding the need to make unique prototypes for each client. Indeed, upon concluding the session, the participants requested a quote to rent a SPARK platform for their upcoming exhibition. In addition to this marked interest in the standard SPARK platform, the participants mentioned their interest for a more portable setup.

7.2.2.4 Samsonite

The participants from Samsonite were interested in evaluating the utility of an SAR platform to support their design review sessions. A particular concern of theirs was the ability of the system to project well onto a large physical prototype, as they wished to work at a 1:1 scale. In order to achieve this, the physical prototype was placed slightly further away from the projectors. This increased the size of the projector envelope; however, this came with a slight loss in the quality of the projection due to a loss of resolution.

As mentioned in Table 38, Samsonite wished to work on two different suitcase models, the Cosmolite and the Neopulse (shown in Figure 95). The session began with the participants working with the Neopulse suitcase model. On switching to the Cosmolite model, it was necessary to deactivate one of the projectors, as during the calibration phase the blending of the two projectors failed for this prototype. For both the Cosmolite and Neopulse models the participants focused primarily on editing the textures, colour, and materials of the suitcases, rather than working on the more detailed elements of the suitcases.

Figure 95. Samsonite Neopulse model physical prototype with SAR projected texture during the calibration phase. The vertical multi-touch screen shown in Figure 93 can be seen in the background on the right of the image. The interface buttons are visible showing options such as “select part” and “add asset”.

On concluding the session, the participants reported that they felt implementing an SAR system within their design process would result in a reduction in prototyping costs and time. The participants felt that an SAR system would enable them to prototype without relying on a large number of foam models, which are regularly used in their design sessions to model the different suitcase configurations (#4). Through the use of SAR, a single physical prototype can be used to model multiple colour and texture configurations (#1).

Additionally, the participants felt that an SAR system would enable them to better communicate and gather feedback from design review sessions, in particular from those with low design knowledge (#5). These individuals with low design knowledge, felt the Samsonite participants, struggle more with interpreting digital models due to the confusion that arises in displaying the scale and dimensions on a computer screen.

The participants also had some reservations about the SPARK platform. The asset management, handled by the Information System, caused technical issues during the design session. During the session, the participants requested a new texture be applied to the physical prototype (#2). Unfortunately, the asset containing this texture had not been provided, and thus not uploaded to the Information System, prior to the start of the session. This resulted in the session having to stop until the asset could be located and added, which also meant that a new session had to be started, thus having to start from a blank canvas.

The Design Manager from Samsonite specifically commented that, in their opinion, the SPARK platform felt complicated to use and that implementing it might negatively impact the amount of time their design team needed to prepare for a design session (#11). In addition, the technical intricacies of the system worried them, as these could halt a session whilst technical errors were resolved. It was suggested that perhaps the presence of a technical expert, fully trained in the use of the platform, would be required to ensure its smooth operation.

On the basis of these concerns, as well as other observations made during the session, the participants made a number of suggestions for future improvements to the SPARK platform. Firstly, the participants mentioned that it would have been nice to modify textures and colours independently of one another. During the session it was not possible to place an asset as a texture and then modify the hue or colour of this asset (#10). It was only possible to change the colour of the background onto which the assets are placed (the blank physical prototype in essence). This means that if a user wishes to review the same texture, but in three different colours, they must have three assets uploaded at the start of the session: one for each desired colour. This also means that if at any point during the session a request is made to view an asset in a new colour, it requires the session to be stopped for a new asset to be uploaded. It was precisely this that lead to the session being paused.

As with other sessions the participants felt that, in order to be seriously considered for adoption, an SAR platform would need to be able to integrate with the existing software and design tools already in use at the company (#3). In addition, increasing the asset and colour management capabilities of the system was identified as a potential improvement. The participants were also concerned about the quality and resolution of the projected images (#6). Finally, the participants mentioned that they would have appreciated having shadow rendering around the placed assets, which would have increased the realism and immersion as well.

Despite the suggested improvements, the representatives from Samsonite felt that an SAR platform, such as the SPARK platform, could be a useful tool in reducing prototyping costs and efficiency of their monthly design sessions. Their interest was such that they requested to be kept updated on any developments regarding the SPARK platform.

7.2.2.5 Wavecontrol

The principal objective for the design session with Wavecontrol was the design and test of a novel user interface for a new product. This product was intended to sport a large touchscreen, which is unlike any other product offered by the company. As such there were limited pre-existing expectations or predispositions regarding the layout and setup of the product. This enabled a more expansive exploration of the design space.

Unlike some of the physical prototypes used in the other sessions the physical prototype used in Wavecontrol's session was uncharacteristically well suited to SAR projection. The size of the prototype, when scaled at a one-to-one scale, allowed for easy projection as it was neither so big to require adjustments to expand the projection envelope nor so small that resolution became an issue. Furthermore, the shape of the physical prototype, rectangular with flat surfaces, made it easy for the projection to map to the object with little distortion. As such the participants quickly engaged with the prototype, feeling comfortable with handling it, and passing the prototype amongst themselves. This led to them being able to review the different user interfaces quickly and effectively. Furthermore, the participants felt able to request and analyse changes with little difficulty.

The participants were impressed with the level and immediacy of the feedback they could obtain whilst designing using an SAR platform (#5). Many of the comments made referred to how the participants felt that they would be able to gain better feedback from members of industry. One example provided highlighted this advantage. The devices designed by Wavecontrol are often intended for field use, where end-users will wear gloves. Being able to interact with the physical prototype, whilst wearing gloves, enables a more realistic impression of how handling their new product would feel to an end user (#8). This leads to a better fine tuning of the ergonomics, not just for the interface, but for the device as a whole.

The success of the session and the positive impact the SAR platform had on the participants had some drawbacks, however. The representatives from Wavecontrol were quick to suggest many technical improvements aimed at increasing the number of features and functionalities of the SPARK platform. These suggestions centred primarily around improvements for the user interface of the platform as well as the addition of certain functionalities, such as the ability to display animations in addition to static images (#11). Furthermore, the participants felt the infrared tracking system was limited. The risk of occlusions happening whilst the physical prototype is being manipulated is high and can quickly result in the projection and the prototype becoming misaligned (#2). This breaks the immersion and can be a jarring or dissuasive experience for those handling the prototype.

Nonetheless, the overall verdict was positive with the participants remarking that the SPARK system was “immersive, fast and dynamic” (Majoral *et al.*, 2018) (#4). The general consensus was that the system could be used to support design sessions and more easily gather feedback from end-users as well as other stakeholders, in particular less technically minded ones. The Wavecontrol’s director in particular remarked that the system could prove advantageous to aiding the marketing and design teams better communicate during the initial stages of a project (#5).

7.2.2.6 Summary

Thus, the main learnings obtained from the participants’ qualitative feedback for each of the sessions were:

#1 Participants praised the speed and ease with which they could switch between artefacts in one SAR session.
#2 The participants felt that there was an abundance of technical setbacks and issues during the session.
#3 An SAR platform should not impair or conflict with their current approach to design but rather wished for it to integrate seamlessly into their current process and workflow.
#4 Participants were pleased with the ability to quickly make changes to the concepts they were presented with and integrate elements from different concepts together.
#5 The participants felt the system aided them in being able to co-create and get better feedback on modifications made in real time.
#6 The participants felt the resolution, colour rendering, and latency of the system were considerable drawbacks.
#7 Participants thought that, unlike some immersive VR setups, such as CAVE, SAR does not require a dedicated room; any standard meeting room can be repurposed to host SAR supported

sessions. The room can also continue to be used as a meeting room when the SAR system is not in use.

#8 The ability to track the prototype as it is moved is especially useful as it increases the immersion and thus improves the co-creative experience.

#9 The participants remarked that, compared to other solutions they had previously evaluated, the projected cost of a SPARK setup was advantageous if the implementation of the system could be kept within the ten to fifteen thousand Euro range.

#10 The participants remarked that the colour accuracy of the system was not always close to the desired outcome.

#11 Cumbersome and feature poor GUI. Lacking in ability to customise and modify assets on the fly while providing poor feedback on user actions and their eventual outcomes.

#12 The ability of the SPARK system to support remote sessions to allow collaboration across different locations was advantageous.

#13 A “presentation mode” where multiple asset configurations could be looped automatically over the physical prototype to display different variations and concepts at trade fairs without input from a designer would add an additional use scenario to the platform.

7.2.3 Survey Data from Sessions

This section provides the results from the questionnaires presented to the participants of the packaging and product design sessions. As described earlier in section 7.2.1.3, each questionnaire contained seven questions, of which four were aimed at collecting quantitative data from the participants. The final three questions in the questionnaire were open ended questions aimed at allowing session participants to express any views or opinions they did not feel they had the chance to lay out during the structured interviews.

The first question in the questionnaire related to the frequency with which the company undertook co-creative design sessions. The answers from the three groups varied considerably. Colruyt organised one session per week; Samsonite stated their sessions occurred approximately once per month, whilst Zobele and Wavecontrol only had one per quarter.

The second question asked participants to rate a number of challenges by their importance for their organization. These challenges, and the average responses for each organization, are shown in Figure 96. Participants from each company were asked to score each challenge on a Likert scale from 1 to 5 with 1 being not very important and 5 being very important. The responses from each participant were then averaged to obtain a score for each company. As can be seen from Figure 96 the majority of participants rated the challenges as important or very important. As such it appears that the SPARK system, which is geared towards addressing these issues, is indeed designed to address problems that companies find relevant to their operations. Samsonite additionally felt that the reduction of the cost of prototypes was particularly important. Wavecontrol was slightly less focused on the other challenges but did remark that generating novel ideas was pivotal. The discrepancy between Samsonite’s focus on the cost of prototypes when compared to Wavecontrol can, in all likelihood, be linked to the frequency with which they conduct these co-design sessions. As Samsonite runs sessions at thrice the rate that Wavecontrol does, it stands to reason that the associated costs may weight more on their minds.

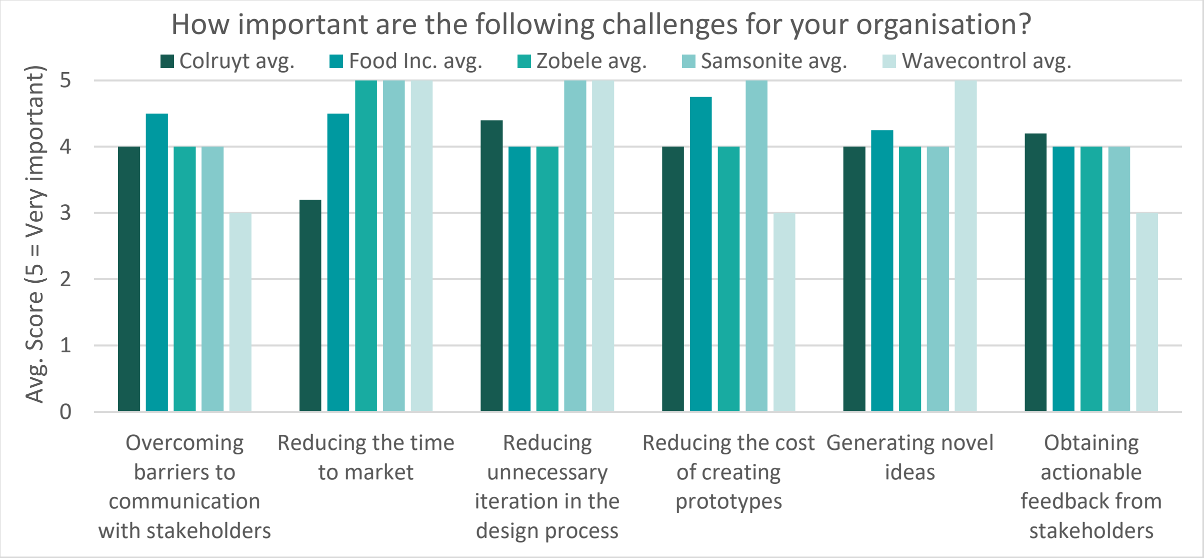


Figure 96. Ratings by the five organizations regarding the importance of six different challenges. Scored on a Likert scale from 1 to 5 with 1 being not at all important and 5 being very important. (Majoral et al., 2018)

Similarly, to question two, question three asked participants to rate the features of the SPARK system. Five features were identified and again participants were asked to rate each one on a Likert scale from 1 to five, with 1 being very poor and 5 being excellent. The results from this question are shown in Figure 97. There was little agreement between the companies on their impressions of the SPARK system's features. While Zobele was most impressed by the ease of use and resolution these were least appreciated by Colruyt. In contrast Colruyt had a positive opinion of the colour rendering, projection alignment, and latency. This is in contrast to Zobele. Food Inc.'s impression of the SPARK system's features appears to have been fairly positive; all the features were scored as above average.

Despite the enthusiastic response by Wavecontrol during and after the session, the feedback provided in the survey showed a more tempered response. Both companies rated the SPARK system's features as average or slightly above average. Interestingly, the participants of the product design sessions rated the ease of use higher than the participants from the packaging design sessions. This is despite rating the other features at the same or lower level when compared to the participants from the packaging design sessions.

These discrepancies can be, in part, explained by the fact that each session took place at a different location, and that in each session different types of technical issues arose to hinder the progress made in the design session. One additional contributing factor to this variation in the responses could be the initial expectations of each participant group. The expectations may have coloured the participants' opinions and thus influenced their reactions when technical issues arose, or they were faced with the current limitations of the system. It should also be noted that the participants' use case for the SPARK platform at their own organization would likely have influenced their expectations and thus their feedback on the various features of the SPARK system. For example, an organization that wished to adopt the SPARK platform as a tool for demonstrating products at trade fairs and other public facing events would be much likelier to value resolution than a company that planned to use the platform internally.

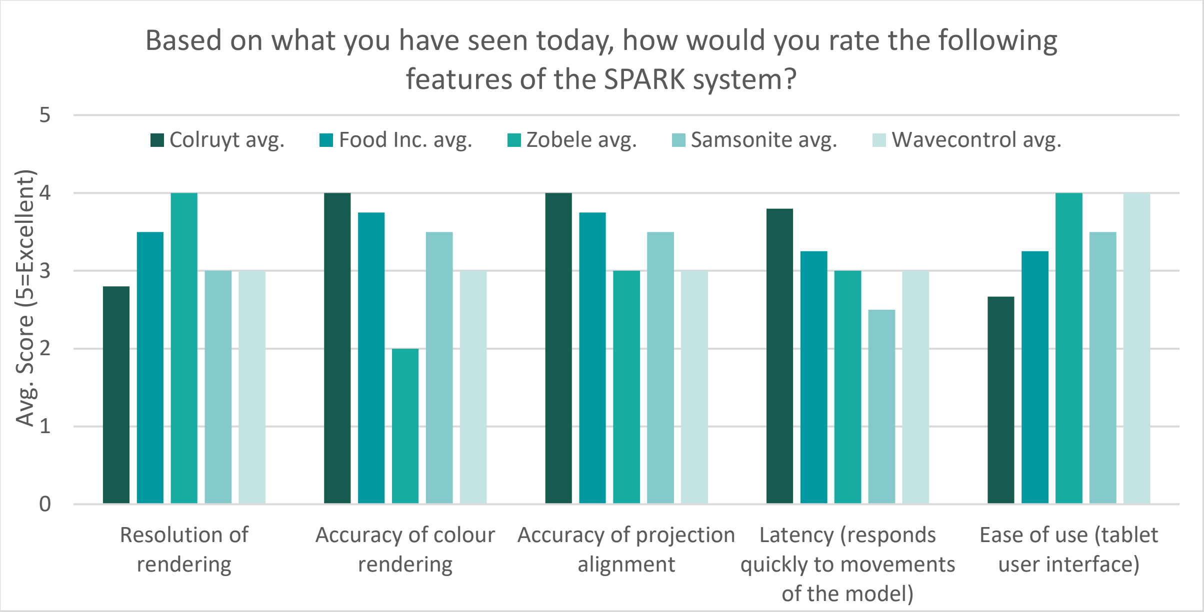


Figure 97. Ratings by the three organizations regarding the importance of five different features of the SPARK system. Scored on a Likert scale from 1 to 5 with 1 being very poor and 5 being excellent. (Majoral et al., 2018)

The results to question four are shown in Figure 98. From the graph it is possible to see that the companies agreed that the implementation of an SAR system could improve their ability to review and filter ideas as well as reduce the time to market. Food Inc. was particularly interested in the ability of an SAR platform to aid in reducing the overall labour and prototyping costs as well as feeling that the implementation of SAR would be beneficial to the company as a whole. Colruyt was poorly disposed to the use of SAR as a tool for idea generation, nor were they convinced it would be easy to implement, but they did feel the system would lead to a reduction in costs, both for labour and prototyping expenses. In contrast, Zobele was not convinced that the system would result in lower costs but did think it would be easy to implement. It is interesting to note that Zobele rated the system as least beneficial overall for the company but was the only company who did ask for a quote to implement the SAR system directly at an upcoming event. Samsonite felt that the largest impact the SAR platform would have would be in the realm of reducing prototyping costs and, to a lesser extent, time to market. They were less convinced that the system could be used for improving idea generation, however. Wavecontrol instead believed the system could aid them in idea review and filtering. Interestingly both companies agreed that the SPARK system would be easy to implement, despite having previously, as shown in Figure 97, shown less conviction in the features of the system.

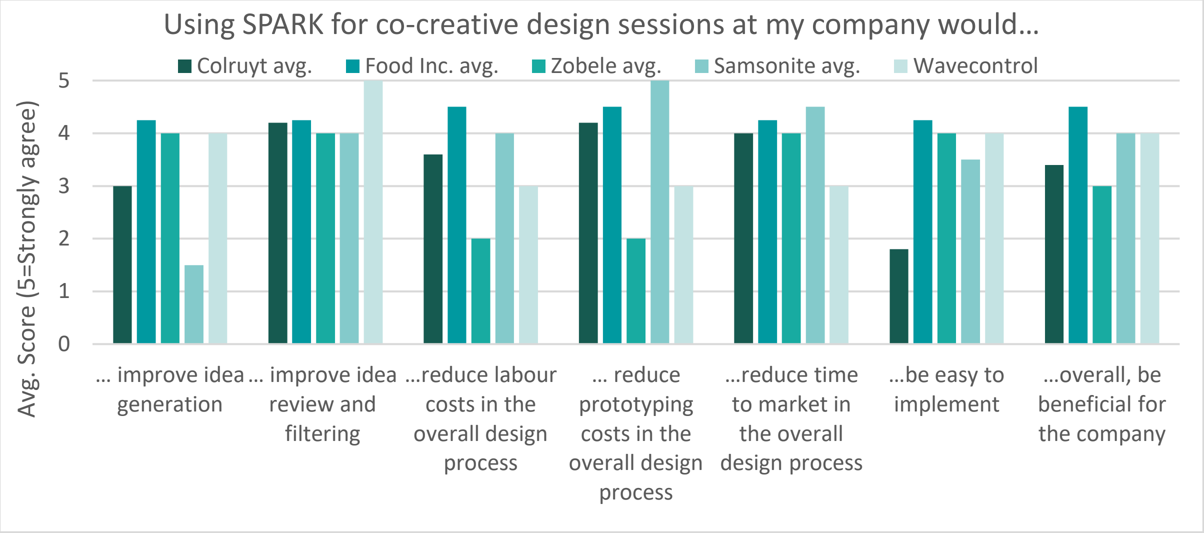


Figure 98. Ratings by the three organizations regarding the importance of seven different potential applications of the SPARK system. Scored on a Likert scale from 1 to 5 with 1 being strongly disagree and 5 being strongly agree. (Majoral et al., 2018)

7.2.4 Log Data from Sessions

Table 40 shows the results obtained from the log files using the parameters listed in the Data Collection Methodology (section 7.2.1.3).

The session with Colruyt was omitted and the session with Food Inc. was divided in two. The results from these sessions can be analysed as one. This was due to errors with the logging system that impacted the generation of the logs and thus the data collection.

Table 40. Packaging design log results for Food Inc. and Zobele sessions (Majoral et al., 2018)

PARAMETER	INSIGHT PROVIDED BY METRIC	FOOD INC. (1)	FOOD INC. (2)	ZOBELE	SAMSONITE	WAVECONTROL
#_FUNCT	Total amount of activities within the sessions	1101	741	2378	1777	3634
#_FUNCT_EFF	Amount of changes made to the design/mixed prototype during session	472	229	482	346	602
#_ASSET	Exploration of alternatives for the mixed prototype used for the evaluation	37	17	18	13	17
#_VARIANT	Number of variations tested with a single mixed prototype	377	140	354	186	412
#_TIME (SECOND)	The time required to run co-creative session	3839	1828	2477	6074	8212
#_EFFECTIVENESS (SECOND)	The time required to switch from a variant to a next one.	10.2	13.1	7	32.7	19.9
#_HCI_GUI (%)	The efficiency of the SPARK UI in terms of number of changes made to the design with respect to the number of activities requested to the interacting user(s)	42.9	30.9	20.3	19.5	16.6

From Table 40 it is possible to gain some insights into the general progression of each session. The #_Time parameter shows that sessions lasted from as little as thirty minutes to well over an hour. However, interestingly, it seems #_Funct does not scale directly with the amount of time the session lasts. The Zobele session, which was of middling duration, had more function calls than the Food Inc. session one, which was the longest. Furthermore, the Zobele session had more #_Funct_Eff calls than any other session. This observation is of particular interest when combined with the fact that the #_Asset value for the Zobele session is comparable to that of the Food Inc. session two. Since the number of assets present in the session was comparatively low it stands to reason that the participants of the Zobele session used the time to mix the available assets to a greater extent and that thus the complexity of a session is not limited exclusively by the number of assets available to the participants. This is reflected by the #_Effectiveness value which shows the amount of time needed to complete each variant. This is lowest for the Zobele session.

One other potential explanation lies, however, in the #_HCI_GUI value. This is lowest for the Zobele session which implies that the designers in this session struggled with implementing their concepts. Lastly, it is important to note that the sessions had different goals and that this would have, in all likelihood, influenced the approach of the participants. Indeed, this is already visible in the discrepancy between Food Inc.'s first and second sessions. As the second session was aimed more at narrowing down and filtering the concepts generated in the first session, it stands to reason that it would involve less idea generation and less time as the designers merely need to display the pre-determined concepts.

The Samsonite and Wavecontrol sessions, also experience considerable variability. The physical prototype used in the Wavecontrol case study was one of the largest, by volume, ever used with the SPARK platform and the Samsonite physical prototypes were even larger. The objectives for each of these sessions also differed in purpose: Samsonite expected to conduct a design review whereas Wavecontrol wished to experiment with interface design. These differences are evident when one compares the #_Assets and the #_Variants. Is should also be noted that these sessions lasted for considerably longer than the sessions conducted with Zobele or Food Inc.

Both the Wavecontrol and Samsonite sessions showed that the SPARK platform can support the exploration of a large number of options over a short span of time. In particular, the Wavecontrol session showed a very high #_Variant score, the highest amongst all sessions with a score of 412. The #_Effectiveness scores were higher in the Wavecontrol and Samsonite sessions this is likely due to the session requirements. The placement of assets in a design session that focuses predominately on the product, rather than the packaging, needs to take less care with the careful placement and arrangement of assets.

7.2.5 Summary of Industry Feedback from Pilots

In total five pilot studies were run with members of industry. In each of these sessions participants from companies worked on developing their products whilst supported by an SAR system. Three of these pilots focused on the packaging design aspect, whereas the remaining two were aimed at product design. The pilot studies conducted enabled the collection of feedback data both in the form of qualitative and quantitative feedback. These studies were run in order to further address RO-5: “Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform” and provide additional insight into RQ-4: “What are the industry requirements for an SAR system to support co-design?”.

Participants from both types of sessions highlighted their frustration with the technical difficulties experienced during the sessions. Furthermore, participants highlighted the need for the SAR system to integrate more directly with their existing design tools as well as their existing design process. Nonetheless, the participants expressed their interest in the system explaining that, if properly implemented, an SAR platform could greatly aid in communicating with end-users.

It is interesting to note that both types of participants, those who focused on packaging and those who focused on product design, felt that reducing the time to market and unnecessary iterations were amongst the most important challenges faced by their organizations. Similarly, both types of participants gave similar ratings regarding the SPARK system’s performance during their respective sessions. Major differences only begin to appear when the expected use cases are discussed.

Participants in the packaging design sessions reported that implementing an SAR system in their design sessions would enable them to improve their idea generation as well as their ability to review and filter ideas. Furthermore, the participants all agreed that the implementation of an SAR system would decrease their time to market. There was less agreement amongst the participants on the other categories, with some reporting they felt that an SAR platform would be easy to implement and others reporting they felt it would decrease prototyping costs.

Participants in the product design sessions were generally in agreement that the implementation of an SAR platform in their design process would improve idea review and filtering. However, unlike the participants in the packaging sessions, the participants in the product design sessions felt that an SAR system would not be particularly useful in reducing time to market. Furthermore, the participants were generally more convinced that the system would be easy to implement and that it would provide a benefit to their organization.

In summary, the participants were fairly satisfied with the potential of the SAR system and how it supported them during their respective sessions. Technical issues aside, the participants were able to identify the potential benefits of such a system to their organizations. However, the participants did identify some key issues and challenges they felt the system would need to address. In particular, participants were interested in seeing a closer integration between the platform and their existing design tools.

The technical issues experienced during some of the sessions appear to not have, overly, negatively coloured the opinions of the participants. The presence of these issues must nonetheless be addressed should the platform wish to become a commercial product. Both stability issues, such as crashes, and a lack of features, such as integration with design tools, can be addressed through additional development of the software. The issues encountered are not indicative of an inherent limitation of the technology that cannot be overcome without fundamental change. As such any improvements required would require a refinement of the technology rather than a return to the drawing board approach.

One interesting observation that can be made as a result of the data collected is that, despite somewhat different use cases, the challenges both types of participants felt their organizations faced were similar. This indicates that an SAR platform would not necessarily need to be tailored to one type of industry. This is supported by the participants from both the packaging and the product design sessions rating the SPARK platform in similar ways across the various categories. It may also be that the two industries are similar enough that the distinction between them is less relevant than originally thought.

Thus, the key findings were:

KF-30. Participants from both types of sessions highlighted their frustration with the technical difficulties experienced
KF-31. Participants highlighted the need for the SAR system to integrate more directly with their existing design tools as well as their existing design process.
KF-32. Participants were pleased with the ability to quickly make changes to the concepts they were presented with and integrate elements from different concepts together.
KF-33. The participants felt the system aided them in co-creating and getting better feedback on modifications made in real time.
KF-34. Participants saw technical advantages in using a SAR platform over VR but also highlighted shortcomings due to the immaturity and lack of features present for the technology

KF-35. Participants felt that reducing the time to market and unnecessary iterations were amongst the most important challenges faced by their organizations

KF-36. Participants were able to identify the potential benefits of such a system to their organizations

7.3 LONGITUDINAL ANALYSIS OF SAR IMPACT ON THE DESIGN PROCESS

The following section discusses studies into the long-term impact of an SAR platform. The two studies presented here showcase the impact that utilizing an SAR platform has on the design process over the course of the entire development of a product, rather than over the course of a single design session. These studies build upon those presented in Chapter 5 and were originally published in "Results of The Experiments Benchmarking The Platform" (Ben-Guefreche *et al.*, 2018) and "D5.1 Validation At End Users' Premises" (Bellucci *et al.*, 2018a). Namely section 7.3.1 builds on the work and results presented in section 5.1. Section 7.3.2 is similarly based on the work and results presented in section 5.2.

Section 7.3.1 lays out the design process efficiency metrics developed to analyse the impact of an SAR platform on the design process. It analyses three historical design cases to develop these. These historical cases were provided by Artefice, one of the SPARK consortium members, and involved one of their clients, Food Inc. Food Inc. had been a participant in the studies mentioned in sections 5.2 and 7.2. Section 7.3.2 uses the metrics to analyse the results from two similar design projects to highlight where the application of SAR has impacted the design process. The work presented in this section serves to provide additional evidence to address RO-3: "Evaluate the efficacy of an SAR platform in Complete Co-Design Sessions" and RO-5: "Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform". In doing so, the section provides additional evidence for the answering of RQ-4: "What are the industry requirements for an SAR system to support co-design?"

7.3.1 Design Process Efficiency in Historical Design Cases

As part of the comparison study aimed at benchmarking the SPARK platform against competing design methodologies, discussed in section 5.1, an analysis of the design process of three design cases was conducted. All the design cases analysed had been completed by the time the analysis was conducted. The purpose of this study was to provide a benchmark which future studies could then use to understand the impact that the SAR platform had.

7.3.1.1 Design Process Efficiency Metrics

The metrics used to evaluate the design process efficiency are shown in Table 41. In total five metrics were used. These metrics attempt to provide an understanding of the time and resources required to bring the design process to completion. The amount of time taken is analysed from two perspectives: Person-Hours Spent on Project and Lead Time. The Lead Time provides a general overview of the duration of the project, from the initial brief to the end of the idea generation phase. The Person-Hours Spent on Project enables a more refined understanding of the amount of time the designers actually spent on the project during the Lead Time. Where Lead Time may be prolonged or shortened due to external factors the Person-Hours Spent on Project enables a more refined understanding of the level of "crunch time" that may otherwise not be captured by a simple analysis of the Lead Time. The Total Development Cost and Cost of Prototype Production metrics capture the amount of money the design agency has had to pour into the development of the designs and their respective prototypes. Lastly, the Re-Work Iterations metric represents the number of times that a design activity had to be repeated until it was of a level deemed satisfactory to move to the next stage of the design process.

Table 41. Design Process Efficiency Metrics (Ben-Guefreche et al., 2018)

METRIC TITLE	**DEFINITION**
PERSON-HOURS SPENT ON PROJECT	All hours spent on project by design agency (including unbilled hours)
LEAD TIME	Number of days between project start date and end of the Ideas Development phase
TOTAL DEVELOPMENT COST	Direct costs incurred by design agency (Only up to End of layout - Ignoring post-production costs)
COST OF PROTOTYPE PRODUCTION	Cost of preparing all design representations used in collaborative sessions or sent to end-user (materials and labour)
RE-WORK ITERATIONS	Total number of co-creative design sessions completed within the project
	Number of versions of the design released to the client up to the end of the creative phase
	Average cost for each version release calculated as the total cost of development divided over the number of iterations for each prototype

7.3.1.2 Analysis of Historical Cases

The three design projects within Artefice analysed were for a pizza, a yoghurt, and a soup. The design processes for each are summarised in Figure 99, Figure 100, and Figure 101 respectively. Each of these three projects was provided by Artefice. These case studies were selected as they were considered to be representative of the type of work that would be conducted using an SAR platform with regards to packaging design. The three design projects all had the same client company, referred to here as “Food Inc.” for anonymity. This company is a regular client of Artefice and as such the two organizations have often worked together on these and other projects. This presents an advantage when analysing future sessions using the metrics developed here as it will not be necessary to control for the client type or attitude to the design project, simplifying any future comparisons. Due to time constraints and the difficulty of traveling to the various industry partners, it was decided to limit the longitudinal analysis to Artefice, who mostly focuses on packaging design. Stimulo, the other industrial partner of the SPARK consortium, who focuses more on product design was also consulted. This was to ascertain whether, should the need arise, an analysis using the same metrics could be conducted of one or more of their design processes. As Stimulo was, in theory, able to provide the data, but not immediately, a decision was made to focus primarily on Artefice’s case studies.

The timelines shown in Figure 99, Figure 100, and Figure 101 were compiled to understand the progression of a typical design project using historical data. It is important to highlight that the sessions being analysed in these figures were not conducted using any form of SAR. The analysis of these sessions was merely to provide a baseline against which SAR supported design sessions could be compared. Each figure shows the stage-gates of the design, going from the initial project brief to the final output. As part of this the timelines also highlight the sessions with the clients where they provide feedback on the design under development and filter these.

Each of the figures is divided into four sections. The topmost section shows the project name as well as a brief description of the scope of the project. Thereunder, in the bar, the design process is noted down. Lines extend down from this bar to separate the concepts generated into three categories (ideas production, ideas development, and ideas execution) mapped against the design process. The timeline section of each figure notes the dates of major events. Finally, the iterations portion of the figure holds all the information pertinent to the prototype iterations developed. Each dot represents an iteration and has information, such as the iteration version, date, and additional information. The dots are interconnected to show the evolution of the design.

In addition to the data shown in the figures, information was collected to deploy the metrics shown in Table 41. Information regarding costs and expenses is considered confidential. As a result, the values themselves cannot be made available. However, the percentage difference across the two conditions can be reported; percentage differences in the values for sessions conducted with and without SAR support will be provided in the relevant sections.

Figure 99 shows the design process for the development of a package for frozen pizza. As can be seen eight versions were created over the course of two months. The design process shows two distinct phases focused on filtering concepts. The first phase begins with the project brief, where a number of different concepts are generated to suit said brief, and slowly converges until the client agrees. Based on the client meeting new proposals are generated and once more filtered down to a single final proposal.

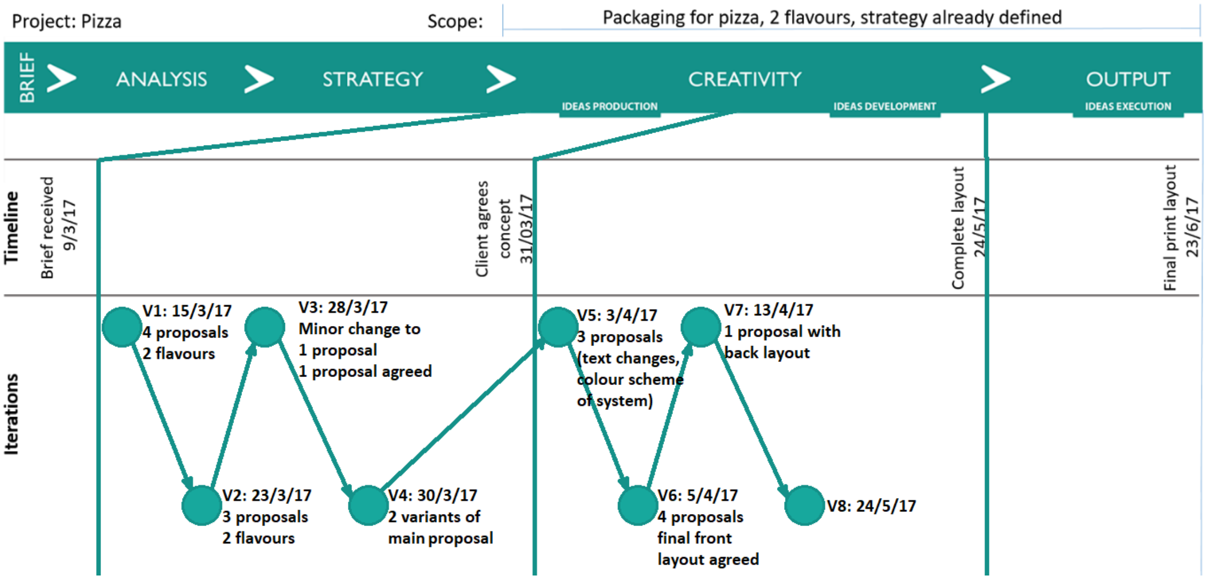


Figure 99. Design Process for a Frozen Pizza Packaging without the use of SAR (Ben-Guefreche et al., 2018)

In contrast with the relatively smooth process shown in Figure 99 for the development of a pizza packaging, the development of a yoghurt packaging, shown in Figure 100, was more turbulent. The process overall took four months and a total of twenty-one iterations were generated. The lead up to the meeting where the client gives their approval shows considerably more rework of the ideas generated with the numbers of layout proposals increasing and decreasing over time. However, once the client had given their approval for the concept, the design process appears to become more streamlined with a simple filtering process resulting in one. Despite this it is possible to see that the number of iterations is still quite high.

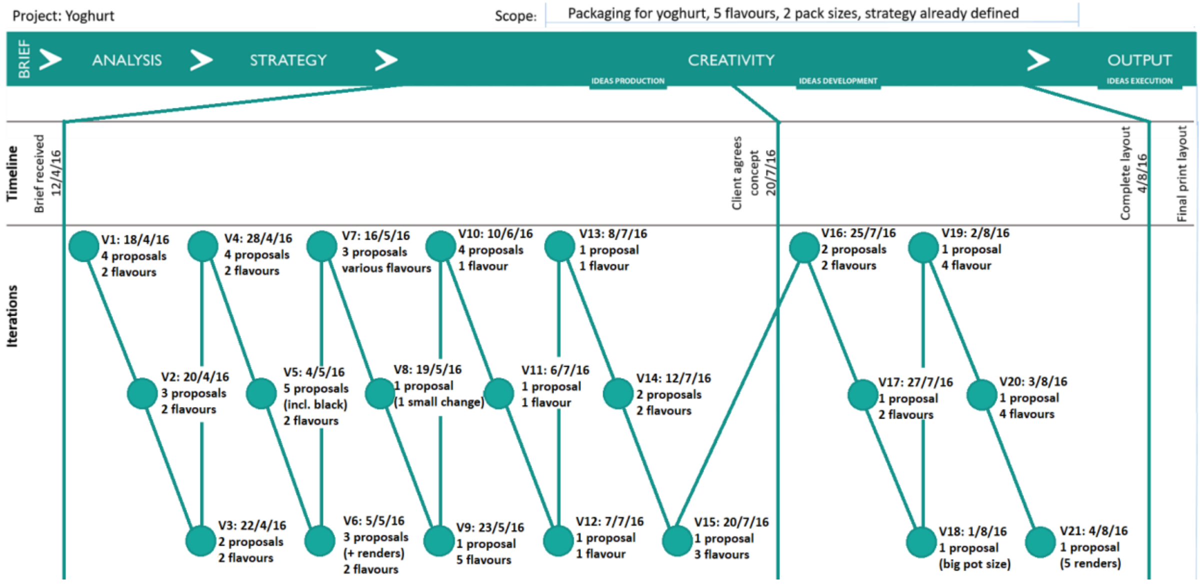


Figure 100. Design Process for a Yoghurt Packaging without the use of SAR (Ben-Guefreche et al., 2018)

Figure 101 shows the design process for the development of a soup package. This design process appears to be more in line with that seen for the pizza packaging. The entire design process took a little over two months and resulted in the development of thirteen iterations. There appeared to have been less uncertainty in the design of the soup packaging than for the design of the yoghurt packaging as can be seen in the number of iterations. In addition, it appears that the design process here followed a divergent-convergent model as the number of proposals seems to change with each iteration. First decreasing with the second iteration, and then increasing with the third iteration, before decreasing once more with the fourth. This follows the general pattern of convergence and divergent design generation.

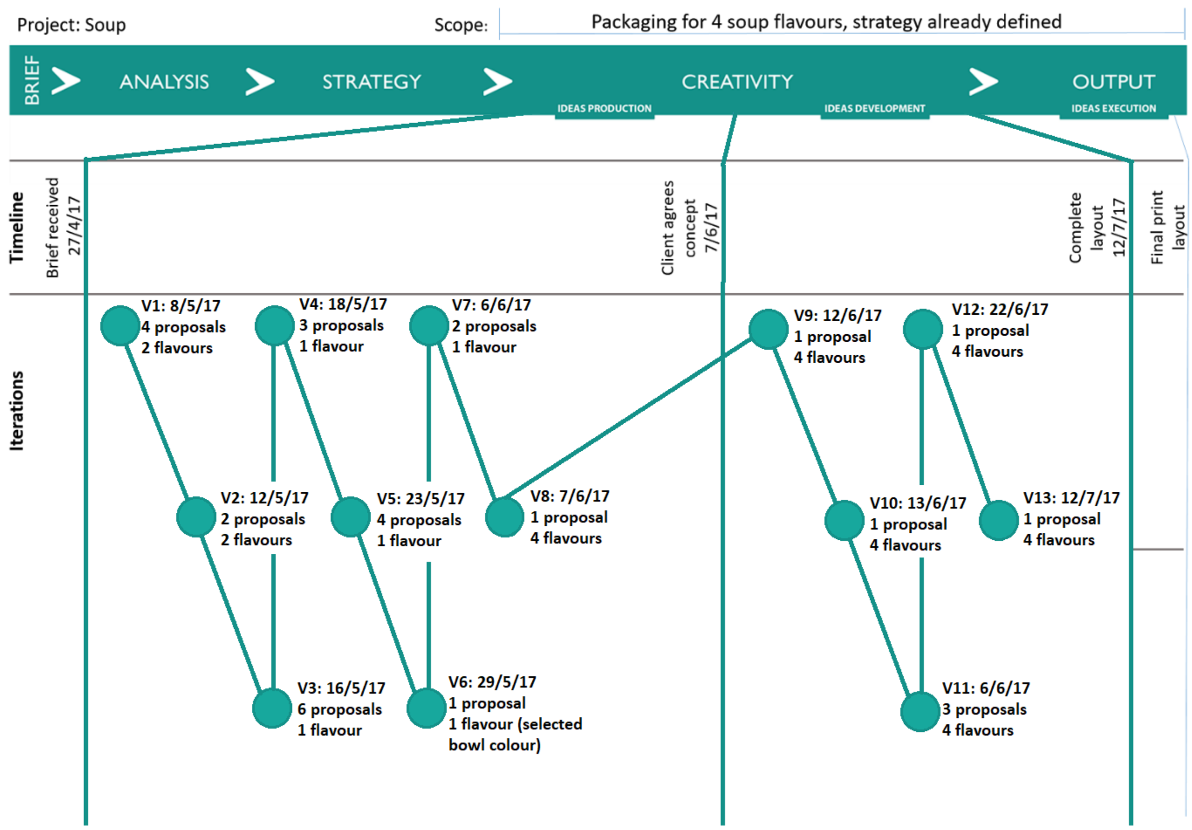


Figure 101. Design Process for a Soup Packaging without the use of SAR (Ben-Guefreche et al., 2018)

7.3.2 SAR Platform Impact on Design Process Efficiency

The analysis of historic design projects conducted without the support of an SAR platform provides a baseline for the comparison of future sessions performed using an SAR platform. It is possible in this way to understand the potential impact that an SAR platform has on the design process overall. As part of the data collection effort reported in section 5.2, it was possible to perform an analysis using the design process efficiency metrics listed in Table 41.

Five design sessions were conducted during the course of the experiments reported in section 5.2. Of these three were with Artefice. Interestingly, Artefice was continuing work with Food Inc. on their development of a fresh pizza packaging. This provided an excellent opportunity to perform a comparison between the historical frozen pizza project analysed in 7.3.1.2 which was also developed in conjunction with Food Inc. The similarity of the design challenge was fortuitous as it decreased the number of assumptions that had to be made regarding the complexity of the design by providing an opportunity to compare projects of equivalent complexity. Furthermore, the fact that the client, Food Inc., was unchanged between the two case studies enabled a reduction in the number of assumptions regarding the expected design output.

The metrics, defined in Table 41, are applied to both the frozen (non-SAR supported design process) and fresh (SAR supported design process) pizzas in Table 42. It should be noted that there is no value available for metric subsection on ‘total number of creative design sessions’ for the non-SAR design process. This is because Artefice’s procedure for this product was to present their work via email to the client and then receive feedback in the same medium. As such no physical meetings occur and a direct comparison of the two is not possible.

Data was collected for each metric through interviews with members of Artefice. They relied on their internal reporting tools, such as timesheets, email chains, and quotes to collect the necessary data.

Table 42. Comparison of Process Efficiency Between SAR and non-SAR Supported Design Processes (Bellucci et al., 2018a)

METRIC	DEFINITION	FROZEN PIZZA (NON-SAR)	FRESH PIZZA (SAR)	PERCENTAGE CHANGE
PERSON -HOURS SPENT ON PROJECT	All hours spent on project by design agency (including unbilled hours)	58.25	39	-33%
LEAD TIME	Number of days between project start date and product launch date	100	21	-79%
TOTAL DEVELOPMENT COST	Direct costs incurred by design agency	Confidential		-20%
COST OF PROTOTYPE PRODUCTION	Cost of preparing all design representations used in collaborative sessions or sent to client (materials and labour)	Confidential		+47%
RE-WORK ITERATIONS	Total number of co-creative design sessions completed within the project	N/A	1	N/A
	Number of versions released up to end of creative phase	8	4	-50%
	Average cost for each version release	Confidential		+34%

The results reported in Table 42 show that the implementation of an SAR platform has a significant impact on the design process, in particular on the amount of time required. The implementation of the SAR platform reduced Lead Time and Person-Hours by 79% and 33% respectively. This can be directly linked to the reduced number of Re-work Iteration, which was halved, prior to agreeing on a final layout with the client. The overall reduction in time also resulted in an overall reduction in the Total Development Cost, by 20%. This is despite the increase in the Cost of Prototype Production, which increased by 47%. This increase in cost can be linked to the increased time required to setup the SAR sessions. When interviewed about the amount of time required to perform this step, the designers involved in setting up the session reported they needed ca. 10.5 hours. This was to prepare the physical model, the UV map, and the assets. It should, however, be noted that this was the first time setting up an SAR supported session for this designer. The comparatively long amount of time required for the setup can, in part, be attributed to this.

The increased costs for prototypes also had a direct impact on the Average Cost for Each Version Release which increased by 34%. It is also important to note that this comparison does not take the

SAR platform's cost into account. This was deemed an acceptable assumption as the SAR platform is not single use and can be implemented in other design projects. In order to calculate the average return on investment for the purchase of an SAR platform additional projects would need to be analysed. This would, however, be complicated by the fact that there is no standardised SAR setup, and thus no definitive cost for the platform. For these reasons, the initial SAR platform cost was omitted from the analysis. Nonetheless it should be stressed that the overall cost for the design process, as well as the time taken, were reduced considerably. As noted, the SAR supported design session experienced a 20% saving over a non-SAR supported designs session. As such the expectation would be that, over the course of multiple design sessions, the SAR platform would pay for itself.

Thus, the key findings were:

KF-37. The metrics adapted to analyse impact of SAR on design process rather than design session appear to provide insight into the longitudinal design process

KF-38. The implementation of the SAR platform reduced Lead Time and Person-Hours by almost four fifths and one third respectively

KF-39. The overall reduction in time also resulted in an overall reduction in the Total Development Cost, by one fifth.

KF-40. This is despite the increase in the Cost of Prototype Production, which increased by almost half

KF-41. The increased costs for prototypes also had a direct impact on the Average Cost for Each Version Release which increased by one third

7.4 OBSERVATIONS ON THE SET OF INDUSTRY CASE STUDIES

Table 43 shows a summary of the findings from the studies presented in this chapter and compares them to the research questions first highlighted in section 3.4.2. The experiments detailed in this chapter were, as discussed in section 3.4.4, Table 12, aimed at addressing RO-5: "Capture industry input to the development of an SAR platform, and analyse their response to the implementation of the SAR platform". The experiments achieved this objective through the broader analysis of potential interested parties at trade fairs, detailed studies with participants from industry, and the cross comparison of SAR and non-SAR supported design projects to understand the long-term impact of an SAR system. In focusing the experiments on these three items, it was possible to explore the interest of industry in the adoption of SAR systems.

It should be noted that the industry most consulted during the course of the studies reported on in this section, as well as the thesis overall, was design consultancies. Stimulo and Artefice, the two SPARK consortium members who provided feedback as well as coordinated with their own clients to provide feedback, are design consultancies. The use of design consultancies may have had an impact on the type of design sessions analysed as well as the type of clients who participated in the more in-depth studies. Nonetheless, the use of design consultancies as the basis for the studies conducted provides a good underpinning for the data gathered, with the potential for future studies to analyse different sectors to assess whether the industry requirements may change.

Table 43. Research Outcomes Compared to Research Questions

RESEARCH QUESTIONS	RESEARCH ACTIVITIES		
	7.1 Industry Feedback from Trade Fairs	7.2 SAR Platform Pilots with Industry Members	7.3 Longitudinal Analysis of SAR Impact on the Design Process
RQ-4: “What are the industry requirements for an SAR system to support co-design?”	• A majority of respondents rated each of the six challenges as important or very important to their organization • Respondents frequently engaged in co-design sessions with stakeholders or customers	• Participants from both types of sessions highlighted their frustration with the technical difficulties experienced • Participants highlighted the need for the SAR system to integrate more directly with their existing design tools as well as their existing design process. • Participants were pleased with the ability to quickly make changes to the concepts they were presented with and integrate elements from different concepts together. • The participants felt the system aided them in co-creating and getting better feedback on modifications made in real time. • Participants saw technical advantages in using a SAR platform over VR but also highlighted shortcomings due to the immaturity and lack of features present for the technology • Participants felt that reducing the time to market and unnecessary iterations were amongst the most important challenges faced by their organizations • Participants were able to identify the potential benefits of such a system to their organizations	• The implementation of the SAR platform reduced Lead Time and Person-Hours by 79% and 33% respectively • The overall reduction in time also resulted in an overall reduction in the Total Development Cost, by 20%. • This is despite the increase in the Cost of Prototype Production, which increased by 47% • The increased costs for prototypes also had a direct impact on the Average Cost for Each Version Release which increased by 34%

8 DISCUSSION AND CONCLUSIONS

Chapter 8 summarises and discusses the results obtained in Chapters 5, 6, and 7. The chapter begins by providing a summary of the major findings for each of the studies. These are then combined into two matrices to aid in understanding how each of the studies, and by extension their findings, have contributed towards achieving the research objective and answering the research questions. The section continues by providing a discussion of the impact that the research has on the development of future SAR systems for co-design as well as a reflection on the studies' contributions towards achieving the research aim. The chapter concludes by providing a set of recommendations for the development and implementation of SAR platforms to be used in the support of collaborative design sessions as well as discussing how the work presented in the thesis may be expanded on in future research.

8.1 SUMMARY OF KNOWLEDGE GAINED THROUGH EXPERIMENTS

This section serves to provide an overview of the major findings from the eight studies performed as part of this thesis. The major findings are cross referenced to the relevant research questions and research objectives in order to map how each of the studies answered the research aim originally set out in section 3.4.1.

8.1.1 Summary of Findings

Table 44 provides an overview of all the findings reported throughout all of the studies performed as part of this thesis. It should be noted that Table 44 reports only on the findings from each of the studies; Table 45 provides a summary of the methodological insights gained from the studies. These insights are kept separate in order to aid the reader in understanding how the research questions were answered, and the research aim achieved, as this understanding relies on the results presented in Table 44. The insights presented in Table 45 serve to provide information regarding the execution of the studies as well as provide additional context, such the impact of specific experimental design choices, for future researchers.

Table 44. Summary of Findings for Each Reported Study

STUDY	SECTION	CONDENSED FINDINGS
Literature Review	2	#1 AR, and in particular SAR, have a potential for application in the field of design. There is currently a lack of sufficient research to understand how SAR could impact design sessions. #2 Currently it is not clear what, if any, causal links exist between the use of SAR and design outcomes #3 No pre-existing metrics were identified that were developed specifically to analyse the impact of a technology on design or the design process. Literature did support the extrapolation of new metrics from pre-existing ones.
Developing Design Research Metrics	2.6	#4 Related metrics were identified that could, if adapted, be adjusted to evaluate the impact of SAR on design sessions. These were customised to meet the expected needs of an SAR session. Testing of these customised metrics revealed their utility and effectiveness in gauging the impact of SAR on design sessions.

Comparing SAR and existing Co-Design Tools	5.1	#5 SAR supported design sessions may have better outcomes than standard design sessions - In particular for Quantity and Quality of Ideas #6 Presently, SAR struggles to support Idea Filtering activities #7 Hardware instability can cause considerable delays and difficulties in the sessions #8 Colour correction and calibration are very difficult to effectively implement when using SAR due to ambient light interference, different colour standards between devices, and the inherent difficulty of projecting dark colours #9 Interaction with the SAR physical model by participants was lower than originally expected
SAR Platform Validation at End-Users' Premises	5.2	#10 Session progress is hindered by poor colour accuracy and visibility #11 Physical prototype size affects the session participants' design and behaviour #12 Scaling up physical prototypes may improve the resolution and rendering of the SAR effect.
Interface Comparison Study	6.1	#13 Accuracy of position, rotation, and scale was comparable between conditions relying on the use of a touch interface (UV Map, Touch Area, and 3D View) to that of the conditions that made use of a mouse interface. #14 The discrepancy in accuracy, in particular the larger variance, found between UBATH and POLIMI could be attributed to the different SAR platform setups used at the two locations. However, the results for conditions A (PC/Mouse) and B4 (PC/Mouse), which were identical at both locations, show a larger variance at UBATH. This may indicate that the SAR platform setup is not the sole contributor to the discrepancy. #15 Conditions UV Map, Touch Area, and 3D View showed comparable levels of accuracy. This indicates that accuracy is independent of these methods of interaction. #16 The analysis of the Participant Efficiency was inconclusive at determining the most efficient interface. There was discrepancy between the results from the two universities, UBATH and POLIMI, which renders a firm conclusion impossible. It was interesting to note, however, that participants from UBATH completed the task in less time than those from POLIMI in all conditions. Whether this is linked to the different SAR setup is, as of yet, unclear. #17 The participants' perception of usability was significantly higher for conditions A (PC/Mouse) and B4 (PC/Mouse). This is to be expected as the mouse and computer interface would have been the most familiar interface with which to complete the task. #18 Of the touch interfaces, condition UV Map was rated with the highest perceived usability.
Impact of SAR on Communication	6.2	#19 Sessions using SAR were noticeably shorter than sessions run using a 3D digital representation on a PC screen.

between Design Session Participants		#20 The percentage of interaction between participants with the shared design representation was higher in the SAR scenarios than in the 3D model scenarios indicating that the SAR system better supports communication in the shared space. #21 Clients in the PC sessions had to rely more on the use of verbal cues to guide the designer to compensate for this. This may also be linked to the amount of data that can be transmitted between participants during the interaction, with the SAR model supporting larger amounts of data transmitted.
Impact of Scale in Design Sessions Supported by an SAR Platform	6.3	#22 The physical model scale used plays a role in how design session participants accomplish their task, influencing how they utilise the assets provided. #23 The scale of the model did not play a major role in how the designers experience the SAR platform itself and their experience whilst designing. #24 The scale of the model appears to have played a role in the designers' satisfaction with the concepts generated. #25 There may be an “optimum” scale at which specific effects, such as exploration or concept generation, are more pronounced. #26 The study identified a consistent "waterfall" asset-by-asset layering design behaviour when generating concepts with SAR.
Industry Feedback from Trade Fairs	7.1	#27 A majority of respondents rated each of the six challenges (listed in Table 37) as important or very important to their organization #28 Respondents frequently engaged in co-design sessions with stakeholders or customers
SAR Platform Pilots with Industry Members	7.2	#29 Participants from both packaging and product design sessions highlighted their frustration with the technical difficulties experienced #30 Participants highlighted the need for the SAR system to integrate more directly with their existing design tools as well as their existing design process. #31 Participants were pleased with the ability to quickly make changes to the concepts they were presented with and integrate elements from different concepts together. #32 The participants felt the system aided them in co-creating and getting better feedback on modifications made in real time. #33 Participants saw technical advantages in using a SAR platform over VR but also highlighted shortcomings due to the immaturity and lack of features present for the technology #34 Participants felt that reducing the time to market and unnecessary iterations were amongst the most important challenges faced by their organizations

		#35 Participants were able to identify the potential benefits of such a system to their organizations
Longitudinal Analysis of SAR Impact on the Design Process	7.3	#36 In the case studies presented, the implementation of the SAR platform reduced Lead Time and Person-Hours by 79% and 33% respectively #37 The overall reduction in time also resulted in an overall reduction in the Total Development Cost of 20%. #38 This is despite the increase in the Cost of Prototype Production, which increased by 47% #39 The increased costs for prototypes also had a direct impact on the Average Cost for Each Version Release which increased by 34%

Table 45 contains some of the methodological insights that were gathered during the execution of the experimental studies. Not all studies are represented in the table as the observations presented resulted predominately from unexpected behaviours or events during the studies' course.

Table 45. Summary of Methodological Insights

Study	Section	Major Insights
SAR Platform Validation at End-Users' Premises	5.2	#40 Fatigue should be taken into account more when designing experiments, participants should be allowed sufficient breaks and time - low task progress was noted for most of these sessions
Interface Comparison Study	6.1	#41 By analysing the time taken and the accuracy, it was possible to note that participants completing condition PC/Mouse were less engaged in the task. This is likely linked to the repetitiveness of the task which caused them to feel disinvested and uninterested. #42 The experiments showed that providing a training scenario to the participants to familiarise them with the interface and task did not create a learning effect that would impact the results of a comparative test. This enables future experiments to be undertaken with participants completing multiple conditions or training cases. #43 The results from both the CSI and SUS surveys were similar, with neither providing additional insight over the other, in order to reduce the time taken to administer subsequent tests it would be advisable to utilise only one.
Longitudinal Analysis of SAR Impact on the Design Process	7.3	#44 The metrics adapted to analyse impact of SAR on design process rather than design session appear to provide insight into the longitudinal design process

Comparison of Findings to Research Questions

Table 46 shows how the findings, listed in Table 44, were used to answer the research questions. The numbers in each cell link to those shown in Table 44 and Table 45.

Table 46. Comparison of Major Findings from each Research Activity to Research Questions

Research Questions	Research Activities									
	Literature Review		Supporting Co-Design through Spatial Augmented Reality		Studies into SAR System Characteristics and Features			Studies with Industry		
	Literature Review	Developing Design Research Metrics	Comparing SAR and existing Co-Design Tools	SAR Platform Validation at End-Users' Premises	Interface Comparison Study	Impact of SAR on Communication between Design Session Participants	Impact of Scale in Design Sessions Supported by an SAR Platform	Industry Feedback from Trade Fairs	SAR Platform Pilots with Industry Members	Longitudinal Analysis of SAR Impact on the Design Process
RQ-1: "How can co-design sessions' efficacy be measured?"	SAR appears to have value as a tool to support design; but a knowledge gap is apparent #1 #2 #3	Existing design session metrics need adjustment to be used to evaluate SAR technology #4								It was possible to successfully adapt metrics to analyse SAR impact on the design process rather than design session #44
RQ-2: "How does an SAR system affect co-design sessions' efficacy?"			Compared to other tools such as AR and classical design tools SAR provides better outcomes in particular with the Quantity and Quality of Ideas #5	Low task progress noted in some SAR supported sessions. Other results consistent with previous findings. #40	Participant analysis was inconclusive at determining the most efficient interface type #16	The ability and freedom of participants to interact with the physical prototype of an SAR system has a positive impact on their ability to perform a design task #19 #20 #21	CSI scores remained unvaried both when scaling up and down for all conditions tested (1:1, 2:1, 3:1, 1:10, and 1:20) #23			
RQ-3: "How do specific SAR characteristics and features affect co-design sessions' efficacy?"					The type of interface, touch or keyboard and mouse, does not seem to influence accuracy when designing #13 #14 #15	SAR sessions took less time, indicating better efficiency of the process. More verbal interaction needed in the PC sessions while ephemeral interaction remains constant across sessions. #19 #20 #21	The scale of the physical prototype appears to have an impact on the quantity of ideas generated, time taken, and the exploration of the assets available #22 #24 #25		Participants highlighted the need for SAR systems to integrate more directly with their existing design tools and processes #30	
RQ-4: "What are the industry requirements for an SAR system to support co-design?"				Limitations encountered due to physical prototype size and colour leading to decreased quality of projection impacting the session. #10 #12				There is an interest from those in industry who do engage in regular co-design sessions for an SAR system to aid them in this process #27 #38	Reducing time to market and unnecessary iterations were amongst the most important challenges faced by respondents #34	SAR use can result in cost and time savings throughout the entire design process, despite higher prototype costs #36 #37 #38 #39

8.2 CONCLUSIONS

The preceding section provided an overview of the key learnings collected through the experiments conducted as part of this thesis. These learnings were compared to the research objectives and research questions in a number of tables. It should be noted that the wide variety of experiments and studies conducted, with a diverse pool of participants, contributes to the robustness of the data collected and the key learnings identified. In the following section a breakdown will be provided to analyse how each learning has addressed the individual research questions and objectives, thereby providing a clearer understanding of how this thesis contributes to the knowledge space surrounding SAR as applied to collaborative design.

8.2.1 Objectives

Six research objectives had originally been set out in section 3.4.3. These were:

RO-1: "DEVELOP A METRIC FRAMEWORK FOR EVALUATING THE EFFICACY OF CO-DESIGN SESSIONS"

RO-1 was predominately addressed by experiments performed as part of the Literature Review, the Developing Design Research Metrics analysis and the experiment focusing on the Longitudinal Analysis of SAR Impact on the Design Process. The Literature Review provided the basic underpinnings for the development of the metric framework, in particular by identifying existing metrics as well as providing insight into co-design sessions. The Developing Design Research Metrics section provided additional insight into the development of the metrics, showing an iterative approach to balance the need for coding live design sessions with the theory. This was done in order to achieve metrics that could be easily and rapidly implemented without compromising their utility and accuracy. The work performed as part of the Longitudinal Analysis of SAR Impact on the Design Process differed slightly as here the focus was on the development of metrics to analyse the entire design process, rather than a single design session.

RO-2: "DESIGN AND DEVELOP AN SAR PLATFORM FOR USE IN CO-DESIGN SESSIONS"

RO-2 was entirely tackled as part of chapter 6: Studies into SAR System Characteristics and Features. The three experiments described in the chapter, namely the Interface Comparison Study, the Impact of SAR on Communication between Design Session Participants, and the Impact of Scale in Design Sessions Supported by an SAR Platform showed the implementation of a functional SAR research platform as originally described in chapter 4: The Development of an SAR Research Platform. The development of the SAR platform discussed in chapter 4 ran concurrently with the experiments discussed in chapter 6; the platform evolving in response to the issues encountered through each experiment and being updated as required to address the issues identified.

RO-3: "EVALUATE THE EFFICACY OF AN SAR PLATFORM IN COMPLETE CO-DESIGN SESSIONS"

RO-3 was tackled by the studies performed as part of the Supporting Co-Design through Spatial Augmented Reality experiments (chapter 5) as well as SAR Platform Pilots with Industry Members and Longitudinal Analysis of SAR Impact on the Design Process which were reported on in chapter 7: Studies with Industry. With the exception of the studies reported on in the Longitudinal Analysis of SAR Impact on the Design Process, the focus for all the experiments was on analysing the impact that SAR had on individual design sessions. The studies reported on in chapter 5 specifically focused on the impact the SAR platform had on the outcomes of the design sessions. SAR Platform Pilots with Industry Members focused instead on evaluating the experience that members of industry had with the SAR platform in order to better understand which challenges and opportunities they felt existed and thus how to promote adoption of the platform. Finally, the Longitudinal Analysis of SAR Impact on the Design Process analysed the impact of the SAR platform from the perspective of the overall design process rather than the individual design session. This study was conducted to provide a more holistic approach to understanding the impact of SAR.

RO-4: "ANALYSE THE IMPACT OF A SAMPLE OF SPECIFIC CHARACTERISTICS AND FEATURES OF THE SAR PLATFORM ON CO-DESIGN SESSIONS"

RO-4 again fell entirely within the experiments performed as part of chapter 6: Studies into SAR System Characteristics and Features. The three studies conducted as part of this chapter analysed three characteristics/features of SAR, namely the SAR platform interface (Interface Comparison Study), the communication (Impact of SAR on Communication between Design Session Participants) and the scale of the physical prototype (Impact of Scale in Design Sessions Supported by an SAR Platform).

RO-5: "CAPTURE INDUSTRY INPUT TO THE DEVELOPMENT OF AN SAR PLATFORM, AND ANALYSE THEIR RESPONSE TO THE IMPLEMENTATION OF THE SAR PLATFORM"

RO-5 was fully achieved by the studies that were part of chapter 7: Studies with Industry research activities. The three studies reported in this chapter collected feedback from members of industry to better understand both the potential for adoption for SAR as a tool to support design as well as provide insight into the requirements for said adoption. The Industry Feedback from Trade Fairs focused predominately on gathering the general opinions and perspectives from a larger population sample. The SAR Platform Pilots with Industry Members focused specifically on the input provided by a select number of industry professionals who provided insight into how SAR could be adopted by their specific industry as well as providing feedback on the requirements for said adoption. Finally, the Longitudinal Analysis of SAR Impact on the Design Process provided insight into the long-term advantages that adopting SAR could have for members of industry.

RO-6: "PROPOSE RECOMMENDATIONS/GUIDELINES FOR THE DEVELOPMENT OF SAR PLATFORMS FOR CO-DESIGN"

Lastly, RO-6 was addressed by all the studies performed with the exception of "Industry Feedback from Trade Fairs" and "Longitudinal Analysis of SAR Impact on the Design Process". Each study reported on in chapters 5, 6, and 7 provided insights into the development of guidelines and recommendations for the development of SAR platforms. This was either in the form of direct feedback from the study participants or from the experiences of running said studies which often revealed shortcomings in the technology.

8.2.2 Research Questions

In achieving the research objectives, it was possible to answer the research questions which were laid out in section 3.4.2. In the interest of legibility and to aid the reader, links are provided to the relevant findings in Table 44 .

RQ-1: "HOW CAN CO-DESIGN SESSIONS' EFFICACY BE MEASURED?"

As shown in Table 46 multiple findings from the studies performed served to answer the research questions. The literature review performed enabled the answering of RQ-1. The literature review did not identify an established set of metrics for evaluating the impact of technologies on design sessions (#1-#3). However, the literature review did establish that existing metrics aimed at similar tasks could be adapted to analyse the impact of SAR on design sessions (#4). These metrics were developed and tested as part of the research establishing the validity of the SPARK project research. The metrics were further adapted to analyse the impact of SAR on the design process rather than the individual design session (#44).

RQ-2: "HOW DOES AN SAR SYSTEM AFFECT CO-DESIGN SESSIONS' EFFICACY?"

The major findings from each of the studies conducted found that, for RQ-2, there is a link between the use of an SAR platform and co-design sessions efficacy. In particular the Quantity and Quality of ideas appears to be positively influenced by the use of an SAR platform (#5). Low task progress was also noted in some sessions (#40). The ability of participants to interact with the physical prototype seems to have an impact on the participants' ability to perform during the design session (#19 - #21). However, the CSI scores did not change with the scale of the physical prototype indicating that the presence of the prototype is more important than the specific scale (#23).

RQ-3: "HOW DO SPECIFIC SAR CHARACTERISTICS AND FEATURES AFFECT CO-DESIGN SESSIONS' EFFICACY?"

In answer to RQ-3, it appears that the user interface does not play a major role in how SAR impacts a co-design session's efficacy (#13-#15). However, the scale of the physical prototype does appear to play a role (#22, #24), in particular the Quantity, time taken and the exploration of the assets available appears to be impacted by the scale of the prototype used (#23, #25). Communication also appears to be aided by the presence of an SAR platform (#19 - #21). Industry feedback highlights the need for integration between the platform and the existing tools and processes of companies (#30).

RQ-4: "WHAT ARE THE INDUSTRY REQUIREMENTS FOR AN SAR SYSTEM TO SUPPORT CO-DESIGN?"

In the case of RQ-4, the industry requirements seem to have focused predominantly on the ability of an SAR platform to integrate with the existing tools and processes currently in industry (#30). Furthermore, it was shown that members of industry are predominately attracted to the idea of implementing an SAR platform into their design process as they expect to be able to use it to reduce iterations and time to market (#27, #34). Adoption of the technology, however, requires improvements in the quality of the projection and stability of the platform (#10, #12, #29). The viability of this was also proven not just by the analysis of single design sessions but also through a longitudinal study that showed a decrease in time and overall costs when using an SAR platform over the course of a full design development process (#36 - #39).

8.2.3 Summary

In achieving the research objectives and answering the research questions it is possible to claim that the aim has been achieved. While further work is required, in particular to analyse the impact of additional features and characteristics of SAR, it appears that, based on the evidence collected, a deeper understanding of the impact that SAR has on co-design sessions is now possible.

As SAR is still not a fully-fledged technology, much of this research has focused on understanding how SAR can be optimised for the future. As such, the aim of this thesis morphed from a direct investigation of a wide band of characteristics and features into a narrower exploration of the impact that specific features and characteristics can have. Furthermore, the effect of SAR on co-design received considerable focus in order to provide a more holistic understanding of how SAR impacts the overall design session and process. This was deemed important as the work performed now allows future researchers the opportunity to delve into specific characteristics and features with more confidence, knowing that their investigations will be focused on characteristics and features that will have a long-term impact on the adoption of SAR technology or the improvement of an SAR platform.

The major findings highlighted by the work performed as part of this thesis allow for an increased understanding is not just for the overall effect that an SAR platform can have on co-design sessions but also on how specific characteristics and features, namely the interface, the physical prototype scale, and the communication impact within co-design sessions.

In conclusion, the following findings were identified as the major learnings from the work performed:

- The use of SAR in design sessions appears to have a positive impact on the design session outcome.
    - In comparison to other AR and classical design tools, SAR performs better; in particular where Quality and Quantity of Ideas is concerned.
- Participants interacting during a co-design session seem to communicate more effectively when using SAR, requiring less time and discussion to bring their point across.
- The use of a physical prototype has an impact on the design session, the presence of a physical object over a fully virtual representation aids with communication.
    - However, the scale of the physical object can impact the ability of the design session participants to generate ideas as well as explore the design space.
- There are major advantages for industry where they to adopt SAR throughout the design process.
    - Cost and time savings are particularly pronounced.
    - These savings are, however, tempered by the current lack of integration of SAR with existing design tools as well as the overall infancy of the technology.

8.3 LIMITATIONS OF WORK PERFORMED

The wide variety and number of studies conducted as part of this thesis contributes to the robustness of the data collected. In addition, many of the studies were conducted in different settings, thereby expanding the dataset, and strengthening the generalizability of the conclusions drawn. However, this variety may also have caused some limitations. Many participants in the studies, in particular those conducted with members of industry as well as Artefice and Stimulo, were not native English speakers. As a result, it is possible that language barriers hindered effective communication between participants in the studies when these were conducted in English. In other experiments, the design sessions were conducted in the participants' native language. This however had the drawback of translation into English for analysis. Nonetheless, the advantage gained by having multinational participants and the advantages this brings in mitigating culturally biased findings outweighs the complexities of running these experiments.

The majority of experiments were performed in controlled conditions. The experiments reported on in subsections 5.1, 6.1, 6.2, and 6.3 all aimed to control as many independent variables as possible in an attempt to isolate the causal or correlational link between the independent variables being tested. However, this resulted in design sessions that were not necessarily perfect representations of real design sessions. As such it is possible that the effects that were found may not be as strong or prevalent in regular design sessions, where external factors may play a role. Furthermore, all these aforementioned sessions, in particular those reported on in subsections 6.1, 6.2, and 6.3, were conducted using students as the main participant group. While all those selected to participate in the experiments were required to have a minimum degree of proficiency in design or engineering, it is possible that professional designers would have behaved differently, if only because of their greater experience in tackling common design challenges. However, simply discounting the results obtained, on the basis that the participants were not professional designers, would also be misleading. While the participants may not have been experienced as designers, they would have been representative of novices just starting their careers as professional designers.

An additional limitation of the work performed as part of this thesis include the limitations regarding the independence of the designers in the sessions performed as part of the experiments aimed at Supporting Co-Design through Spatial Augmented Reality and the Studies with Industry. In both these sets of experiments the design cases studied were put forward by Stimulo and Artefice, two design agencies that were in the SPARK consortium. Furthermore, the two agencies also led their own design sessions for all the experiments mentioned. As such it was not possible to perform any form of blinding to prevent bias. Despite this limitation, it should be noted that, in the experiments conducted with real clients, the client would likely have prevented excessive formation of bias in favour of the SAR platform.

8.4 RECOMMENDATIONS FOR SAR PLATFORM DEVELOPMENT

Table 47 and Table 48 summarise recommendations for future development of SAR platforms. These are based on the feedback received during the course of the experiments as well as from the experimental data itself. The purpose of this table is to provide a general set of “dos and don’ts” to inform whoever may, in the future, wish to develop SAR technologies for the purpose of supporting co-design sessions.

Table 47. Summary of Recommendations for Improvement of SAR Platforms (Basic Technical Problems to Be Addressed)

NR.	RECOMMENDATIONS (BASIC TECHNICAL PROBLEMS TO BE ADDRESSED)
1	Colour accuracy needs to be improved. The colour seen on the user interface should match the colour projected as closely as possible.
2	The SAR platform should seamlessly integrate with existing design tools such as Adobe Photoshop and/or Illustrator.
3	It should be possible to export the results of an SAR supported design session to a common format (.pdf, .png, .tiff, etc.) in such a way that it can easily be shared with others.
4	When interacting with the physical prototype the digital overlay should update as quickly and seamlessly as possible to avoid breaking the illusion.
5	Calibration of the tracked object should be as fast and painless as possible as this is often a major bottleneck to setting up an SAR session and can limit the types of physical prototypes that are used.
6	GUI should be clear and responsive. It should be obvious what each option does, and if an error is made it should be easy to backtrack.
7	GIF/animated image support should be available in order to prototype interactions and allow the physical prototype to appear responsive to user inputs during user testing.
8	Physical prototype tracking should occur in such a way that it is as stable as possible, minimizing the impact that occlusions have on the tracking.
9	Resolution needs to be improved in such a way that projection onto small surfaces remains accurate.
10	Improved cable management to reduce the number of power supplies needed as well as the number of cables required to operate the SAR platform, in particular for any portable SAR system
11	Multi input GUI, enabling touch interfaces, mouse and keyboard, or other input types.
12	Localised power supply to allow a portable platform to be deployed outdoors or in locations where mains power is unavailable.

Table 48. Additional Observations from the Experience of Doing the Research (Features and Affordances to Be Developed Further)

NR.	OBSERVATIONS (FEATURES AND AFFORDANCES TO BE DEVELOPED FURTHER)
1	Rudimentary design tools should be available to the designer within the GUI. The designer should not have to rely exclusively on pre-uploaded assets but be able to create them as they work.
2	Interactivity tools should be added to simulate, for example, an interface. Designers should be able to create the illusion of interactivity when participants interact with the physical prototype
3	Colour projection range should be improved. At present dark colours in particular are very hard to project; a method of projecting these in well lit rooms would greatly improve the colour rendition of the SAR platform
4	As the scale of the physical prototype appears to impact the designers' choices during the design process, it may be valuable to have a tool that assists the designer in choosing the scale of said physical prototype based on the constraints they have and the desired session outcomes.
5	Multiuser input system to allow multiple designers to simultaneously work on the same design without having to pass the GUI amongst themselves.
6	Automated deployment system for a portable SAR platform. This would allow the users to save calibration settings (such as projector distance) and re-use these as necessary when on location.

8.5 FUTURE WORK

Future work should be aimed at better cataloguing and analysing the impact of additional features and characteristics of SAR. Of particular interest would be a more detailed exploration of the impact of scale on design outcomes, aimed specifically at identifying if an optimum scale exists for specific design tasks. Additionally, understanding how to better support the user experience surrounding the use of SAR in design would greatly improve the adoption potential of the technology. In doing so it will be possible to better understand how SAR can be adjusted to gain more traction as a design tool.

It is clear that the experiments presented in this thesis are able to answer the research questions. However, due to the small sample sizes used and the small number of conditions tested firm conclusions regarding the actual impact of SAR on co-design using only the data provided here are limited. Future work should also focus on the replication and expansion of the experiments presented in this thesis.

In alternative to this approach, were the goal to promote the adoption of SAR in design, rather than to understand the causal relationship between SAR use and design output, an iterative approach could be pursued. Such an approach would focus on collaboration with industry partners to gain a deeper understanding of their needs and the limitations they experience when using SAR. An iterative approach such as this, focusing on co-designing future SAR platforms would greatly aid the development of new SAR tools aimed specifically at promoting SAR.

Finally, any future work should also focus on the expansion and development of a new SAR platform, utilizing some or all of the recommendations mentioned in Table 47 and Table 48. Replication of the studies presented in this thesis using such a platform would avoid some of the technical issues that frustrated users. This would likely lead to better outcomes with a clearer understanding of the exact impact that SAR is having on the design process. Furthermore, the use of an upgraded SAR platform would allow for the exploration of new use cases for SAR as a tool to support collaborative design. This would aid in promoting the use of SAR in industry.

PUBLICATIONS

Conference:

O'Hare, J.A., Dekoninck, E., **Giunta, L.**, Boujut, J.F. and Becattini, N., 2018, May. Exploring The Performance Of Augmented Reality Technologies In Co-Creative Sessions: Initial Results From Controlled Experiments. In *DS92: Proceedings of the DESIGN 2018 15th International Design Conference* (pp. 405-416).

Dekoninck, E., O'Hare, J., **Giunta, L.**, Masclet, C. and Cascini, G., 2018. Exploring ways to speed up the application of metrics to assess co-creative design sessions. In *DS 89: Proceedings of The Fifth International Conference on Design Creativity (ICDC 2018), University of Bath, Bath, UK* (pp. 69-76).

Giunta, L., Dekoninck, E., Gopsill J., O'Hare J.,2018 'A Review of Augmented Reality Research for Design Practice : Looking to the Future', in *NordDesign 2018*.

Giunta, L., Dekoninck, E. and Gopsill, J. (2020) 'Investigating the Impact of Scale in Design Sessions Supported By a Spatial Augmented Reality (Sar) Tool', *Proceedings of the Design Society: DESIGN Conference*, 1, pp. 131–140. doi: 10.1017/dsd.2020.148.

Giunta, L., Ben Guefrache, F., Dekoninck, E., Gopsill, J., O'Hare, J. and Morosi, F. (2019) 'Investigating the Impact of Spatial Augmented Reality on Communication between Design Session Participants - A Pilot Study', *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), pp. 1973–1982. doi: 10.1017/dsi.2019.203.

Journal:

Cascini, G., O'Hare, J., Dekoninck, E., Becattini, N., Boujut, J. F., Ben Guefrache, F., Carli, I., Caruso, G., **Giunta, L.** and Morosi, F. (2020) 'Exploring the use of AR technology for co-creative product and packaging design', *Computers in Industry*, 123. doi: 10.1016/j.compind.2020.103308.

REFERENCES

Akaoka, E., Ginn, T. and Vertegaal, R. (2010) 'DisplayObjects: Prototyping Functional Physical Interfaces on 3D Styrofoam, Paper or Cardboard Models', in *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction - TEI '10*. New York, New York, USA: ACM Press, p. 49. doi: 10.1145/1709886.1709897.

Akin, Ö. and Lin, C. (1995) 'Design protocol data and novel design decisions', *Design Studies*, 16(2), pp. 211–236. doi: 10.1016/0142-694X(94)00010-B.

Ballestin, G., Chessa, M. and Solari, F. (2019) 'Assessment of optical see-through head mounted display calibration for interactive augmented reality', in *Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019*, pp. 4452–4460. doi: 10.1109/ICCVW.2019.00546.

Bangor, A., Staff, T., Kortum, P., Miller, J. and Staff, T. (2009) 'Determining what individual SUS scores mean: adding an adjective rating scale', *Determining what individual SUS scores mean: adding an adjective rating scale*, 4(3), pp. 114–123.

Beaudouin-Lafon, M. and Mackay, W. E. (2012) 'Prototyping Tools and Techniques', in Jacko, J. A. (ed.) *The Human–Computer Interaction Handbook*. 3rd edn. CRC Press, pp. 1081–1104.

Becattini, N., Cascini, G., O'Hare, J. A. and Masclet, C. (2018) 'Coding schemes for the analysis of ict supported co-creative design sessions', *Proceedings of International Design Conference, DESIGN*, 2, pp. 533–544. doi: 10.21278/idc.2018.0532.

Bellucci, G., Becattini, N., Cascini, G., O'Hare, J., Majoral, X., Boujut, J.-F. and Ben-Guefrache, F. (2018a) *D5.1 VALIDATION AT END USERS' PREMISES*. Available at: http://spark-project.net/sites/default/files/file-wp/D5.1_WP5_Validation at end-users premises.pdf.

Bellucci, G., Niccolo, B., Cascini, G., Hare, J. O., Martens, P. and Morosi, F. (2018b) *D5.4 Show-cases for Increasing the Awareness of SPARK Platform*. Available at: www.spark-project.net/wp-deliverables.

Ben-Guefrache, F., Masclet, C., Prudhomme, G., Cascini, G. and O'Hare, J. A. (2018) 'Real-Time Coding Method For Capture Of Artefact-Centric Iinteractions In Co-Creative Design Sessions', in *International Design Conference - Design 2018*, pp. 33–44. doi: 10.21278/idc.2018.0468.

Ben-Guefreche, F., Boujut, J.-F., Masclet, C., Becattini, N., O'Hare, J., Giunta, L. and Begnoni, T. (2017) *D4.1 DEFINITION OF THE EXPERIMENTAL PROTOCOL FOR A CREATIVE DESIGN PROCESS AND CASE STUDIES*.

Ben-Guefreche, F., Boujut, J.-F., Masclet, C., Poulin, M., Prudhomme, G., Becattini, N., Carbone, N., O'Hare, J., Giunta, L., Dekoninck, E. and Cascini, G. (2018) *RESULTS OF THE EXPERIMENTS BENCHMARKING THE PLATFORM*. Available at: http://spark-project.net/sites/default/files/file-wp/D4.2_WP4_Results_of_the_experiments_benchmarking_the_platform.pdf.

Bergman, M., Lyytinen, K. and Mark, G. (2007) 'Boundary Objects in Design: An Ecological View of Design Artifacts', *Journal of the Association for Information Systems*, 8(11), pp. 546–568. doi: Article.

Billinghurst, M. and Kato, H. (2002) 'Collaborative augmented reality', *Communications of the ACM*, 45(7). doi: 10.1145/514236.514265.

Bimber, O. and Raskar, R. (2005) *Spatial Augmented Reality Merging Real and Virtual Worlds*. Wellesley: A K Peters.

Bimber, O. and Raskar, R. (2006) ‘Modern approaches to augmented reality’, in *ACM SIGGRAPH 2007 courses on - SIGGRAPH ’07*. New York, New York, USA: ACM Press, p. 1. doi: 10.1145/1281500.1281628.

Blessing, L. T. M. and Chakrabarti, A. (2002) ‘DRM, a Design Research Methodology’, in *Proceedings of Les Sciences de la Conception*. London: Springer London. Available at: http://link.springer.com/10.1007/978-1-84882-587-1.

Blessing, L. T. M. and Chakrabarti, A. (2009) *DRM, a Design Research Methodology*, *Focus*. London: Springer London. doi: 10.1007/978-1-84882-587-1.

Blomkvist, J. and Holmlid, S. (2011) ‘Existing Prototyping Perspectives: Considerations for Service Design’, in *Nordic Design Research Conference*, pp. 1–10.

Boa, D. R. and Hicks, B. (2016) ‘Discriminating engineering information interaction using eye tracking and an information operations model’, in *Proceedings of International Design Conference, DESIGN 2016*, pp. 1–10.

Bordegoni, M., Cugini, U., Caruso, G. and Polistina, S. (2009) ‘Mixed prototyping for product assessment: A reference framework’, *International Journal on Interactive Design and Manufacturing*, 3(3), pp. 177–187. doi: 10.1007/s12008-009-0073-9.

Bottani, E. and Vignali, G. (2019) ‘Augmented reality technology in the manufacturing industry: A review of the last decade’, *IISE Transactions*. Taylor & Francis, 51(3), pp. 284–310. doi: 10.1080/24725854.2018.1493244.

Boujut, J. and Blanco, E. (2003) ‘Intermediary Objects as a Means to Foster Co-operation in Engineering Design’, *Computer Supported Cooperative Work (CSCW)*, 12(2), pp. 205–219.

Boujut, J., O’Hare, J. and Becattini, N. (2017) ‘WP4 Test and validation in relevant environment’. SPARK Consortium.

Brooke, J. (1996) ‘SUS: A “Quick and Dirty” Usability Scale’, in *Usability Evaluation In Industry*. CRC Press, pp. 207–212. doi: 10.1201/9781498710411-35.

Bruno, F., Barbieri, L., Marino, E., Muzzupappa, M., D’Oriano, L. and Colacino, B. (2019) ‘An augmented reality tool to detect and annotate design variations in an Industry 4.0 approach’, *International Journal of Advanced Manufacturing Technology*. The International Journal of Advanced Manufacturing Technology, 105(1–4), pp. 875–887. doi: 10.1007/s00170-019-04254-4.

Burke, B. and Smith, D. (2019) *Hype Cycle for Emerging Technologies, 2019*, *Gartner*. Available at: https://www.gartner.com/en/documents/3956015.

Buschfeld, D., Dilger, B., Heß, L. S., Kurt, S. and Voss, E. (2011) *IDENTIFICATION OF FUTURE SKILLS NEEDS IN MICRO AND CRAFT (-TYPE) ENTERPRISES UP TO 2020*.

Butt, J. (2020) ‘A Strategic Roadmap for the Manufacturing Industry to Implement Industry 4 . 0’, pp. 1–30. doi: 10.3390/designs4020011.

Cakmakci, O. and Rolland, J. (2006) ‘Head-Worn Displays: A Review’, *Journal of Display Technology*, 2(3), pp. 199–216. doi: 10.1109/JDT.2006.879846.

Calife, D., Bernardes, J. L. and Tori, R. (2009) ‘Robot Arena: An Augmented Reality Platform For Game Development’, *Computers in Entertainment*, 7(1), p. 1. doi: 10.1145/1486508.1486519.

Calixte, X. and Leclercq, P. (2017) ‘The Interactive Projection Mapping as a Spatial Augmented Reality to Help Collaborative Design: Case Study in Architectural Design', in Luo, Y. (ed.) *International Conference on Cooperative Design, Visualization and Engineering 14th International Conference*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 143–152. doi: 10.1007/978-3-319-66805-5_18.

Camere, S. and Bordegoni, M. (2016) 'A lens on future products: An expanded notion of prototyping practice', in Dorian, M., Mario, S., Neven, P., Nenad, B., and Stanko, S. (eds) *Proceedings of the DESIGN 2016 14th International Design Conference*, pp. 155–164. Available at: https://www.designsociety.org/publication/38825/A+LENS+ON+FUTURE+PRODUCTS%3A+AN+EXPANDED+NOTION+OF+PROTOTYPING+PRACTICE.

Carlile, P. R. (2002) 'A Pragmatic View of Knowledge and Boundaries : Boundary Objects in New Product Development', *Organization Science*, 13(4), pp. 442–455. doi: 10.1287/orsc.13.4.442.2953.

Carrasco, M. D. O. and Chen, P.-H. (2021) 'Application of mixed reality for improving architectural design comprehension effectiveness', *Automation in Construction*, 126(September 2020), p. 103677. doi: 10.1016/j.autcon.2021.103677.

Caruso, G., Carli, I., Noël, F., Trlin, M., Garzia Garza, I. and Dutreve, L. (2016a) *D2.3 SPARK modules prototype*. Available at: www.spark-project.net/wp-deliverables.

Caruso, G., Carulli, M. and Bordegoni, M. (2015) 'Augmented Reality System for the Visualization and Interaction with 3D Digital Models in a Wide Environment', *Computer-Aided Design and Applications*, 12(1), pp. 86–95. doi: 10.1080/16864360.2014.949579.

Caruso, G., Lucia, D. De, Becattini, N., Noël, F., Pusch, A., Martens, P., Garza, I. G. and Dutreve, L. (2016b) *TECHNOLOGIES AND TECHNIQUES – STATE OF THE ART UPDATES*. Milano.

Chai, K. and Xiao, X. (2012) 'Understanding design research: A bibliometric analysis of Design Studies (1996-2010)', *Design Studies*. Elsevier Ltd, 33(1), pp. 24–43. doi: 10.1016/j.destud.2011.06.004.

Chalhoub, J. and Ayer, S. K. (2018) 'Using Mixed Reality for electrical construction design communication', *Automation in Construction*. Elsevier, 86(November 2017), pp. 1–10. doi: 10.1016/j.autcon.2017.10.028.

Cherry, E. and Latulipe, C. (2014) 'Quantifying the creativity support of digital tools through the creativity support index', *ACM Transactions on Computer-Human Interaction*, 21(4). doi: 10.1145/2617588.

Chin, R. (2013) *eDrawings for iOS with Augmented Reality*, *Dassault Systemes*. Available at: http://www.dexigner.com/news/26246 (Accessed: 29 April 2019).

Clough, P. and Nutbrown, C. (2012) *A Student's Guide to Methodology*. 3rd edn. SAGE Publications.

Colruyt Group (2020) *Annual Report with sustainability reporting 2019/20*. Halle. Available at: https://www.colruytgroup.com/wps/wcm/connect/cg/e91deee1-d72d-41c2-868a-7e47895b0243/812801_JV20_ONLINE_EN+DEF.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE.Z18_NI0C1CS0O81620QKM9KAC02H94-e91deee1-d72d-41c2-868a-7e47895b0243-nezZTjc.

Columbus, L. (2019) 'What's New In Gartner's Hype Cycle For AI, 2019', *Forbes*, September. Available at: https://www.forbes.com/sites/louiscolumbus/2019/09/25/whats-new-in-gartners-hype-cycle-for-ai-2019/.

Cortes, G., Marchand, E., Brincin, G. and Lécuyer, A. (2018) 'MoSART: Mobile spatial augmented reality for 3D Interaction with tangible objects', *Frontiers Robotics AI*, 5(AUG), pp. 1–13. doi:

10.3389/frobt.2018.00093.

Cross, N., Christiaans, H. and Dorst, K. (1996) *Analysing design activity*. New York, New York, USA: Wiley.

Dalinger, T., Thomas, K. B., Stansberry, S. and Xiu, Y. (2020) 'A mixed reality simulation offers strategic practice for pre-service teachers', *Computers & Education*. Elsevier, 144(September 2019), p. 103696. doi: 10.1016/j.compedu.2019.103696.

Dasey, D. (2017) *IKEA Place Augmented Reality App*, *IKEA*. Available at: https://highlights.ikea.com/2017/ikea-place/ (Accessed: 5 April 2019).

Dekoninck, E. A., O'Hare, J. A., Giunta, L., Masclet, C. and Cascini, G. (2018) 'Exploring ways to speed up the application of metrics to assess co-creative design sessions', in *Proceedings of the Fifth International Conference on Design Creativity*, pp. 1–8.

Dervojeda, K., Verzijl, D., Nagtegaal, F., Lengton, M., Rouwmaat, E., Monfardini, E. and Frideres, L. (2014) *Design for Innovation: Co-creation design as a new way of value creation*, *Business Innovation Observatory*.

Dey, A., Billinghurst, M., Lindeman, R. W. and Swan, J. E. (2018) 'A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014', *Frontiers in Robotics and AI*, 5(April). doi: 10.3389/frobt.2018.00037.

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M. and Hernandez, N. V. (2015) 'Empirical Studies of Designer Thinking : Past , Present , and Future', *Journal of Mechanical Design*, 137(2). doi: 10.1115/1.4029025.

Eder, M., Spitzer, M., Hebenstreit, M. and Ramsauer, C. (2021) 'Development and Evaluation of a Mixed Reality Assistance System in the Context of Manual Assembly', *SSRN Electronic Journal*. doi: 10.2139/ssrn.3858456.

European Commission (2020) *User guide to the SME Definition*, *Publications Office of the European Union*. Luxembourg. doi: https://doi.org/10.2873/255862.

Fenn, J. and Blosch, M. (2018) *Understanding Gartner's Hype Cycles*. Available at: https://www.gartner.com/en/documents/3887767/understanding-gartner-s-hype-cycles.

Franke, N. and Piller, F. (2004) 'Value Creation by Toolkits for User Innovation and Design: The Case of the Watch Market', *Product Innovation Management*, 21, pp. 401–415.

Gero, J. S. and Mc Neill, T. (1998) 'An approach to the analysis of design protocols', *Design Studies*, 19(1), pp. 21–61. doi: 10.1016/S0142-694X(97)00015-X.

Gerriets (2018) *Special Effect Film PEPPER'S GHOST*. Available at: https://www.gerriets.com/tr/mirrorefect-film-pepper-s-ghost (Accessed: 8 September 2018).

Giunta, L. (2017) *SPARK Assessment Metrics Refining and Applying Metrics for the Testing of the SPARK Platform*. Available at: http://www.bath.ac.uk/library/dissertations/upload.bho/lg413.pdf.

Giunta, L., Dekoninck, E. and Gopsill, J. (2020) 'Investigating the Impact of Scale in Design Sessions Supported By a Spatial Augmented Reality (Sar) Tool', *Proceedings of the Design Society: DESIGN Conference*, 1, pp. 131–140. doi: 10.1017/dsd.2020.148.

Giunta, L., Dekoninck, E., Gopsill, J. and Hare, J. O. (2018) 'A Review of Augmented Reality Research for Design Practice : Looking to the Future', in *NordDesign 2018*. Available at:

https://www.designsociety.org/publication/40967/A+Review+of+Augmented+Reality+Research+for+Design+Practice%3A+Looking+to+the+Future.

Giunta, L., Ben Guefrache, F., Dekoninck, E., Gopsill, J., O'Hare, J. and Morosi, F. (2019) 'Investigating the Impact of Spatial Augmented Reality on Communication between Design Session Participants - A Pilot Study', *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), pp. 1973–1982. doi: 10.1017/dsi.2019.203.

Golafshani, N. (2003) 'Understanding Reliability and Validity in Qualitative Research', *The Qualitative Report*, 8(4), pp. 597–607.

Gopsill, J., Snider, C., Shi, L. and Hicks, B. (2016) 'Computer aided design user interaction as a sensor for monitoring engineers and the engineering design process', in *Proceedings of International Design Conference, DESIGN 2016*, pp. 1707–1718.

Gultekin-atasoy, P., Lu, Y., Bekker, T., Eggen, B. and Brombacher, A. (2014) 'Evaluating Value Design Workshop at Collaborative Design Sessions', in Laakso, M. and Ekman, K. (eds) *DS 81: Proceedings of NordDesign 2014*, pp. 335–345.

Hallgrimsson, B. (2012) *Prototyping and Modelmaking for Product Design*. London: Laurence King Publishin.

Hansen, C. A., Jensen, L. S., Özkil, A. G. and Martins Pacheco, N. M. (2020) 'FOSTERING PROTOTYPING MINDSETS IN NOVICE DESIGNERS WITH THE PROTOTYPING PLANNER', in *Proceedings of the Design Society: DESIGN Conference*, pp. 1725–1734. doi: 10.1017/dsd.2020.132.

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

Hashimoto, S., Ishida, A., Inami, M. and Igarashi, T. (2013) 'TouchMe: An augmented reality interface for remote robot control', *Journal of Robotics and Mechatronics*, 25(3), pp. 529–537. doi: 10.20965/jrm.2013.p0529.

Henrik (2020) *File:Focal-length.svg*. Available at: https://commons.wikimedia.org/wiki/File:Focal-length.svg (Accessed: 27 March 2021).

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004) 'Design Science in Information Systems Research', *MIS Quarterly*, 28(1), p. 75. doi: 10.2307/25148625.

Hua, H., Brown, L. D. and Zhang, R. (2011) 'Head-Mounted Projection Display Technology and Applications', in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 123–155. doi: 10.1007/978-1-4614-0064-6_5.

Irlitti, A. and Itzstein, S. Von (2013) 'Validating constraint driven design techniques in spatial augmented reality', in *AUIC '13 Proceedings of the Fourteenth Australasian User Interface Conference - Volume 139*, pp. 63–72. doi: http://dx.doi.org/10.1016/j.sigpro.2003.07.018.

Israel, J. H., Wiese, E., Mateescu, M., Zöllner, C. and Stark, R. (2009) 'Investigating three-dimensional sketching for early conceptual design—Results from expert discussions and user studies', *Computers & Graphics*, 33(4), pp. 462–473. doi: 10.1016/j.cag.2009.05.005.

Jang, C., Bang, K., Moon, S., Kim, J., Lee, S. and Lee, B. (2017) 'Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina', *ACM Transactions on Graphics*, 36(6), pp. 1–13. doi: 10.1145/3130800.3130889.

Jensen, L. S., Özkil, A. G. and Mortensen, N. H. (2016) ‘Prototypes in engineering design: Definitions and strategies’, in Dorian, M., Mario, S., Neven, P., Nenad, B., and Stanko, S. (eds) *Proceedings of the DESIGN 2016 14th International Design Conference*, pp. 821–830. Available at: https://www.designsociety.org/publication/38892/PROTOTYPES+IN+ENGINEERING+DESIGN%3A+DEFINITIONS+AND+STRATEGIES.

Jensen, M. B., Balters, S. and Steinert, M. (2015) ‘Measuring prototypes-a standardized quantitative description of prototypes and their outcome for data collection and analysis’, in *Proceedings of the International Conference on Engineering Design, ICED*, pp. 1–14.

Kaufmann, H. and Csisinko, M. (2011) ‘Wireless Displays in Educational Augmented Reality Applications’, in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 157–175. doi: 10.1007/978-1-4614-0064-6_6.

Kemmoku, Y. and Komuro, T. (2016) ‘AR Tabletop Interface using a Head-Mounted Projector’, in *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, pp. 288–291. doi: 10.1109/ISMAR-Adjunct.2016.0097.

Kerr, J. and Lawson, G. (2020) ‘Augmented Reality in Design Education: Landscape Architecture Studies as AR Experience’, *International Journal of Art & Design Education*, 39(1), pp. 6–21. doi: 10.1111/jade.12227.

van Krevelen, D. W. F. and Poelman, R. (2010) ‘A Survey of Augmented Reality Technologies, Applications and Limitations’, *The International Journal of Virtual Reality*, 9(2), pp. 1–20.

Lauff, C. A., Weidler-Lewis, J., Kotys-Schwartz, D. and Rentschler, M. E. (2018) ‘Prototypes as intermediary objects for design coordination in first-year design courses’, *International Journal of Engineering Education*, 34(3), pp. 1085–1103.

Lauff, C., Kotys-Schwartz, D. and Rentschler, M. E. (2017) ‘What is a prototype? Emergent roles of prototypes from empirical work in three diverse companies’, *Proceedings of the ASME Design Engineering Technical Conference*, 7, pp. 1–13. doi: 10.1115/DETC201767173.

Lauff, C., Menold, J. and Wood, K. L. (2019) ‘Prototyping Canvas: Design Tool for Planning Purposeful Prototypes’, in *Proceedings of the Design Society: International Conference on Engineering Design*, pp. 1563–1572. doi: 10.1017/dsi.2019.162.

Liestøl, G. (2011) ‘Situated Simulations Between Virtual Reality and Mobile Augmented Reality: Designing a Narrative Space’, in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 309–319. doi: 10.1007/978-1-4614-0064-6_14.

Ligar (2005) *File:DOF-ShallowDepthofField.jpg*. Available at: https://commons.wikimedia.org/wiki/File:DOF-ShallowDepthofField.jpg (Accessed: 27 March 2021).

Linsey, J. S., Clauss, E. F., Kurtoglu, T., Murphy, J. T., Wood, K. L. and Markman, A. B. (2011) ‘An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods’, *Journal of Mechanical Design*, 133(3), p. 031008. doi: 10.1115/1.4003498.

Linsey, J. S., Green, M. G., Murphy, J. T., Wood, K. L. and Markman, A. B. (2005) ‘“Collaborating To Success”: An Experimental Study of Group Idea Generation Techniques’, in *Volume 5a: 17th International Conference on Design Theory and Methodology*. ASME, pp. 277–290. doi: 10.1115/DETC2005-85351.

Löchtefeld, M., Krüger, A. and Rohs, M. (2011) ‘Mobile Projection Interfaces for Augmented Reality

Applications', in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 177–197. doi: 10.1007/978-1-4614-0064-6_7.

López-Mesa, B. and Vidal, R. (2006) 'NOVELTY METRICS IN ENGINEERING DESIGN EXPERIMENTS', in *DS 36: Proceedings DESIGN 2006*, pp. 557–564.

Ma, D., Fan, X., Gausemeier, J. and Grafe, M. (2011) *Virtual Reality & Augmented Reality in Industry*. 1st edn. Edited by D. Ma, X. Fan, J. Gausemeier, and M. Grafe. Berlin: Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-17376-9.

Majoral, X., Becattini, N., O'Hare, J., Bellucci, G. and Martens, P. (2018) *D5.3 DEMONSTRATION WITH OTHER CREATIVE INDUSTRIES AND WITH CUSTOMERS*. Available at: http://spark-project.net/sites/default/files/file-wp/D5.3_WP5_Demonstration_with_other_creative_industries_and_with_customers.pdf.

Marner, M. R., Smith, R. T., Porter, S. R., Broecker, M. M., Close, B. and Thomas, B. H. (2011) 'Large Scale Spatial Augmented Reality for Design and Prototyping', in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 231–254. doi: 10.1007/978-1-4614-0064-6_10.

Marner, M. R. and Thomas, B. H. (2010) 'Augmented foam sculpting for capturing 3D models', in *2010 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, pp. 63–70. doi: 10.1109/3DUI.2010.5444720.

Marner, M. R., Thomas, B. H. and Sandor, C. (2009) 'Physical-virtual tools for spatial augmented reality user interfaces', in *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. IEEE, pp. 205–206. doi: 10.1109/ISMAR.2009.5336458.

Masood, T. and Egger, J. (2020) 'Adopting augmented reality in the age of industrial digitalisation', *Computers in Industry*. Elsevier B.V., 115, p. 103112. doi: 10.1016/j.compind.2019.07.002.

Mathias, D., Hicks, B., Snider, C. and Ranscombe, C. (2018) 'Characterising the affordances and limitations of common prototyping techniques to support the early stages of product development', *Proceedings of International Design Conference, DESIGN*, 3, pp. 1257–1268. doi: 10.21278/idc.2018.0445.

McAlpine, H., Hicks, B. J., Huet, G. and Culley, S. J. (2006) 'An investigation into the use and content of the engineer's logbook', *Design Studies*, 27(4), pp. 481–504. doi: 10.1016/j.destud.2005.12.001.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1995) 'Augmented reality: a class of displays on the reality-virtuality continuum', in Das, H. (ed.) *Telemanipulator and Telepresence Technologies*, pp. 282–292. doi: 10.1117/12.197321.

Mombeshora, M., Dekoninck, E., O'hare, J., Boujut, J.-F. and Cascini, G. (2017) 'Applying multiple metrics in the performance measurement of design sessions in industry: A co-design case study', *21st International Conference on Engineering Design (ICED17)*, 2(DS87-2), pp. 457–466.

Morosi, F., Carli, I., Caruso, G., Cascini, G., Dhokia, V. and Ben-Guefrache, F. (2018a) 'Analysis of Co-Design Scenarios and Activities for the Development of a Spatial-Augmented Reality Design Platform', in *International Design Conference - Design 2018*, pp. 381–392. doi: 10.21278/idc.2018.0504.

Morosi, F., Carli, I., Caruso, G., O'Hare, J. and Ben-Guefrache, F. (2018b) 'D5.2 Validation with Students'. Available at: http://www.spark-project.net/wp-deliverables.

Mugge, R., Schoormans, J. P. L. and Schifferstein, H. N. J. (2009) 'Incorporating consumers in the design of their own products. The dimensions of product personalisation', *CoDesign*, 5(2), pp. 79–97. doi: 10.1080/15710880802666416.

Nelson, B. A., Wilson, J. O., Rosen, D. and Yen, J. (2009) 'Refined metrics for measuring ideation effectiveness', *Design Studies*. Elsevier Ltd, 30(6), pp. 737–743. doi: 10.1016/j.destud.2009.07.002.

O'Hare, J. A., Dekoninck, E., Giunta, L., Boujut, J. and Becattini, N. (2018a) 'Exploring the Performance of Augmented Reality Technologies in Co- Creative Sessions : Initial Results From Controlled Experiments', in *International Design Conference - Design 2018*, pp. 405–416. doi: 10.21278/idc.2018.0391.

O'Hare, J., Dekoninck, E., Mombeshora, M., Martens, P., Becattini, N. and Boujut, J.-F. (2018b) 'Defining requirements for an Augmented Reality system to overcome the challenges of creating and using design representations in co-design sessions', *CoDesign*. Taylor & Francis, 00(00), pp. 1–24. doi: 10.1080/15710882.2018.1546319.

O'Hare, J., Mombeshora, M., Varvatis, C., Bellucci, G., Majoral, X., Martens, P. and Becattini, N. (2016a) *D1.1 Case Studies and Evaluation Criteria*.

O'Hare, J., Mombeshora, M., Varvatis, C., Ben-Guefrache, F., Masclet, C., Prudhomme, G., Martens, P. and Becattini, N. (2016b) *RESULTS FROM THE EXPERIMENTAL ACTIVITIES AND PRESENTATION OF THE RESEARCH METRICS FRAMEWORK*. Available at: http://spark-project.net/wp-deliverables.

OED Online (2020) *research, n.1*, *Oxford University Press*. Available at: https://www.oed.com/view/Entry/163432?rskey=PCFeh3&result=1&isAdvanced=false.

Ong, S. K., Zhang, J., Shen, Y. and Nee, A. Y. C. (2011) 'Augmented Reality in Product Development and Manufacturing', in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 651–669. doi: 10.1007/978-1-4614-0064-6_30.

Oxford University Press (2018a) *in vitro*, *Oxford University Press*. Available at: https://en.oxforddictionaries.com/definition/in_vitro.

Oxford University Press (2018b) *in vivo*, *Oxford University Press*. Available at: https://en.oxforddictionaries.com/definition/in_vivo.

Pahl, G., Beitz, W., Feldhusen, J. and Grote, K. H. (2007) *Engineering Design A Systematic Approach*. 3rd edn. Edited by K. Wallace and L. Blessing. London: Springer. Available at: https://link.springer.com/content/pdf/10.1007%2F978-1-84628-319-2.pdf%0Ahttps://ezproxy.eafit.edu.co:2183/content/pdf/10.1007%2F978-1-84628-319-2.pdf.

Panetta, K. (2017) *Top Trends in the Gartner Hype Cycle for Emerging Technologies, 2017*, *Gartner*. Available at: https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/ (Accessed: 4 July 2018).

Panetta, K. (2019) *5 Trends Appear on the Gartner Hype Cycle for Emerging Technologies, 2019*, *Gartner*. Available at: https://www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019/.

Panetta, K. (2021a) *3 Themes Surface in the 2021 Hype Cycle for Emerging Technologies*, *Gartner*. Available at: https://www.gartner.com/smarterwithgartner/3-themes-surface-in-the-2021-hype-cycle-for-emerging-technologies.

Panetta, K. (2021b) *5 Trends Drive the Gartner Hype Cycle for Emerging Technologies, 2020*, *Gartner*. Available at: https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020.

Papadimitriou, K. and Pellegrin, C. (2007) ‘Dynamics of a project through Intermediary Objects of Design (IODs): A sensemaking perspective’, *International Journal of Project Management*, 25(5), pp. 437–445. doi: 10.1016/j.ijproman.2006.11.002.

Park, H. and Moon, H.-C. (2013) ‘Design evaluation of information appliances using augmented reality-based tangible interaction’, *Computers in Industry*. Elsevier B.V., 64(7), pp. 854–868. doi: 10.1016/j.compind.2013.05.006.

Park, J. (2008) ‘Augmented Reality Based Re-formable Mock-Up for Design Evaluation’, in *2008 International Symposium on Ubiquitous Virtual Reality*. IEEE, pp. 17–20. doi: 10.1109/ISUVR.2008.22.

Park, M. K., Lim, K. J., Seo, M. K., Jung, S. J. and Lee, K. H. (2015) ‘Spatial augmented reality for product appearance design evaluation’, *Journal of Computational Design and Engineering*. Elsevier, 2(1), pp. 38–46. doi: 10.1016/j.jcde.2014.11.004.

Payne, A. F., Storbacka, K. and Frow, P. (2008) ‘Managing the co-creation of value’, *Journal of the Academy of Marketing Science*, 36(1), pp. 83–96. doi: 10.1007/s11747-007-0070-0.

Peddie, J. (2017) *Augmented Reality*. 1st edn. Cham: Springer. doi: https://doi.org/10.1007/978-3-319-54502-8.

Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007) ‘A design science research methodology for information systems research’, *Journal of Management Information Systems*, 24(3), pp. 45–77. doi: 10.2753/MIS0742-1222240302.

Perttula, M. K., Krause, C. M. and Sipilä, P. (2006) ‘Does idea exchange promote productivity in design idea generation ?’, *CoDesign*, 2(3), pp. 125–138. doi: 10.1080/15710880600797942.

Petrakis, K., Hird, A. and Wodehouse, A. (2019) ‘The Concept of Purposeful Prototyping: Towards a New Kind of Taxonomic Classification’, in *Proceedings of the Design Society: International Conference on Engineering Design*, pp. 1643–1652. doi: 10.1017/dsi.2019.170.

Poh, Y. L., Nee, A. Y. C., Youcef-Toumi, K. and Ong, S. K. (2005) ‘Facilitating Mechanical Design with Augmented Reality’, *Innovation in Manufacturing Systems and Technology*, pp. 1–5. Available at: https://dspace.mit.edu/handle/1721.1/7448%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Facilitating+Mechanical+Design+with+Augmented+Reality#0.

Porter, S. R., Marner, M. R., Smith, R. T., Zucco, J. E. and Thomas, B. H. (2010) ‘Validating spatial augmented reality for interactive rapid prototyping’, in *2010 IEEE International Symposium on Mixed and Augmented Reality*. IEEE, pp. 265–266. doi: 10.1109/ISMAR.2010.5643599.

Potts, D., Loveys, K., Ha, H., Huang, S., Billinghurst, M. and Broadbent, E. (2019) ‘ZenG: AR Neurofeedback For Meditative Mixed Reality’, in *Proceedings of the 2019 on Creativity and Cognition*. New York, NY, USA: ACM, pp. 583–590. doi: 10.1145/3325480.3326584.

Prahalad, C. K. and Ramaswamy, V. (2004) ‘Co-creation experiences: The next practice in value creation’, *Journal of Interactive Marketing*. Elsevier, 18(3), pp. 5–14. doi: 10.1002/dir.20015.

Probst, L., Frideres, L., Demetri, D., Moujahid, S., Vomhof, B. and Lonkeu, O.-K. (2014) *Customer Experience Customer incentives and involvement*.

Radkowski, R. and Linnemann, M. (2009) ‘Applicability of Image-Based Lighting for an Augmeted Reality-Based Design Review’, in Norell Bergendahl, M. ., Grimheden, M. ., Leifer, L. ., Skogstad, P. ., and Lindemann, U. (eds) *ICED09: 17th International Conference on Engineering Design*, pp. 253–264. Available at: https://www.designsociety.org/publication/28697/applicability_of_image-based_lighting_for_an_augmented_reality-based_design_review.

Ben Rajeb, S. and Leclercq, P. (2013) ‘Using Spatial Augmented Reality in Synchronous Collaborative Design’, in Luo, Y. (ed.) *Cooperative Design, Visualization, and Engineering 10th International Conference*. Springer Berlin Heidelberg, pp. 1–10. doi: 10.1007/978-3-642-40840-3_1.

Rauschnabel, P. A., Brem, A. and Ivens, B. S. (2015) ‘Who will buy smart glasses? Empirical results of two pre-market-entry studies on the role of personality in individual awareness and intended adoption of Google Glass wearables’, *Computers in Human Behavior*. Elsevier Ltd, 49, pp. 635–647. doi: 10.1016/j.chb.2015.03.003.

Reyes, A. C. C., Del Gallego, N. P. A. and Deja, J. A. P. (2020) ‘Mixed Reality Guidance System for Motherboard Assembly Using Tangible Augmented Reality’, in *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*. New York, NY, USA: ACM, pp. 1–6. doi: 10.1145/3385378.3385379.

Robson, C. and McCartan, K. (2016) *Real World Research*. 4th edn. Chichester: John Wiley & Sons Ltd.

Sanders, E. B.-N. (2006) ‘Design Research in 2006’, *Design Research Quarterly*, 1(1), pp. 1–8.

Sanders, E. B.-N. and Stappers, P. J. (2008) ‘Co-creation and the new landscapes of design’, *CoDesign*, 4(1), pp. 5–18. doi: 10.1080/15710880701875068.

Savov, V. (2017) ‘Google Glass gets a second chance in factories, where it’s likely to remain’, *The Verge*, 18 July. Available at: https://www.theverge.com/2017/7/18/15988258/google-glass-2-enterprise-edition-factories.

Scrivener, S. A. R. (2005) ‘Editorial’, *CoDesign*, 1(1), pp. 1–4. doi: 10.1080/15710880412331289935.

Scrivener, S. A. R., Ball, L. J. and Andrée, W. (2000) *Collaborative Design: Proceedings of CoDesigning 2000*, *Journal of Chemical Information and Modeling*.

Shah, J. J., Kulkarni, S. V and Vargas-Hernandez, N. (2000) ‘Evaluation of Idea Generation Methods for Conceptual Design : Effectiveness Metrics and Design’, *Journal of Mechanical Design*, 122(4), pp. 377–384. doi: 10.1115/1.1315592.

Shah, J. J., Vargas-Hernandez, N. and Smith, S. M. (2003) ‘Metrics for measuring ideation effectiveness’, *Design Studies*, 24(2), pp. 111–134. doi: 10.1016/S0142-694X(02)00034-0.

Shah, J. J., Vargas-Hernandez, N., Summers, J. D. and Kulkarni, S. (2001) ‘Collaborative Sketching (C-Sketch) – An Idea Generation Technique for Engineering Design’, *Journal of Creative Behavior*, 35(3), pp. 168–198. doi: https://doi.org/10.1002/j.2162-6057.2001.tb01045.x.

Smparounis, K., Mavrikios, D., Pappas, M. and Xanthakis, V. (2007) ‘A virtual and augmented reality approach to collaborative product design and demonstration’.

Snider, C. M., Dekoninck, E. A. and Culley, S. J. (2012) ‘IMPROVING CONFIDENCE IN SMALLER DATA SETS THROUGH METHODOLOGY : THE DEVELOPMENT OF A CODING SCHEME’, in *DS 70: Proceedings of DESIGN 2012*, pp. 1253–1264.

Sola, C. (2016) *Spatial Augmented Reality for Design Representations in Product Design and Development*. University of Bath.

SPARK Consortium (2015) *SPARK Literature Review*.

SPARK Consortium (2017) *SPARK platform mixed prototype visualization: Alce Nero tomato sauce jar.* Available at: https://www.youtube.com/watch?v=_BcM1P9c8MM.

SPARK Consortium (2018a) *Interested in how Augmented Reality can revolutionize your creative processes?*, *Instagram*. Available at: https://www.instagram.com/p/BhOKfhunf9I/?taken-by=sparkh2020.

SPARK Consortium (2018b) ‘Plug and play Augmented Reality platform?’ Kortrijk. Available at: https://twitter.com/SparkH2020/status/1060455546480943105/photo/1.

SPARK Consortium (2018c) ‘Ready to discover more about #minispark ?’ Available at: https://twitter.com/SparkH2020/status/1069897231250153473/photo/1.

SPARK Consortium (2018d) ‘SAVE THE DATE!!’ Barcelona. Available at: https://twitter.com/cangurman/status/1003217821692809217/photo/1.

SPARK Consortium (2019a) ‘Augmenting the reality’. Milan. Available at: https://www.instagram.com/p/BtLMPWxHsih/.

SPARK Consortium (2019b) ‘Inside Spark room’. Milan. Available at: https://www.instagram.com/p/ByZ04t7is8P/.

Srivathsavai, R., Genco, N., Hölttä-Otto, K. and Seepersad, C. C. (2010) ‘Study of Existing Metrics Used in Measurement of Ideation Effectiveness’, in *ASME Proceedings | 22nd International Conference on Design Theory and Methodology*, pp. 355–366.

Star, S. L. and Griesemer, J. R. (1989) ‘Institutional Ecology , ’ Translations ’ and Boundary Objects : Amateurs and Professionals in Berkeley ’ s Museum of Vertebrate Zoology , 1907-39’, *Social Studies of Science*, 19(3), pp. 387–420. doi: 10.2307/285080.

Statista (2018) *Share of virtual reality and augmented reality (VR and AR) users in the United States as of 2018, by type of device.* Available at: https://www.statista.com/statistics/830508/us-virtual-augmented-reality-users-by-device/.

Stutzman, B., Nilsen, D., Broderick, T. and Neubert, J. (2009) ‘MARTI: Mobile Augmented Reality Tool for Industry’, in *2009 WRI World Congress on Computer Science and Information Engineering*. IEEE, pp. 425–429. doi: 10.1109/CSIE.2009.930.

Syberfeldt, A., Danielsson, O. and Gustavsson, P. (2017) ‘Augmented Reality Smart Glasses in the Smart Factory: Product Evaluation Guidelines and Review of Available Products’, *IEEE Access*. IEEE, 5, pp. 9118–9130. doi: 10.1109/ACCESS.2017.2703952.

Szebeko, D. and Tan, L. (2010) ‘Co-designing for society’, *Australasian Medical Journal*, 3(9), pp. 580–590. doi: 10.4066/AMJ.2010.378.

Tawara, T. (2011) ‘Interactive Volume Segmentation and Visualization in Augmented Reality’, in *Handbook of Augmented Reality*. New York, NY: Springer New York, pp. 199–210. doi: 10.1007/978-1-4614-0064-6_8.

Ullman, D. G. (2003) *The mechanical design process*. 3rd edn. New York: McGraw-Hill.

Ulrich, P. V., Jo Anderson-Connell, L. and Wu, W. (2003) ‘Consumer co-design of apparel for mass customization’, *Journal of Fashion Marketing and Management: An International Journal*, 7(4), pp. 398–412. doi: 10.1108/13612020310496985.

Uva, A. E., Gattullo, M., Manghisi, V. M., Spagnulo, D., Cascella, G. L. and Fiorentino, M. (2018) ‘Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations’, *International Journal of Advanced Manufacturing Technology*. The International Journal of Advanced Manufacturing Technology, 94(1–4), pp. 509–521. doi: 10.1007/s00170-017-0846-4.

Valentini, P. P. (2009) ‘Interactive virtual assembling in augmented reality’, *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 3(2), pp. 109–119. doi: 10.1007/s12008-009-0064-x.

Valentini, P. P. and Biancolini, M. E. (2018) ‘Interactive Sculpting Using Augmented-Reality, Mesh Morphing, and Force Feedback: Force-Feedback Capabilities in an Augmented Reality Environment’, *IEEE Consumer Electronics Magazine*. IEEE, 7(2), pp. 83–90. doi: 10.1109/MCE.2017.2709598.

Valtonen, A. (2005) ‘Six decades – and six different roles for the industrial designer .’, in *Nordes Conference, In the Making, 30-31st May*, pp. 1–10.

Verlinden, J., van Duijnen, F. and Horvath, I. (2010) ‘Validating the strengths and weaknesses of Interactive Augmented Prototyping in Industry’, in *Proceedings of IDMME*.

Vinck, D. and Jeantet, A. (1994) ‘Mediating and comissioning objects in the sociotechnical process of product design: a conceptual approach’, in *Management and new technology: design, networks and strategies. Proceedings from COST A3 workshop*, pp. 111–129.

Wai, Y. J. and Abd Manap, N. B. (2018) ‘Interactive Objects for Augmented Reality by Using Oculus Rift and Motion Sensor’, *Journal of Telecommunication, Electronic and Computer Engineering*, 10(2–6), pp. 149–153.

Wavecontrol (2016) *About Wavecontrol EMF Test Equipment*. Available at: https://www.wavecontrol.com/rfsafety/en/company/about-wavecontrol (Accessed: 23 November 2020).

de Wit, G. C. (1999) ‘Retinal scanning display: light sources moving over the retina.’, *Science progress*, 82, pp. 135–149. doi: 10.1177/003685049908200203.

Xin, M., Sharlin, E. and Sousa, M. C. (2008) ‘Napkin sketch’, in *Proceedings of the 2008 ACM symposium on Virtual reality software and technology - VRST ’08*. New York, New York, USA: ACM Press, p. 223. doi: 10.1145/1450579.1450627.

Yamaguchi, M., Mori, S., Mohr, P., Tatzgern, M., Stanescu, A., Saito, H. and Kalkofen, D. (2020) ‘Video-Annotated Augmented Reality Assembly Tutorials’, in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. New York, NY, USA: ACM, pp. 1010–1022. doi: 10.1145/3379337.3415819.

Zhang, J., Ong, S. K. and Nee, A. Y. C. (2010) ‘Development of an AR system achieving in situ machining simulation on a 3-axis CNC machine’, *Computer Animation and Virtual Worlds*, 21(2), pp. 103–115. doi: 10.1002/cav.327.

Zhen, L., Jing, C., Zixiang, Z., Qiushuo, T. and Ningsheng, H. (2017) ‘An Optical See-Through Augmented Reality System with Gesture-Based Interaction’, in *Proceedings - 2016 International Conference on Virtual Reality and Visualization, ICVRV 2016*. IEEE, pp. 447–452. doi: 10.1109/ICVRV.2016.82.

Zobele Group (2014) *Product Categories*. Available at: https://www.zobele.com/categories (Accessed: 23 November 2020).

APPENDIX

A. QUESTIONNAIRE

Copy of the interview questions used in the version 4 metrics. First discussed in section 2.6.8.

Pre-session Interview

Imagine if a new designer had just joined your project team and that you wanted them to participate in this session. What important information about the project would they have to know to be able to contribute effectively?
[TASK CHART] **Are there any action points or design tasks from previous meetings and discussions that have not yet been addressed?** **If so could you list and mark them as High, Medium or Low importance?**
What are the objectives of this meeting?
On a scale of 1 to 10, 1 being little and 10 being a lot how important will it be for you to filter ideas during this session?
On a scale of 1 to 10, 1 being little and 10 being a lot how important will it be for you to generate new ideas during this session?
How many concepts or ideas would you like to end the session with (in order to be ready to move to the next stage of the project)?

Post-session Interview

[Screenshots] These are the screenshots you took during the session. Are there any ideas or concepts are missing? If there are, could you please describe them?
(OPTIONAL) During the session you spent some time discussing [Description of IDEA] but you did not take a screenshot of it. Why did you not take a screenshot of it? Should it be added here?

B. Forms

Copy of the charts and sheets used in the version 4 metrics to track idea generation and session progress. First discussed in section 2.6.8.

B.1. Morphological Chart

		Idea Elements					
		Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
Function or Purpose							

B.2. Idea Sheet

Idea Code	Idea To Be Taken Forward		Novelty Rating (From 1 to 10)
	No	Yes	
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			

22			
23			

B.3. Task Rating Sheet

Task		Status			Participant Ranking		
		Still Open	Created	Closed	Low	Med	High
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							

C. INSPIRATION BOARDS

Copy of the inspiration board used in Impact of SAR on Communication between Design Session Participants. First discussed in section 6.3.2.

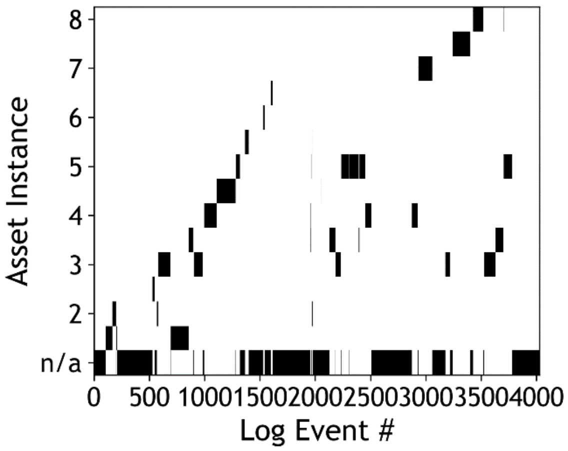
C.1. Scale up

C.2. Scale Down

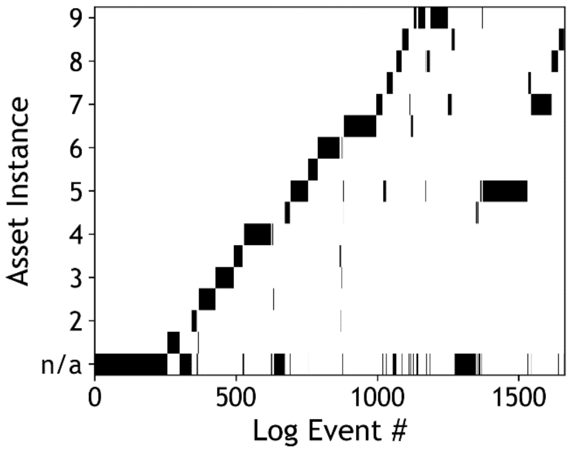
D. ASSET INSTANCE VS LOG EVENT GRAPHS

Full set of interaction behaviour graphs, showing waterfall effect, for all conditions tested. First discussed in section 6.3.6.

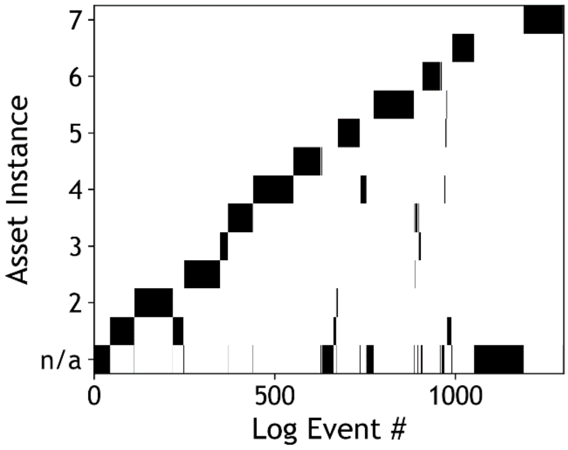
D.1. 1:1 Scale



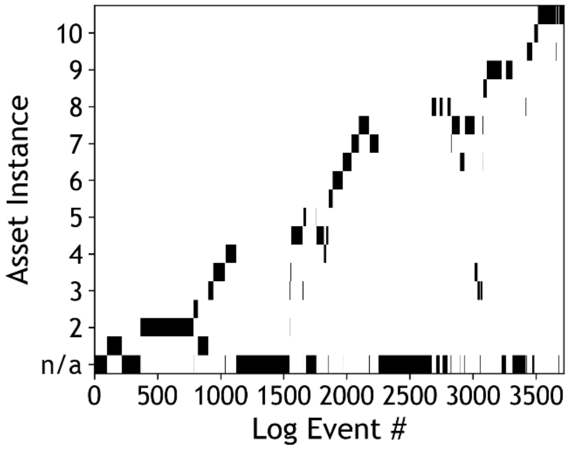
Session 1



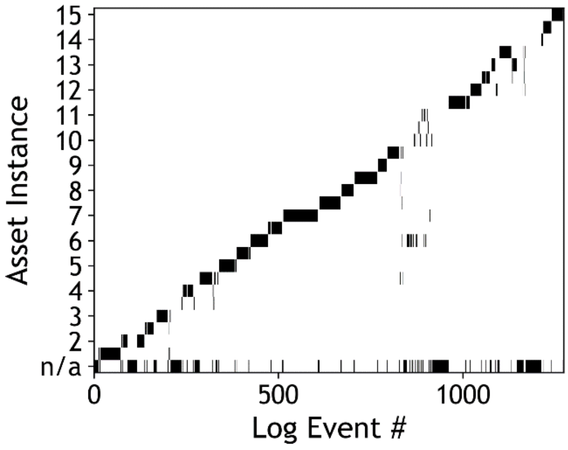
Session 2



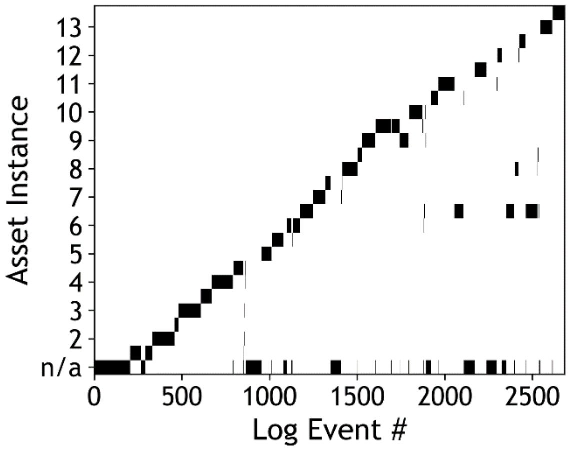
Session 3



Session 4

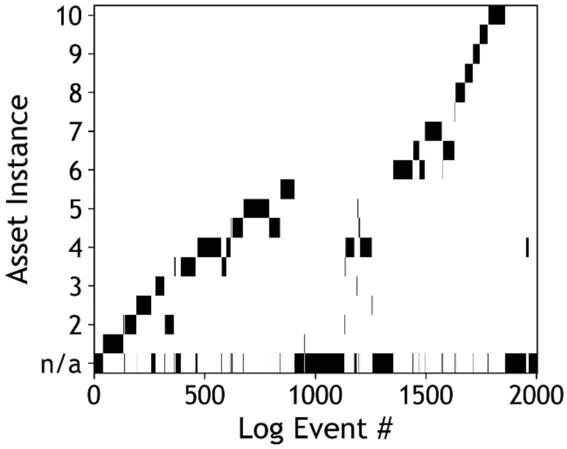


Session 5

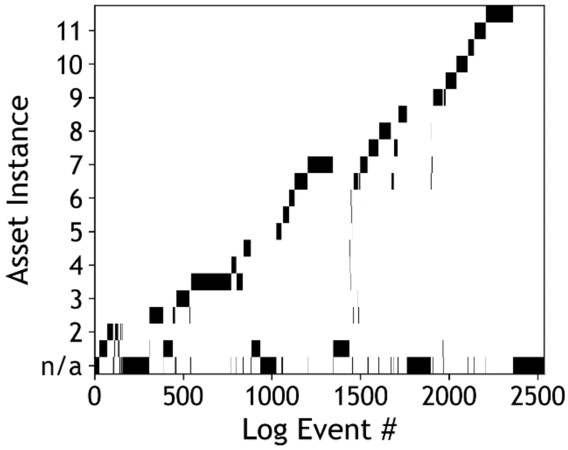


Session 6

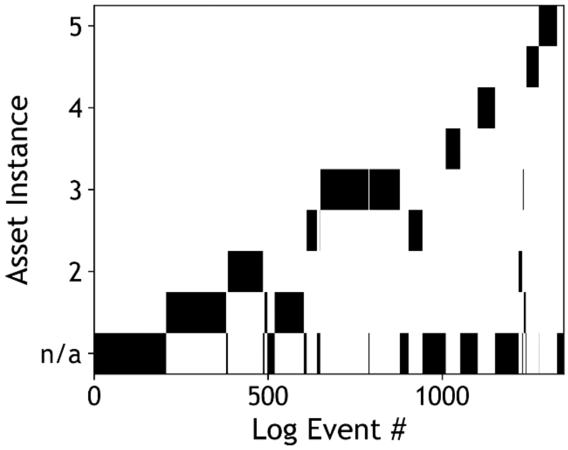
D.2. 2:1 Scale



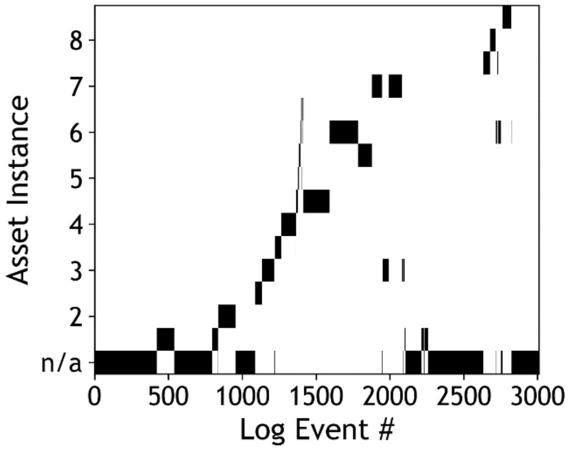
Session 1



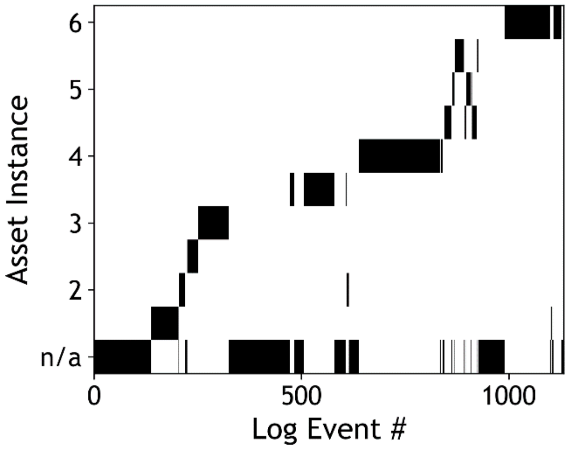
Session 2



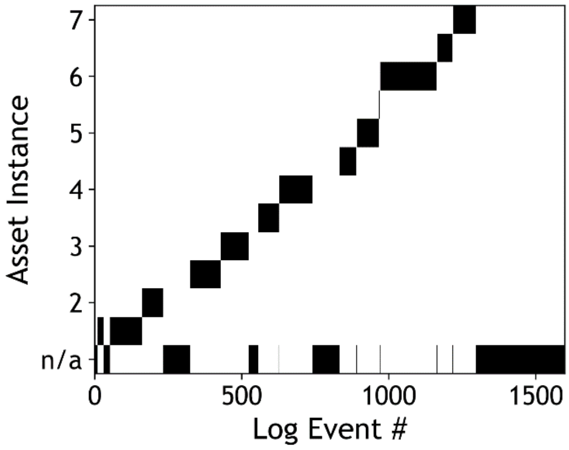
Session 3



Session 4

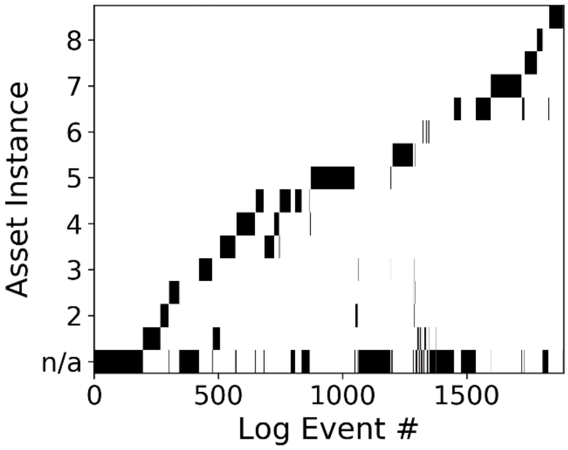


Session 5

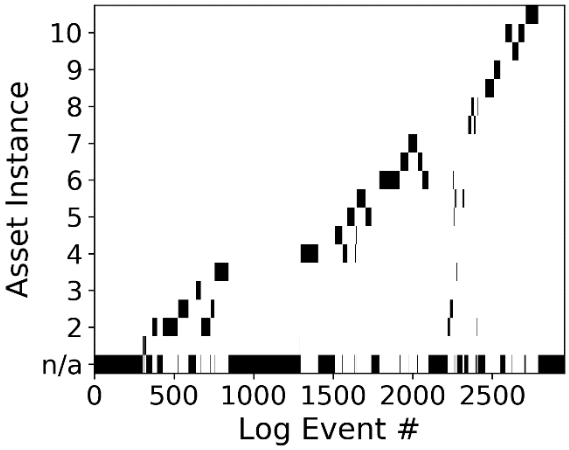


Session 6

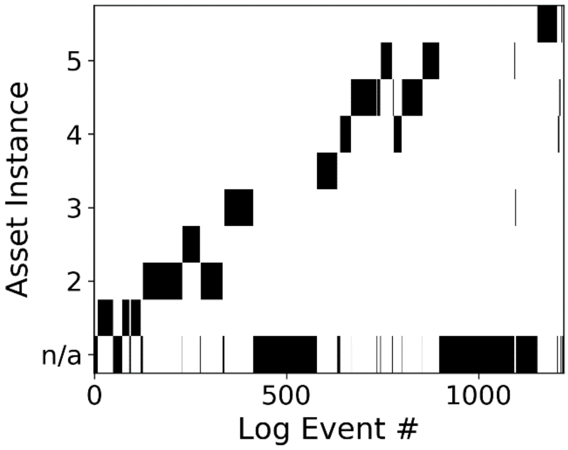
D.3. 3:1 Scale



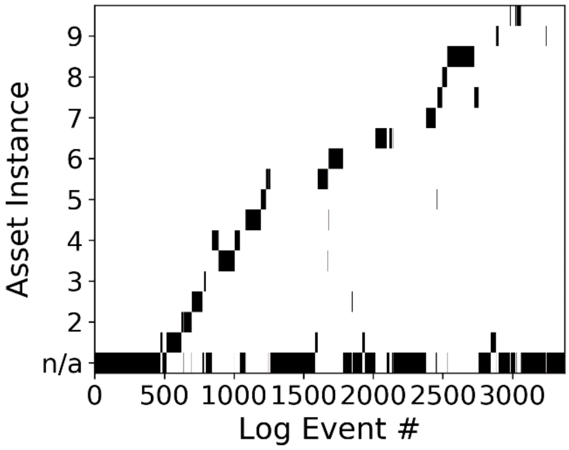
Session 1



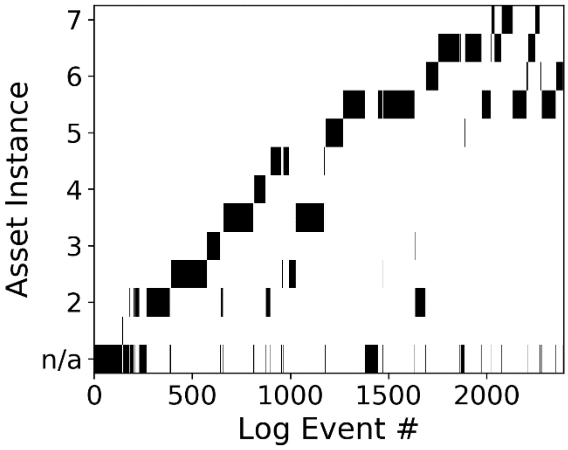
Session 2



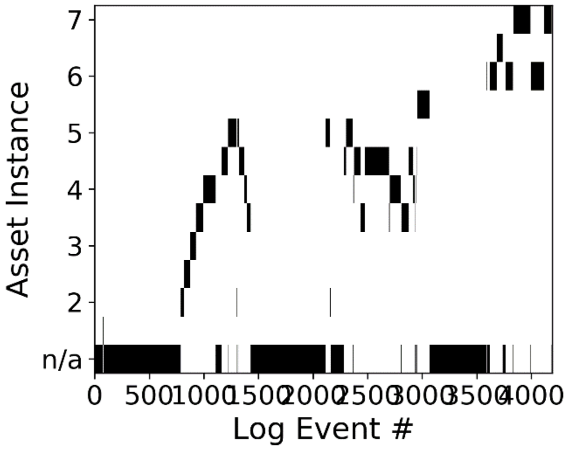
Session 3



Session 4

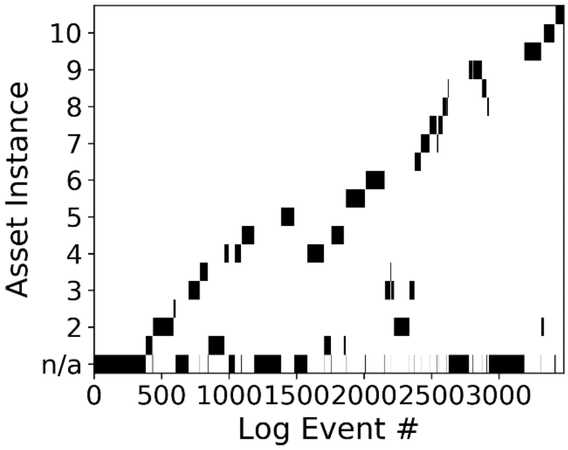


Session 5

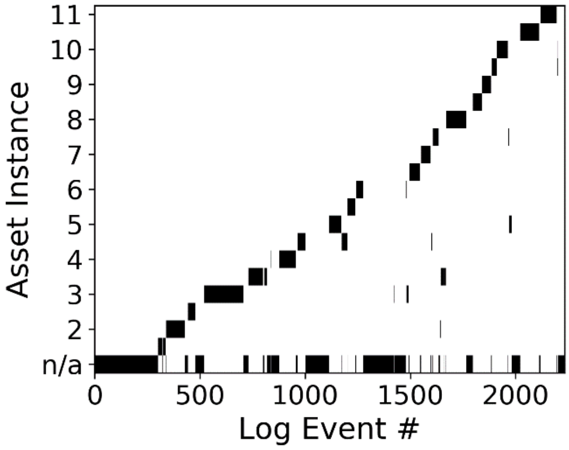


Session 6

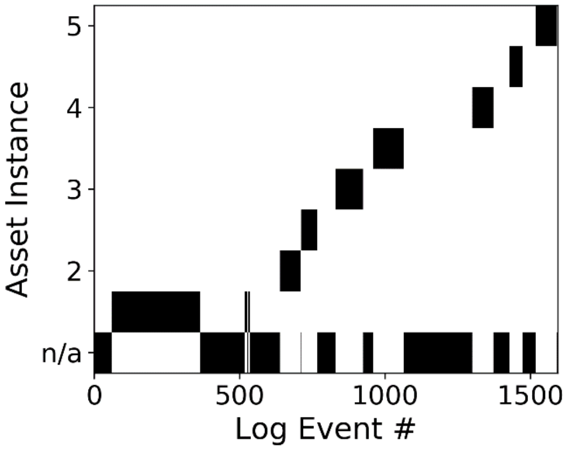
D.4. 1:10 Scale



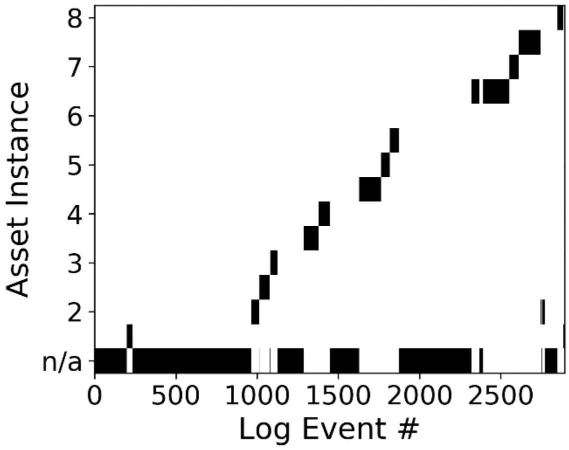
Session 1



Session 2

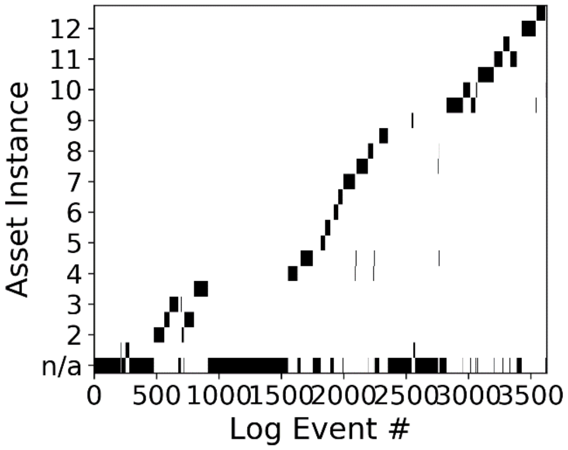


Session 3

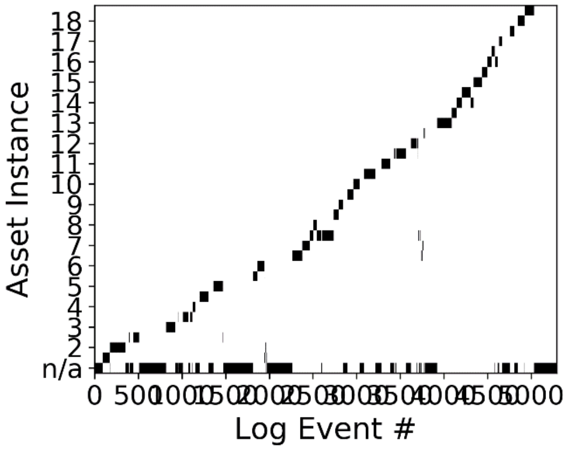


Session 4

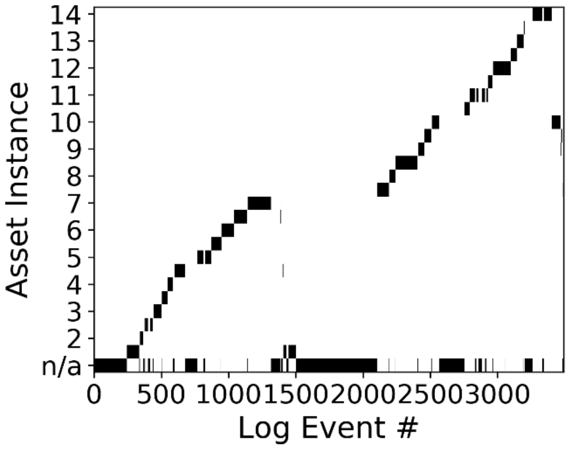
D.5. 1:20 Scale



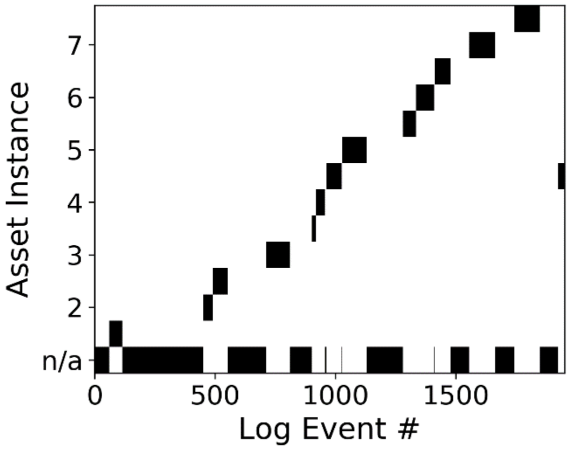
Session 1



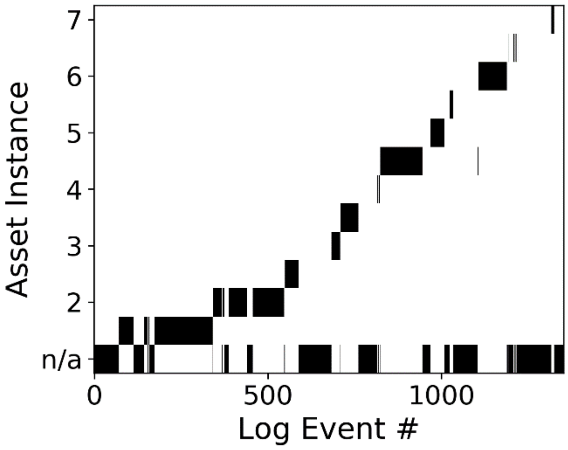
Session 2



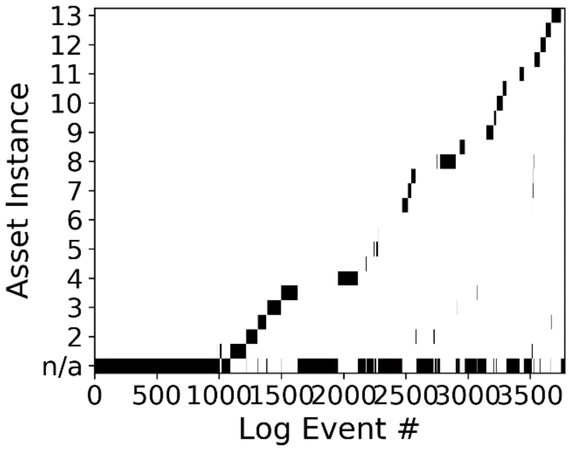
Session 3



Session 4



Session 5



Session 6

E. CREATIVITY SUPPORT INDEX QUESTIONNAIRE SHEET

Copy of the questions used in the CSI survey questionnaire. First discussed in section 6.1.2.2.

Name: **ICT technology used:**

Please rate your agreement with the following statements:

I was satisfied with what I got out of the system or tool.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

It was easy for me to explore many different ideas, options, designs or outcomes, using this system or tool.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

The system or tool allowed other people to work with me easily.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

I would be happy to use this system or tool on a regular basis.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

I was able to be very creative while doing the activity inside the system or tool.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

My attention was fully tuned to the activity, and I forgot about the system or tool that I was using.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

I enjoyed using this system or tool.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

The system or tool was helpful in allowing me to track different ideas, outcomes or possibilities.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

What I was able to produce was worth the effort I had to exert to produce it.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

The system or tool allowed me to be very expressive.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

It was really easy to share ideas and designs with other people inside this system or tool.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

I became so absorbed in the activity that I forgot about the system or tool that I was using.

Highly Disagree ☐ ☐ ☐ ☐ ☐ ¦ ☐ ☐ ☐ ☐ ☐ **Highly Agree**

When doing this task, it’s most important that I’m able to…

Explore many different ideas, outcomes, or possibilities	☐	☐	Work with other people
Be creative and expressive	☐	☐	Produce results that are worth the effort I put in
Enjoy using the system or tool	☐	☐	Become immersed in the activity
Become immersed in the activity	☐	☐	Produce results that are worth the effort I put in
Work with other people	☐	☐	Enjoy using the system or tool
Produce results that are worth the effort I put in	☐	☐	Explore many different ideas, outcomes, or possibilities
Be creative and expressive	☐	☐	Become immersed in the activity
Work with other people	☐	☐	Produce results that are worth the effort I put in
Be creative and expressive	☐	☐	Enjoy using the system or tool
Explore many different ideas, outcomes, or possibilities	☐	☐	Become immersed in the activity
Work with other people	☐	☐	Be creative and expressive
Produce results that are worth the effort I put in	☐	☐	Enjoy using the system or tool
Explore many different ideas, outcomes, or possibilities	☐	☐	Be creative and expressive
Work with other people	☐	☐	Become immersed in the activity
Explore many different ideas, outcomes, or possibilities	☐	☐	Enjoy using the system or tool

F. SYSTEM USABILITY SCALE

Copy of the questions used in the SUS questionnaire. First discussed in section 6.1.2.2.

	Strongly Disagree				Strongly Agree
1. I think that I would like to use this system frequently	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
2. I found the system unnecessarily complex	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
3. I thought the system was easy to use	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
4. I think that I would need the support of a technical person to be able to use this system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
5. I found the various functions in this system were well integrated	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
6. I thought there was too much inconsistency in the system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
7. I would imagine that most people would learn to use this system very quickly	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
8. I found the system very cumbersome to use	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
9. I felt confident using the system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
10. I needed to learn a lot of things before I could get going with this system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

G. PRE-PREPARED STATEMENT FOR PRIMING IN SCALE EXPERIMENTS

Copy of the pre-prepared statements used in Impact of SAR on Communication between Design Session Participants. First discussed in section 6.3.2.

G.1. Scale Up

Imagine the following: you work for a design agency and have been contacted by a company to design a new watch. You have been provided with some pictures of the types of watches your watch is expected to compete with.

Your task is this: using the SAR platform you must design a number of options for your client company to start thinking about which kind of design they want to pursue. Take as much time as you need to work and develop as many ideas as you think are necessary. Just keep in mind: whenever you have a concept you like and would like to capture it for your client please ring the bell next to you. You can generate however many concepts as you want: just keep in mind your client will have to review all of them.

Keep in mind that the client will use the SAR platform to review your designs: this will mean that they will not see the interface at all, so if you have any doubts or notice any differences between the SAR platform and the interface you should go with what you see on the SAR platform.

G.2. Scale Down

Every year Team Bath Racing redesigns its car by setting up a competition. You have decided that you want to join this competition together and already have a list of the previous winning designs from past years. To top all of them with an even better one, you are going to design the best possible livery for TBR.

Your task is this: using the SAR platform you must design a number of options for TBR to start thinking about. Take as much time as you need to work and develop as many ideas as you think are necessary. Just keep in mind: whenever you have a concept you like and would like to capture it for TBR please ring the bell next to you. You can generate however many concepts as you want: However your client will have to review all of them.

Lastly remember this: TBR will use the SAR platform to review your designs: this will mean that they will not see the interface at all, so if you have any doubts or notice any differences between the SAR platform and the interface you should go with what you see on the SAR platform.

H. TRADE FAIR SURVEY

Copy of the questionnaire used in Industry Feedback from Trade Fairs. First discussed in section 7.1.1.

1. How often does your company currently hold co-creative design sessions (or product development review meetings) with internal stakeholders, customers or end users?

One or more times per week [ ]
Around once per month [ ]
Around once per quarter [ ]
Around once per year [ ]
Never [ ]

2. How important are the following challenges for your organisation?

	Not at all important				Very important
Overcoming barriers to communication with stakeholders	[ ]	[ ]	[ ]	[ ]	[ ]
Reducing the time to market	[ ]	[ ]	[ ]	[ ]	[ ]
Reducing unnecessary iteration in the design process	[ ]	[ ]	[ ]	[ ]	[ ]
Reducing the cost of creating prototypes	[ ]	[ ]	[ ]	[ ]	[ ]
Generating novel ideas	[ ]	[ ]	[ ]	[ ]	[ ]
Obtaining actionable feedback from stakeholders	[ ]	[ ]	[ ]	[ ]	[ ]

3. What type of organisation do you work in?

Other [ ]
Home furnishings sector [ ]
Consumer products sector [ ]
Packaging sector [ ]
Fashion & jewellery sector [ ]
Manufacturer - Other sector [ ]
Food and beverage sector [ ]
Design agency/consultancy [ ]

4. Approximately how many people are employed in your organisation?

Less than 10 people [ ]
11-50 people [ ]
51-250 people [ ]
More than 250 people [ ]

I. SEMI-STRUCTURED INTERVIEW SURVEY

Copy of the questionnaire used in SAR Platform Pilots with Industry Members. First discussed in section 7.2.1.3.

1. How often does your company currently hold co-creative design sessions (or product development review meetings) with internal stakeholders, customers or end users?

One or more times per week	☐
Around once per month	☐
Around once per quarter	☐
Around once per year	☐
Never	☐

2. How important are the following challenges for your organisation?

	Not at all important				Very important
Overcoming barriers to communication with stakeholders	☐	☐	☐	☐	☐
Reducing the time to market	☐	☐	☐	☐	☐
Reducing unnecessary iteration in the design process	☐	☐	☐	☐	☐
Reducing the cost of creating prototypes	☐	☐	☐	☐	☐
Generating novel ideas	☐	☐	☐	☐	☐
Obtaining actionable feedback from stakeholders	☐	☐	☐	☐	☐
Other – please specify below	☐	☐	☐	☐	☐

Other challenge: ____

3. Based on what you have seen today, how would you rate the following features of the SPARK system?

	Very poor				Excellent
Resolution of rendering	☐	☐	☐	☐	☐
Accuracy of colour rendering	☐	☐	☐	☐	☐
Accuracy of projection alignment	☐	☐	☐	☐	☐
Latency (responds quickly to movements of the model)	☐	☐	☐	☐	☐
Ease of use (tablet user interface)	☐	☐	☐	☐	☐

4. To what extent do you agree or disagree with the following statements.

Using SPARK for co-creative design sessions at my company would…	Strongly disagree				Strongly agree
… improve idea generation	☐	☐	☐	☐	☐
… improve idea review and filtering	☐	☐	☐	☐	☐
…reduce labour costs in the overall design process	☐	☐	☐	☐	☐
… reduce prototyping costs in the overall design process	☐	☐	☐	☐	☐
…reduce time to market in the overall design process	☐	☐	☐	☐	☐

…be easy to implement	[ ]	[ ]	[ ]	[ ]	[ ]
…overall, be beneficial for the company	[ ]	[ ]	[ ]	[ ]	[ ]

5. In your own words, what are the things that you like most about this new technology?

6. In your own words, what are the things that you would most like to improve in this new technology?

7. Any other comments?

Name:

Company:

Position:

J. LETTER OF CONFIRMATION OF CONTRIBUTION

POLITECNICO
MILANO 1863

Milano, 04 November 2020

To Whom It May Concern

DIPARTIMENTO DI
MECCANICA
Prof. Gaetano Cascini

Contribution confirmation

I would hereby like to confirm Lorenzo Giunta's research contributions to the SPARK Project from until 02 Oct 2017 until 31 Dec 2018.

Although Lorenzo did not appear as the main contributing author for Bath in the deliverables, I can confirm that he:

- designed a number of the data collection methods;
- designed and built the experimental set-up at Bath;
- and conducted data collection and analysis.

This work resulted in research contributions from Lorenzo to work package 4 (deliverables D4.1 and D4.2) and work package 5 (deliverables D5.1, D5.2, D5.3, and D5.4) as detailed in the table below.

Yours sincerely,

Prof. Gaetano Cascini

WP	Contribution	Description
WP4 - D4.1	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.
	Paper writing and review	Created graphs and relevant explanations, created tables and relevant explanations
WP4 - D4.2	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.
	Paper writing and review	Created graphs and relevant explanations, created tables and relevant explanations
WP5 - D5.1	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.
WP5 - D5.2	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.
WP5 - D5.3	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.
WP5 - D5.4	Data Collection	Performed interviews with study participants, administered surveys, recorded participant behaviour during sessions
	Data Analysis	Evaluated data from the surveys administered, collated data collected from live sessions recordings and sanitized this data, analysed interviews to gather qualitative and quantitative information.