OXFORD BROOKES UNIVERSITY

DOCTORAL THESIS

Title of the Doctoral Thesis

Performance Augmentation: Immersive Technology for Workplace Training

Submitted by

William A. A. Guest

in fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Oxford Brookes University, 2023

Degree programme code as it appears on the student record sheet:	PHD-TD
Degree programme as it appears on the student record sheet:	Doctorate in Computer Science
Director of Studies:	Dr. Tjeerd Olde Scheper
Postgraduate Tutor:	Dr. Matthias Rolf
2nd Supervisor:	Prof. Dr. Fridolin Wild

Contents

Acknowledgements

The research supporting this thesis would not have been accomplished without the considerable support of a large number of mentors, collaborators, and friends. I am foremost thankful to my family for showing me the value of communication—an appreciation that has only deepened as a result of this undertaking.

My director of studies, Dr. Tjeerd Olde Scheper, has shown unwavering confidence and support, and I remain indebted to him for many fascinating conversations as well as the compassion and patience he has shown me over the years. My initial director of studies and both a mentor and collaborator, Prof. Dr. Fridolin Wild, has been an enthusiastic source of inspiration and knowledge. In my sometimes-meandering path, he provided consistency and friendly critique at every turn, allowing me to become the researcher I am today. I am grateful for the many opportunities that he offered me, even after his departure to the Open University. I am also thankful to Dr. Matthias Rolf, my postgraduate tutor, whose keen eye and rigorous insight has made this work stronger.

The research was funded by the European Commission, under the Horizon 2020 programme, as part of WEKIT (grant agreement no. 678669), TCBL (646133), and ARETE (856533). During COVID-19, the author received financial aid from Oxford Brookes University, for which I am very thankful. This research was also made possible with funding from The Open University through the OpenReal project, which has built on the software described here and open-sourced it as MirageXR[1].

I cannot speak highly enough of the many collaborators with whom I have been fortunate to work. I am thankful to Alla Kitov for helping me to appreciate the importance of beauty in immersive design. I am very grateful to Prof. Dr. Roland Klemke, Dr. Bibeg Limbu, Dr. Jan Schneider, and Dr. Daniele Di Mitri at the Open University of the Netherlands, as well as Jaakko Karjalainen, Kaj Helin, and Timo Kuula of VTT, and Carl Smith, Dr. Mark Ransley, and Dr. Puneet Sharma, who supported the creation of the experience capture software and built hardware prototypes, allowing for the capture of necessary data. Their expertise was only matched by their generosity; they showed me how the whole can be greater than its parts.

Lastly, I am deeply appreciative of my friends, who have accompanied me through the ups and downs of this period of my life. It is because of you that I strive to make things better and make better things.

[1]Found at: https://github.com/WEKIT-ECS/MIRAGE-XR/

Abstract

This doctoral thesis investigates the use of immersive technology for workplace training with the purpose of improving competence building and providing a framework for future development. Although rooted in computer science, it reaches out to other disciplines, including technology-enhanced learning, pedagogy, and statistical psychology, in order to inform and evaluate the development of an immersive training system.

Built around a core of software engineering principles and practices, this research's theoretical landscape includes Affordance Theory and Social Systems Theory. It takes structure from models of technology acceptance (UTAUT2) and Anderson and Krathwohl's taxonomy of cognitive processes, using these to shed light on the attitudes and expectations that surround the use of this technology.

An 'experience capture system' is described, which allows an expert to compose sequences of augmented instructions while carrying out the activity. As well as a head-mounted display, the system also used body-worn sensors to capture movement and muscle activation, connecting data streams that could be recorded, stored, and edited through the immersive interface. The inclusion of the recorded data streams connected the trainee to a learning dimension not accessible with flat-screen technology.

This system was tested across medical, aeronautic, and astronautic use cases with more than 400 people over two iterations, each of whom also participated in a technology acceptance study investigating their attitudes towards augmented reality and wearable technology. Analysis was done using structural equation modelling, where constructs and relations were investigated with confirmatory factor analysis and path analysis, respectively. Model optimisation was conducted using a combination of theoretical review and statistical metrics, including modification indices and all common absolute and relative fit indices.

In the technology acceptance studies, closer interoperability was consistently found to predict a reduction in the expected effort required to use the system. The attitudinal constructs of individual- and activity-technology fit, developed as part of this research, were found to play important roles in predicting the acceptance of immersive technology. Unexpectedly, a closer perceived fit between the technology and the activity was found to accompany a sense of being ill-equipped for the task.

As well as methodological contributions on acceptance model optimisation and interpretation, this work also puts forward the affordance as a unit of measurement for technology-enhanced learning and describes three categories of immersive learning affordances: propositional, embodied, and meta-constructive. These are presented alongside cognitive processes as an 'affordance dimension', with the goal of operationalising them in the context of immersive technology development.

In summary, through this thesis, I have demonstrated a theoretical outline for categorising affordances, examples of their use with immersive technology, and a quantitative evaluation protocol. Despite constraints from device capabilities, I have shown that immersive technology is enjoyable, easy to use, and effective in supporting learning in the workplace.

Chapter 1: Introduction

This thesis charts the construction of a set of tools that support workplace training using immersive technology. In doing so, it seeks to understand the experience of the trainee when using the technology and attempts to provide a coherent picture of the impacts and influences that the technology brings to bear on this experience. It looks at the relationships between the learning objectives, the experience of the trainee, and a workplace augmented with digital content. Drawing on research from a variety of disciplines, this work strives to give a functional understanding of how these relationships can be structured, explicated, and measured.

Although rooted in computer science and centred around the construction and use of an 'Experience Capture System' (ECS), this research extends into the fields of wearable computing, the production and integration of body-worn devices, and statistical psychology as it investigates the perspectives of those using the immersive training framework. Each area produced data and insights of a different kind. Through the use of extensible software patterns and connections to wearable sensors, complex descriptions of activities could be constructed that contain not only the instructions but also the physical actions of the trainee. Building theoretical models of user acceptance and evaluating the system across two iterations provided data-driven conclusions about the technology's success as well as a rigorous method for evaluating similar implementations.

In the conceptualisation, development, and assessment of the tools presented here, two theories were of particular importance: Social Systems Theory and Affordance Theory. The first offers an interpretation of the trainee as consisting of three functionally distinct, though connected, systems: biological, cognitive, and communicative. The second, working on a more fine-grained level, provides practical support through the definition and operationalisation of the notion of an affordance: an opportunity for action involving the person, their environment, and any information connecting the two. Both theories are relational in nature, taking action in context and understanding that the information around us is also intimately connected to our own knowledge and needs.

Included under the banner of 'immersive technology' are two groups of devices: one is augmented reality (AR) head-mounted displays (HMD) that permit the overlay of digital media onto the physical world. The other is wearable technology, capable of either sensing or directing movement and providing information on our physiological state. In practical terms, this work investigates how these tools, when coupled with experts' knowledge of a task, can facilitate competence-building. It takes an in-depth look into how propositional knowledge and embodied competence can be systematically acquired, codified, and communicated to a trainee. It offers a conceptual and technical basis for testing and development, as well as a structured approach to data collection and analysis. Both the construction of models and the interpretations of trends extracted from the data follow theoretically defined patterns, and the recommendations for iterative change as well as new research are similarly founded.

1.1 Key Concepts

Before pressing on with the details of the research, it will be useful to provide clarification of the meaning and use of the key themes that are under investigation. Given that we are addressing the use of immersive technology for workplace training, three concepts are of central importance: competence (of both trainer and trainee), immersive technology (in general), and augmented reality (as a means of conveying information). In addition, the concept of affordances will be described as it resonates strongly, both in theory and practice, with the use of immersive technology for training.

Competence in Practice

Our notion of competence is closely connected to the outcomes of our efforts; we generally assume that greater competence will allow us to work more efficiently or produce consistently better results. The use of technology, in general, resonates with this concept, and we improve the deftness of its use with repeated exposure, focused training, or associated knowledge. In the case of immersive training, we are thus obliged to consider a person's competence in using the technology as well as in performing the task at hand. Frameworks that attempt to offer taxonomies of competences often reach considerable size and complexity [19][12], illustrating that many competences are highly transferable and, especially when it comes to technology, hard to categorise. A long-standing tension exists in understanding how people apply experience from one area to another, or *how* they enact these competences [98].

One popular idea is that we possess a degree of implicit knowledge behind what is explicitly communicated. Polanyi's work [88] centred around the idea that such knowledge is not communicable through propositional means but rather acquired through experience. More recently, Schmidt has discussed the idea of tacit knowledge at length, in particular looking at how to approach didactic practice [107, p.52] where he concludes that "to make progress on this front is to investigate the actual didactic practices as part of cooperative work practices and how they are integrated with those cooperative work practices.".

Later, Schmidt tackled the idea more broadly, approaching the field of computer-supported cooperative work (CSCW) [106], which takes a multidisciplinary approach to understanding how digital technologies can support practice rather than viewing them simply as purveyors of semantic information. Studies such as this are relevant here because the experience of immersive technology use is more akin to a collaboration than the specification of a set of explicit tasks.

Kuutti and Bannon, writing in 2014, go further, pointing to a 'turn to practice' within human-computer interaction (HCI) research, going as far as to coin a paradigm based on the notion of practice-enabling digital tools, which they place alongside the more established 'Interaction paradigm'. They make the distinction between the two in the following way: "For the Interaction paradigm, the scope of the intervention is viewed as changing human actions by means of novel technology. For the Practice paradigm, a whole practice is the unit of intervention; not only technology, but everything related and interwoven in the performance is under scrutiny and potentially changeable, depending on the goals of the intervention"[63, p.3544]. This highlights both the pluralistic nature of the evolving activity as well as the goal-oriented posture taken when considering pedagogical intervention.

The activities described in this work largely consist of embodied actions that, through practice, are expected to build competence. This topic was discussed in a paper by de Cavalho et al. [22], who draw on many of the previous sources and identify the potential for the use of augmented reality in knowledge-intensive environments. They also specifically cite an absence of research into "how embodied action captured through AR systems can be used for knowledge and expertise sharing" [22, p.10]. This research broadens the scope to include wearable sensors to measure that action, but it is closely aligned with its goal of addressing this gap.

The connection between practice and design is described succinctly by Ludwig et al.: "Instead of simply considering the role of design intervention as changing human actions by introducing novel technology, it needs to be understood that human actions and interactions are just a part of entire practices. Practices emphasize the fabric of action, the knowledge and reasoning that surround that action, and the context in which it takes place" [70, p.4]. Here, the importance of a proper context for training is seen as equally as important as the content of the training, making immersive technology well-suited to the activity of developing meaningful practice.

Immersive Technology and Augmented Reality

Witmer and Singer [136, p.227] define immersion as "a psychological state characterised by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences". Thus, immersion requires both a thing to be immersed as well as a medium in which it is contained. This relational perspective is closely aligned with the constructivist action of meaning-making (as the result of interaction) as well as the notion of affordances, described below.

The term 'Augmented Reality' (AR) is used here to refer to those tools that allow the addition of virtual content to our world in real time and in three dimensions [5]. They are distinct from entirely digitised environments (Virtual Reality) or efforts in miniaturisation that project information onto a near-eye display (such as Google Glass), but do not align this information to the physical structure of space around it.

Technology that allows us to perturb both the content of our experience as well as its context is considered immersive. This distinction therefore includes AR systems, which create a world-locked effect in the creation of digital content. This is accomplished with the combined use of several complex subsystems, including infra-red scanning, depth mapping, device localisation, miniaturised light engines, and optical wave-guides.

'Immersive Technology' also covers wearable devices that are able to provide feedback, either directly or via other, connected devices. These "wearables" can collect data about various forms of biophysical activity, but only become immersive in their function when they offer information to the wearer that modifies their experience, either by providing haptic feedback or visualising data with a head-mounted display (HMD) [14].

In categorising types of display, both Milgram and Benford have offered useful dimensions with which to think about the functionality that they contain. Milgram et al.'s reality-virtuality continuum [78] made a clear distinction between VR and AR glasses. Benford et al., in adding two other dimensions, transportation and spatiality, encouraged us to consider the value of remote presence (at one end of the scale) and also how the spatial frame is constructed with regard to each person using it [8].

Affordances

The term 'affordance' was coined by James Gibson to describe the opportunities for action that the physical environment presents to animals. It is from this ecological standpoint that Gibson discussed our interaction with our world, pointing to the communicative nature of learning. "An affordance", he said, "points two ways, to the environment and to the observer. So does the information to specify an affordance." [36, p.141]. This dual action does not, however, indicate duality. Rather, he states, it "is wholly inconsistent with dualism in any form, either mind-matter dualism or mind-body dualism" (ibid.).

The concept of affordance was further developed by Norman in his book 'The Psychology of Everyday Things' [84]. Here, he also applies the concept to user interaction within software. In later work, he also highlighted the importance of an affordance being perceived, pointing to the difference in usage and accessibility of both graphics and text [83]. More recent work by Volkoff and Strong (2017) identified six principles to guide the use of affordances in information science research [127]. The first two are the reinforcement of the affordance as a user/artefact relationship rather than belonging to one or the other, and the distinction between affordance and actualisation, reminding us that an affordance represents an opportunity rather than a goal. The next two talk about the formulation of the affordances, suggesting action-focused descriptions (e.g., "connecting", "communicating") and a need for an appropriate level of granularity. The last two are system-level, encouraging the discovery of interacting affordances or those that are social in nature, such as collective goals.

When using immersive technology, the term 'affordance' is often used in a more general sense to refer to the benefits that the technology confers on a learner, without necessarily taking into consideration the principles mentioned above [116] [79] [138]. One thing that they agree on, however, is the importance of understanding how the technology can connect people to activities in ways that produce better outcomes, both in terms of the objective metrics around task completion as well as the more subjective notion of task performance.

A Luhmannian Lens

The final element in need of introduction is the constructivist lens through which the theoretical concepts are interpreted and the ground from which new constructs emerge, namely Niklas Luhmann's Social Systems Theory [71]. This theory is important in the context of immersive technology because it offers a way to conceptualise the inclusion of such technology in a learning environment. In a pluralistic turn, it considers the construction of knowledge as the result of several interrelated systems, each considering the others as part of its environment. From a broad, meta-systematic stance, this theory is used to connect the various strands of research (see Chapter 7). By looking at how this technological system interacts with both the trainee and the activity, we aim to draw out the operational connections that make each system distinctive and, in this case, beneficial for training.

1.2 Research Outline

Research Question

> How can we improve workplace training and boost trainee competence using immersive technology?

Research Objectives

In tackling the research question, this thesis also aims to provide solutions to, or at least suggestions for, three previously posed conundra. Each of these also demonstrates a key research objective.

The first objective relates to technical architecture, originally posed by Ludwig et al., when summarising their work on the 'Internet of Practices': "How to capture related resonance activities across communities? Furthermore, if this can be done, how might one approach designing technological support for them?" [70, p.13]. This is the primary objective associated with the construction of an immersive authoring tool, based on information from experts. Put another way, the goal is to build software that is sufficiently context-agnostic that it can be said to have a general purpose but is detailed enough that expertise is properly communicated, such that it improves the performance of the trainee.

On the topic of physical practice in the context of HCI, this research is aligned with a goal raised by Suchman in 1987: how to "explicate the relationship between structures of action and the resources and constraints afforded by physical and social circumstances." [117, p.179]. A similar framing is also made by de Carvalho et al., who suggest that "the missing piece is a focus on the possibility of transmitting non-propositional knowledge through observations of embodied action." [22, p.8]. This work aims to tackle this need using wearable sensors, integrating their data streams into the system, and allowing for the visualisation of wearable data-linked augmentations in order to improve a trainee's performance.

The third objective is concerned with the validation and evaluation of the results obtained from technology acceptance studies. Instead of an outward-looking stance, we are interested in the nuance and descriptive power of the findings, together with quantitative metrics that determine the level of confidence in a particular avenue of development. This goal is equivalent to asking, "How can we specify and use immersive technology acceptance models to evaluate our work and provide insight for iterative development?"

1.3 Originality and Rigour

This thesis probes some of the core elements of augmented reality and wearable technology systems in an effort to improve our understanding of how they can be used in a systematic way to enhance training outcomes.

As a research document, it combines principles of software development and structural equation modelling (SEM) for statistical metrics of attitude and behaviour and blends these together using social science and social system theories. Although each chapter is of a different theme, they each contribute a perspective towards the main research question: how to leverage these technologies in such a way that the trainee reaches their learning objectives with additional speed, competence, or associated benefit?

In each of these fields, this work puts forward novel features. The software, developed to act as an experience recorder for trainers, demonstrated a context-agnostic system that could be used to design AR training procedures. The more advanced features of this system, including the integration of wearable devices, are unique and contribute to the production of an ARLEM data model. The ECS described here was one of the first capable of constructing such an activity model entirely from immersive user interaction.

This work demonstrates novelty in the construction of an immersive technology acceptance model, which was used through two iterations within three industrial use cases, ultimately referencing a dataset of over 400 entries. During this process, new attitudinal constructs were developed and tested as part of a technology-specific extension of the popular UTAUT2 acceptance framework. This model was successful in showing robust correlations, allowing for confident conclusions, as well as highlighting unexpected correlations that prompt new questions and hypotheses.

Both the software engineering and data analysis were part of a wider research effort called Wearable Experience for Knowledge Intensive Training (WEKIT), which investigated the use of AR in industry. Several of the associated avenues of work and related publications are described in more detail in Section 2.3.

The research conduct as well as data collection and storage procedures were subject to oversight and scrutiny both by individual academic partners and, through project reviews, the European Commission. Industry partners were also instrumental in ensuring that data collection instruments were used in an even-handed way and that the participation of the trainees was voluntary and fully informed. More details on this engagement can be found in the relevant sections in Chapter 6.

The trials followed a pre-defined routine that became yet more standardised as it was conducted with more people. Data collection was carried out under strict guidelines, maintaining clear communication with and anonymity for the participants. The project received ethical approval in both countries where the use cases were carried out—Norway and Italy—and there was oversight from Oxford Brookes University. The research received ethical approval from the University Research Ethics Committee of Oxford Brookes University[2].

With regards to statistical or analytical rigour, several steps were taken to mitigate potential bias in the dataset. This included giving attention to the framing of questions and the strict treatment of collected data with regards to anonymity and its use in models. In the analytical phase, the potential for overfitting was dealt with using high-fold Monte Carlo cross-validation, and the presentation of the statistical findings was held to tight semantic constraints.

1.4 Contribution to Knowledge

In this thesis, three strands of investigation have been woven together to provide insight into how immersive technology works and, in each case, provide a novel demonstration of what can be accomplished using it. The three lines of inquiry are programmatic design, wearable data capture, and an analysis of technology acceptance.

Software Engineering

The first contribution comes from work done in the creation of an AR experience capture system designed to be used in knowledge-intensive tasks across a variety of industry environments. The purpose of the system was to allow experts to annotate the performance of a task with digital media, connecting it to a location in space as well as within the task sequence. Other data streams, such as those from body-worn sensors, could also be used to annotate the work flow.

Wearable Data Capture

The second area that this thesis tackles is that of the collection and use of data from wearable sensors connected to the AR system. Data streams from biometric sensors as well as motion capture devices are described and shown in the context of an experience authoring tool to improve performance. The work described here aims to inform researchers looking at how physical movement is connected to competence and how it can be leveraged to provide useful feedback to the trainee.

[2] UREC Reference No. 171156

Immersive Technology Acceptance

The third thread involves the use of attitudinal models to better understand what motivates people to use immersive technology. This thesis develops a technology acceptance model through two iterations of testing, using statistical analysis to constrain and optimise a theoretical framework that supports the development of new technology. Not only are there general recommendations produced by the model, but the process of hypothesis generation and testing using this approach is also expected to be beneficial to other researchers in the field.

1.5 Structure of the Thesis

Chapter 2 presents work in the areas of immersive learning system design as well as the state of the art in AR devices, as well as the software systems that offer authoring functionality. There is also an overview of recent technology acceptance studies and their application in immersive learning environments.

Chapter 3 methodology, consisting of the use of instructional design models, the study of upper-body physiology in support of body-sensor networks, a collection of analytical tools used to evaluate the system, and work as an inference engine linking statistical perturbation to the production of testable hypotheses.

Chapter 4 focuses on the ability to capture the expertise of trainers. Given the intimate connection to the environment as well as the activity being recorded, the functions that the system performs are made to be agnostic to either of these features. The system also produces standardised models of both activity and workplace, allowing the transference of the system's instructional information to other devices or workplaces through the use of the Augmented Reality Learning Experience Model (ARLEM).

Moving from propositional knowledge to practical know-how, Chapter 5 describes the construction of additional features that extend the system's ability to observe physical movement, muscle activation, and other physiological states and convey information about these data to the trainee.

Chapter 6 describes the evaluation of the system based on technology acceptance and structural equation modelling. Statistical analysis informs the specification and optimisation of these models and, over two project iterations, demonstrates how these findings are used to generate new hypotheses.

Chapter 7 provides a summary of the insights gathered over the entire period of work, framing some of the results and indications using the theories mentioned above. It looks at the practical implications of the framework for the expert, the user, and the developer.

Chapter 8 contains some final thoughts, returning to the initial research questions to look at where progress was made and identifying areas for future work.

Chapter 2: Literature Review

The opening chapter identified the key themes and ambitions that frame this work, centred around the goal of enhancing immersive training scenarios. This chapter looks at the research and development landscape of immersive technology for workplace training, beginning with an overview of recent trends in head-mounted displays—the core technology that allows AR to exist. To address the applied use of immersive technology, this is followed by a systematic review of AR authoring software. In the final section, the role of peripherals in the immersive computing paradigm is also summarised, in particular looking at body sensor networks and motion capture tools.

2.1 A Brief History of AR Glasses

The practical development of AR HMDs began with Heilig's Telesphere Mask [46], patented in 1957. Work by Ivan Sutherland and others at Harvard University led to a prototype, built in 1968, of another HMD inspired by a system used by helicopter pilots in which the wearer's movement was coupled to the orientation of a mounted camera. There was a resurgence of interest in HMDs in the 1990s, led by work at Boeing [20], where Tom Caudell coined the term 'Augmented Reality' (AR), and at Armstrong laboratories, under the auspices of the United States Air Force, with the work of Louis Rosenberg.

In the first decade of the 21st century, AR HMDs moved gradually from research to industry, and the first commercial units were made available. Monocular projector-based systems were a theme amongst the first commercial products, and companies such as Vuzix, Google, and Epson led the field, with later iterations adding binocular display and, in some cases, gesture control. Earlier models tended to use OLED displays, with the exception of Google Glass, which instead employed a liquid crystal on silicon (LCoS) near-eye display.

Many of the devices that have come on the market have failed to find traction, either due to compromises made when making them affordable or technical limitations impacting the user experience. Several early devices are summarised by Syberfeldt et al. [118] and, though the devices they cover are almost all now discontinued, their paper clearly highlights the trade-off between form factor and usability. Notably, of the twelve devices in that paper, none featured simultaneous localisation and mapping (SLAM). Of those companies that produce binocular displays, only Epson is still producing HMDs.

Despite the disruptive potential and high hopes for immersive technology, the adoption of new tools in existing markets has been fairly slow. This is generally attributed to two factors: the complexity of the devices (and their resultant cost) and the lack of a standardised framework in which to design and deploy new software applications that successfully leverage the hardware's capabilities.

In using AR systems, many established patterns now require re-thinking, for example, user interface design or input methods, and striking new areas of work have emerged, such as hand gesture recognition and AR-enabled telepresence, each with its own nuance and hurdles.

Billinghurst, Clark, and Lee's survey in 2015, at the same time as providing definitions for many AR concepts now in regular use, clearly outlined the hurdles that should be overcome to make the digital content immersive [9] and pointed to a prototype from Maimone and Fuchs [75] that was a forerunner to the waveguide display that now features in newer AR HMDs.

It was with Microsoft's release of the HoloLens in 2016 that AR headsets reached a commercial level of readiness, setting a benchmark for user experience, hardware performance, precision, and build quality. This device's tracking accuracy, display quality, and interoperability made it highly adaptable. With approximately centimetre-scale precision when anchoring augmentations to the spatial map and high frame-rate updates (more than 25 per second), this device's performance and improved affordability (around 3,500 USD at the time of release) were major milestones in the coming-of-age of AR head-mounted displays. The Meta 2, while sporting a large field of view, relied on combiner optics that made it extremely bulky and less bright than other systems.

Released	**Company**	**Device**	**Display Type**	**SLAM?**
Apr. 2016	Microsoft	HoloLens	Waveguide	Yes
May 2016	Meta	Meta 2	LCD	Yes
Jan. 2017	DAQRI	Smart Helmet	Waveguide	Yes
Apr. 2017	ODG	R-7	OLED	No
Aug. 2018	Magic Leap	Magic Leap 1	Waveguide	Yes
Oct. 2018	Third Eye Gen	X2	Waveguide	Yes
Mar. 2019	Shadow Creator	Action One	Waveguide	Yes
Nov. 2019	Microsoft	HoloLens 2	Waveguide	Yes
Feb. 2020	Vuzix	M400	OLED	No
Jul. 2020	Rokid	Glass 2	Waveguide	Yes
Jun. 2021	Epson	BT-40	OLED	No

Table 1: Recent AR HMDs

An article in the Journal of Nanophotonics by Microsoft engineers Kress and Chatterjee [61] is an in-depth resource for those wishing to understand the technological hurdles and practical considerations of waveguide design and production.

2.2 Authoring Content in Augmented Reality

Given the AR context described in the previous chapter, an "authoring system" or "authoring tool" is a piece of software that is used to craft AR experiences by adding, arranging, and configuring digital content so that it offers a coherent experience to a consumer of the designed experience. This multimodal content must be delivered in such a way that it supports an educational process rather than providing a particular piece or set of information. To be included in this data set, the system described should be considered to have three features or characteristics.

They should first include some form of content selection by the user. An experience in which existing objects are only modified (in colour, size, position, etc.), or where all connections to virtual content are predetermined, are not authoring tools; rather, they are AR applications that feature some user interaction. Second, they should expose lower-level functionality through their operation, simplifying the control or orchestration of components or subsystems, reducing the overall complexity, and lowering the barriers to entry for non-specialists. This simplification allows for a greater focus on designing the learning sequence, with the goal of more faithfully representing the learning objectives. Specifically, the systems should be usable by non-programmers (i.e. the people who constructed them). Third, the system should provide some distinction between the mode in which the objects are assembled or structured and the one in which they are then used (i.e. an expert/trainer mode and a viewer/trainee/learner mode).

2.2.1 Bloom's Revised Taxonomy

Bloom's Revised Taxonomy [4] is a pedagogical tool that describes learning outcomes in a pragmatic, "hands-on" style, emphasising the importance of semantic knowledge, practice and reflection. The revised taxonomy (RT) addresses aspects of learning, teaching, and assessment, categorising learning activities across two dimensions: the knowledge dimension and the cognitive process dimension.

With regards to the learning process, there is a hierarchical structure where actions are spread across six levels. These levels are generally sequential in terms of expertise-building but do not necessarily require preceding actions to be recognised in their own right. This will become evident later, when the taxonomy is applied.

Knowledge Dimension (excluding metacognitive knowledge)

Here we are concerned with the type of knowledge that the AR authoring tool provides, or those elements in the system that have the user's attention. Four knowledge types are identified: factual, conceptual, procedural, and metacognitive. Each of these is described in Table 2, which is a summary of that found in the RT. The continuum that these categories represent is one ranging from concrete (factual) to abstract (metacognitive). The two other types, conceptual and procedural knowledge, overlap with one another, though procedural knowledge is generally considered to be the more abstract of the two. Factual knowledge is characterised by its specificity and forms the basis of expertise in a particular field. Specific facts, together with symbols, notation, or jargon, are examples of factual knowledge that aid those familiar with a subject to accurately and efficiently communicate their ideas. Similarly, these aid in the comprehension of new ideas for those learning the subject and constitute the basis for the second knowledge type. Conceptual knowledge refers to the connections between facts, symbols, or concepts within a larger, functional framework.

In this work, metacognitive knowledge will not be studied for two reasons. First, because the field of augmented reality is still young with respect to information technology. As was the case in the earlier years of desktop computing, the most optimal arrangement of sensors, user interfaces, and input methods is still being worked out, and many of the studies here describe previously untouched areas of work. While immersive technology has yet to mature, little can be said about the cognitive impacts of our intentional use of it and less about its impact on our self-knowledge.

Second, this same lack of technological readiness means that the complexity of development pipelines is high, limiting the functionality of the software. Often, a variety of development environments, toolkits and custom-written applications are needed to produce immersive, authentic results. The consequence of this is that AR authoring applications either did not acquire the complexity within functionality needed to reach meta-cognitive processes or that there was simply too much else to do, given the possible scope for technical development (see [9]) that AR provides.

Table 2: The Knowledge Dimension

Major Type	**Description**	**Sub-types**
Factual Knowledge	The basic elements students must know to be acquainted with a discipline or solve problems in it.	Knowledge of (Ko.) terminology Ko. specific details and elements
Conceptual Knowledge	The interrelationships among the basic elements within a larger structure enable them to function together.	Ko. classifications and categories Ko. principles and generalisations Ko. theories, models, & structures
Procedural Knowledge	How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods.	Ko. subject-specific skills, algorithms Ko. techniques and methods Ko. criteria for determining when to use appropriate procedures
Metacognitive Knowledge	Knowledge of cognition in general as well as awareness and knowledge of one's own cognition.	Strategic knowledge Knowledge about cognitive tasks Self-knowledge

Cognitive Process Dimension

In this dimension we are looking at which learning activities are supported by the authoring implementation. Actions that go into a particular activity can be grouped along the lines described by the RT, arranged accordingly to their complexity. At one end of the scale is remembering, the action of recalling previously learnt knowledge. A mainstay of educational systems, this process is the easiest for which to construct objectives. The remaining processes—to understand, apply, analyse, evaluate, and create—allow the learner to progress from novice to expert and, though this route to competence will feature each of these parts, it will not be a sequential experience of each type of cognitive function. As anyone who has mastered a skill knows, each is visited many times and accessed when needed, though it is not until we have a basis at one level of complexity that we can develop working knowledge of the next.

Understanding, the most varied in its linguistic expression, follows remembering and includes the concept of representation, taken here as the transcription of information from one medium to another, such as playing music by ear or creating a written transcript of a conversation. This sub-type is often used in AR authoring tools, given the many media channels available with the technology. Several other sub-types exist, including exemplifying, interpreting, summarising, and inferring, all of which point to a more structured, detailed or nuanced comprehension of knowledge.

Next up is applying, a cognitive process closely linked to procedural knowledge, which can be further delineated according to the familiarity of the learner; when a task is known, the process is to execute knowledge; when unfamiliar, it is regarded as implementation. In this case, the understanding of some conceptual knowledge is thought to be a prerequisite to the application of procedural knowledge [4]. The fourth cognitive process, 'Analyse', involves a systematic look at the constituent elements of a structure and seeks to find knowledge about their interrelation and role in the system as a whole.

Evaluate, according to criteria or standards, may be a qualitative or quantitative exercise and relate to prior understanding and, potentially, analysis. Judgements about efficiency, effectiveness, or consistency address internal criteria while critiques or checks address those external to the system in question. This complex cognitive process can be applied to material objects or abstract concepts and may be used iteratively with 'understand' and 'analyse'.

The last category, Create, focuses on the putting together of entire structures, models, or patterns in a novel way. In the learning context, this is distinct from purely free expression as it relates to a learning objective, though it involves some degree of creativity. The RT breaks the creative process into three phases: problem representation, solution planning, and solution execution, with corresponding cognitive processes: generating, planning and producing [4].

It is important to note, especially given the mode of operation of AR, that the cognitive processes take place in a learning context that cannot be disentangled from the process itself. While this does complicate their examination, it is also a significant boon for AR research, where the experience of the physical environment can be altered to intentionally improve the learning context.

Table 3: The Cognitive Process Dimension

Major Type	Description	Sub-types and Synonyms
Remember	Retrieve relevant knowledge from long-term memory.	Recognising, identifying Recalling, retrieving
Understand	Construct meaning from instructional knowledge, including oral, written, and graphic communication.	Interpreting, clarifying, representing Exemplifying, illustrating Classifying, categorising, subsuming Summarising, generalising Inferring, concluding, predicting Comparing, contrasting, matching Explaining, constructing models
Apply	Carry out or use a procedure in a given situation.	Executing, carrying out Implementing, using
Analyse	Break material into constituent parts and determine how the parts relate to one another and to an overall structure or purpose.	Differentiating, discriminating Organising, integrating, structuring Attributing, deconstructing
Evaluate	Make judgement based on criteria or standards.	Checking, detecting, monitoring Critiquing, judging, testing
Create	Put elements together to form a coherent or functional whole, reorganise elements into a new pattern or structure.	Generating, hypothesising Planning, designing Producing, constructing

Using the taxonomy to assess AR learning affordances

Based on the revised taxonomy and a deeper appreciation of how the dimensions relate to AR authoring, we can draw up a table of learning affordances that we might expect to see. Following the spirit of Anderson et al.'s work, the focus here is on the application of the RT through examples. Table 2.2.1 provides descriptions and illustrative comments on how certain features of an AR authoring system or toolkit can provide learning affordances across the various knowledge types and cognitive processes.

Cognitive Process Dimension

Knowledge Dimension	1. Remember	2. Understand	3. Apply	4. Analyse	5. Evaluate	6. Create
A. Factual Knowledge	The system supports the recall or naming of details or visual elements. (E.g. matching markers to specific virtual objects.)	The system supports the aggregation, comparison of factual elements.	The authoring tool supports the use of terminology or encourages the user to make use of an augmentation.	The system aids the user in analysing real or virtual information, perhaps exposing the underlying structure of it.	The authoring tool supports or facilitates the evaluation or critiquing of facts.	The tool supports the generation of hypotheses or the production of new factual data.
B. Conceptual Knowledge	System aids in the recognition or grouping of principles. Identify a known model, such as a set of map directions.	The software supports the abstraction, re-imagining or grouping of principles or ideas.	The software supports the user in applying a principle or model in a new situation.	The software supports the integration of disparate concepts or helps to structure	The tool provides sufficient insight into a concept or model so that its validity or consistency can be verified.	The authoring tool helps in planning or devising a model or theory that explains observed phenomena.
C. Procedural Knowledge	The AR system prompts the use of a technique, skill, or method (competence).	The system contextualises a skill or technique or compares practice.	Knowledge of the sequence or structure of the procedure is enhanced using augmentations.	The system supports the user in (re)organising the procedure according to their knowledge.	The software allows the user to compare procedure knowledge and supports decision-making.	The software supports the design of complete, coherent models or their reorganisation into new forms.
D. Meta-cognitive Knowledge	Not included in this review. See note.					

Table 4: Authoring Learning Affordance Examples

2.2.2 Other AR Literature Reviews

Over the past few years, there has been enormous growth in the number of examples of industrial implementations of AR technology. There have been systematic reviews conducted of AR in industry [114] and education [32] and others that are industry-specific (e.g. automotive, aeronautic) or function-specific (e.g. maintenance, manufacturing). A systematic review was also conducted by Limbu et al. on the use of sensors when training people with AR [69].

Egger and Masood (2019) published an extensive review of industrial AR based around the technology, organisation, environment framework [27]. Palmarini et al., in 2019, focused on maintenance applications [85], extracting 30 studies published over two decades to 2017, noting a preponderance of marker-based studies. Bottani and Vignali's review, also in 2019, showed a marked increase in applications and technical papers, beginning in 2013 [13]. In education, Ibañez and Kloos (2018) performed a systematic review of AR in STEM settings [50], providing insight for instructional design processes. Parmaxi and Demetriou (2020) focused on AR language learning, highlighting trends in both activity classification—nearly two-thirds of papers supported skills training—and specific skills—around a quarter tackled vocabulary [87].

Industry-specific reviews have also been conducted by Boboc et al. (2020) for the automotive industry, Fraga-Lamas et al., looking at the shipyard of the future [31], Safi et al. for aerospace [99] and Cheng et al., who looked at the areas of architecture, engineering, construction, and operation (AECO) [16]. The authors did not find any other reviews addressing authoring functionality or those that extract learning affordances from technical implementations of AR systems.

2.2.3 Review Methodology

This review summarises the state of the art in AR authoring tools. It seeks to understand how the various elements within Bloom's Revised Taxonomy are represented using such systems by looking at how the technology was used in practice and inspecting the technological capabilities from the perspective of learning affordances. This literature review follows the PRISMA structure for the construction of the dataset. This consists of a sequence of four phases, shown in Figure 1.

After selecting a set of search terms—words commonly used to describe the process of creating AR experiences that are dynamically structured with selected content by the user (see Section 2.2). The data were then cleaned, removing empty rows or malformed entries. Next was a process of screening, first more generally by inspecting titles and abstracts, and then by reading the full text. The final set was then assessed and categorised along the lines of Bloom's revised taxonomy, providing a picture of how the authoring toolkits support learning through the provision of affordances.

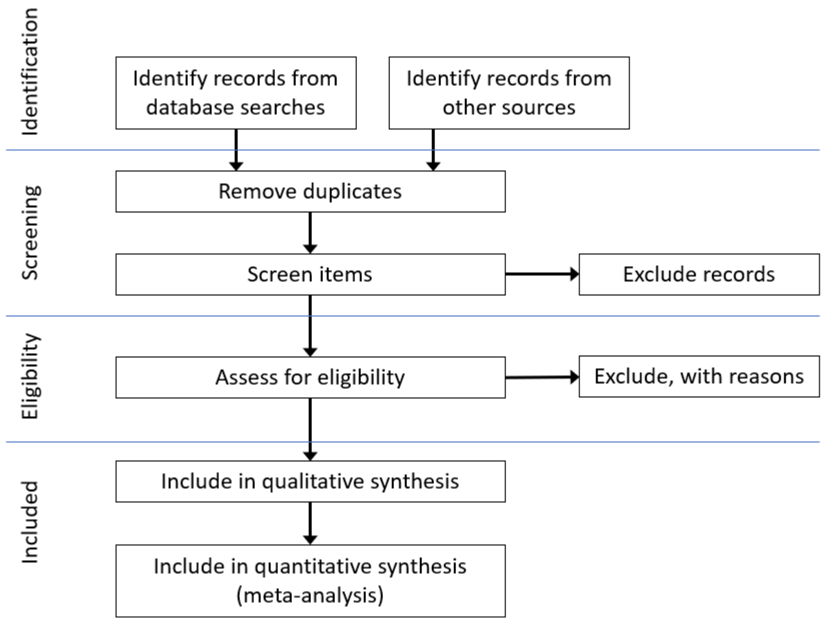


Figure 1: PRISMA systematic review phases, redrawn from [80]

Identification

The terms “augmented reality” and “mixed reality” are used interchangeably in the literature and, for the purposes of the search, were used interchangeably. In addition to this, the terms “authoring” (the most common), “composing” and “creating” were selected to capture research focused on the activity described in Section 2.2. Two other terms—"editor"” and “content management"—were also included to capture those papers that were grammatically focused on a noun describing the system rather than the function it performs. The final list of search terms is shown in Table 5.

Aside from the manual addition of papers that were considered relevant, two sources were used for the identification of research. The first was Google Scholar, accessed through the application ‘Publish or Perish’, which provides a structured format for search results. The search was constructed using the application’s user interface. The second was Lens.org, an engine referencing Microsoft Academic and Crossref, and the final search string used was: ‘title:("authoring" OR "composing" OR "creating" OR “editor” OR “content management”) AND title:("augmented reality" OR "mixed reality")’.

"Augmented Reality"		"Authoring"
		"Composing"
OR	AND	"Creating"
		"Editor"
"Mixed Reality"		"Content Management"

Table 5: Search terms

Screening

The screening process removed results that were incomplete or incorrectly formed, those that were not accessible in English, and those that could not be found or accessed through online portals. Duplicate entries were also identified and excluded, and patents were separated from the list. Publications were excluded if they met one or more of the following criteria:

EC1) was not peer-reviewed.
EC2) was itself a review paper or did not provide a demonstration of an implemented system.
EC3) the study was found to be a poster or workshop.
EC4) the study was not related to or applicable to AR authoring.

The inclusion criteria were based on the definitions of 'AR' and 'authoring' given above:

IC1) a primary study that uses AR
IC2) a primary study that demonstrates authoring capabilities

Eligibility

During this phase, a deeper level of assessment was made of the subject matter of each paper. First, the abstract was read, and it was determined if the paper met the criteria for demonstrating both AR and authoring. If there was ambiguity or these topics were not mentioned, the full text was read to determine its relevance.

Determining whether the research item was on topic involved an assessment of the relevance of the work; it should demonstrate the implementation and/or design of an AR authoring toolkit. Once the general theme of the paper was found to be relevant, the definitions (given above) of both AR and AR authoring were taken into consideration. If the paper was found not to meet either one, the entry was marked as "not true AR" or "not authoring", respectively. These categories can also show the amount of research that is excluded simply as a result of drawing up these definitions.

Categorisation

The assessment of the research and the assignment of categories to it is based on an understanding of the principles of both AR authoring and the RT. Despite the clarity of the RT, the categorisation process relied on some assumptions being made about the learning affordances expressed, for the following reasons:

- The papers did not explicitly use the RT to design their learning outcomes, so there is no explicit intent to reveal or otherwise support a particular knowledge type or cognitive process.
- More broadly, there were rarely cases in which the learning outcomes were explicitly stated.
- Often, the details given about the experience of the user are limited so the reader or analyst must infer which learning outcome is being met.
- Though an authoring tool may have designed affordances, in implementation several reported functional limitations or errors in the system, restricting what the user could do and, as a consequence, impacting their potential for learning.
- The functions that provide evidence for specific results will not overlap perfectly with those that support learning, so while a paper may focus on a feature that provides academic novelty, for instance, the learning outcomes might be unrelated, leading to little space being given for their description or explanation.

To account for these discrepancies, for each paper that was categorised, one or more quotes were also recorded that provided evidence of the choice or highlighted a key aspect of the authoring tool.

It is hoped that these examples, together with the discussion below, will make the method of interpretation of the RT more clear. Rather than provide an objective standard for what constitutes a (reported) learning affordance, we aim to present a lens through which to view AR authoring systems that others can use and which can contribute to the discussion on how to make better experiences with this technology.

2.2.4 Results

The searches conducted in the identification phase yielded 553 papers and 71 patents, which were not included in this work. The initial screening phase saw this number reduced to 521 papers by removing duplicates and incomplete or malformed entries. A second round of screening was then carried out to identify those papers that were clearly off-topic, not locatable online, or not available in English.

After the eligibility phase, in which the research was assessed in detail for its relevance to the topic, 131 papers remained, which were downloaded and collated in preparation for categorisation. During this categorisation process,

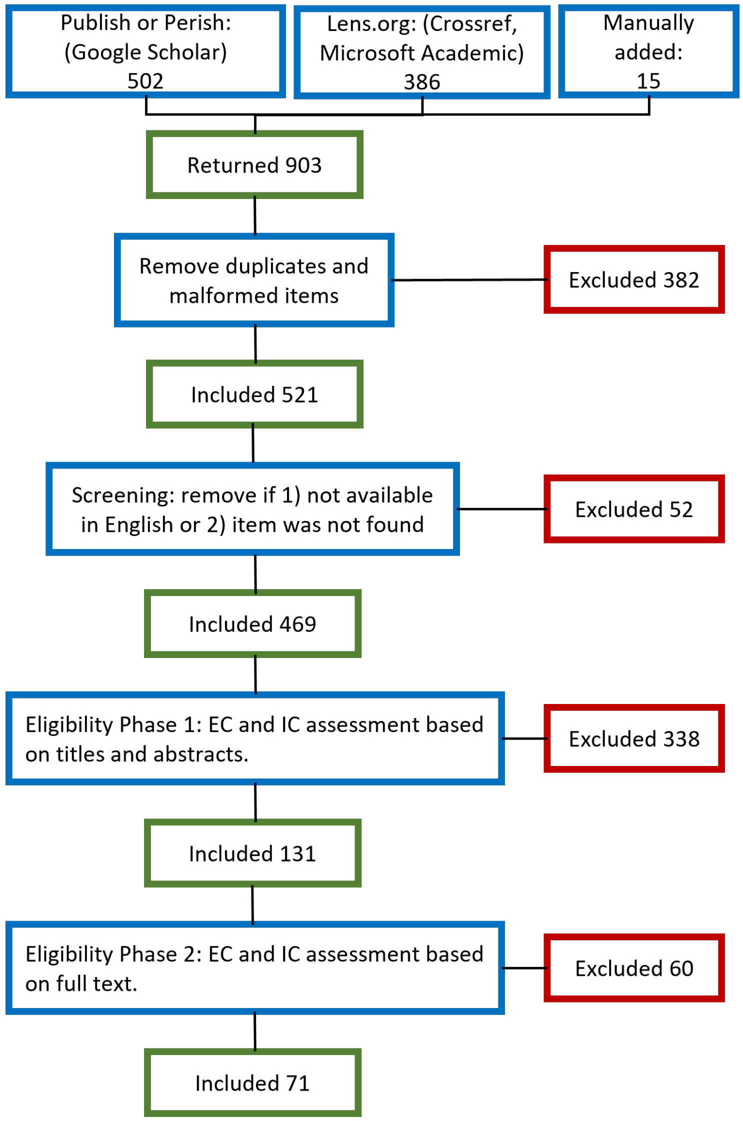


Figure 2: Selection process for AR authoring studies

the specific details of the papers were scrutinised in terms of learning affordances. This led to additional exclusions, the majority due to the work being marked as “not AR Authoring”, bringing the dataset to its final size. The final dataset of 71 items is presented chronologically in Figure 4.

Table 6: Breakdown of reasons for exclusion from dataset.

Source	# excl.	Type	# excl.	Content	# excl.
Not found / no access	22	Poster or workshop	4	Not "AR"	28
Not available in English	30	Review paper	17	Not "Authoring"	43
		Not peer reviewed	10	Off-topic	242
Total excluded	*396*				

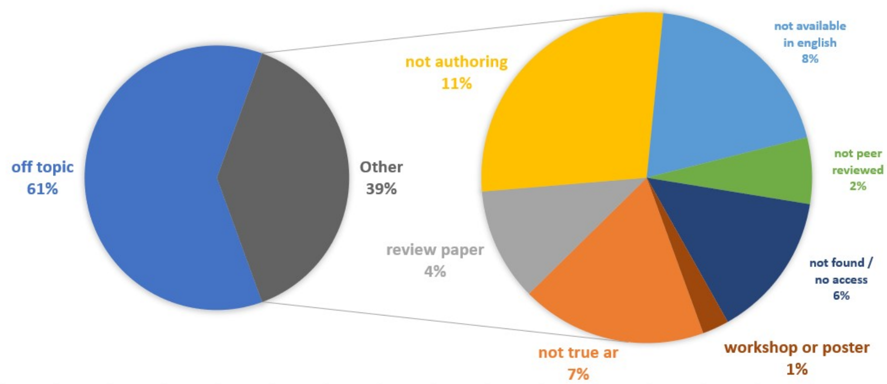


Figure 3: Overview of reasons for exclusion from dataset.

The chronology of the papers can be seen as having three phases, corresponding to improving levels of technology readiness and computational performance. Early work, followed by a 3-year hiatus, was then picked up by a larger community of researchers in a second wave. Better equipped with smart phones and tablets, the period between 2009 and 2012 showed considerable interest, with 2012 delivering nine works on AR authoring. The trend again slowed, but there has since been a continuously expanding body of knowledge, fuelled by recent technological developments in depth mapping hardware and their integration into head-mounted devices, and from there to the most modern form, the waveguide-based HMD.

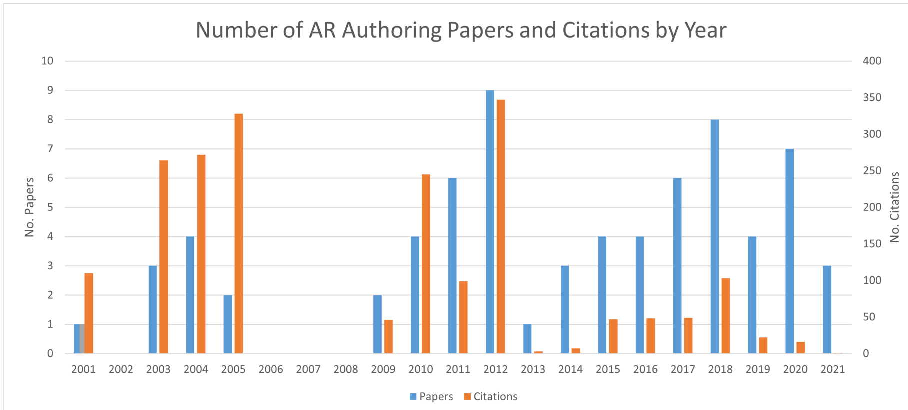


Figure 4: AR Authoring publications and citations, 2001 to 2021 (N=71)

2.2.5 Categorisation

This section will trace the development of key systems that allowed these authoring tools to exist, focusing on the learning affordances that were designed by the developers and researchers involved. To clarify, we are interested in those functions that provide affordances to the learner, rather than the teacher or designer of the experience.

Early work (2001–2005)

Many of the authoring solutions use similar underlying technologies, such as ARToolkit, which was developed by Hirokazu Kato, Mark Billinghurst, and Ivan Poupyrev and first released in 1999. This was one of the first tracking libraries that would allow developers to connect markers to augmentations and is frequently cited in the research found in this dataset. The first open-source version was released through HIT Labs in Washington in 2001.

The first example of an AR authoring system, Tiles, comes from the same group of researchers (Poupyrev et al. 2001). "The Tiles system is a prototype tangible augmented reality authoring interface that allows a user to quickly layout virtual objects in a shared workspace and easily manipulate them without the need for special-purpose input devices" [89, p.7]. The virtual content was arranged using square-printed markers (tiles), which were each linked to a 3D visualisation.

The assignment of content to marker objects can support recall (A1) and the use of functional tiles (such as the 'delete' option) requires understanding of the object (A2). Tiles could also be used in conjunction with others, producing results such as 'copy-and-paste', or providing application-level functionality. These types support learning about the interrelations of the objects (B1) and provide a deeper understanding of the content at hand (B2).

Most other early examples come from the Authoring Mixed Reality (AMIRE) framework. Abawi et al. (2004) highlight four tasks that the author can perform: "qualification of MR components, adaptation of the selected MR components, combination of the MR components to allow interaction between them and calibration of real objects and virtual objects in the application" [1, p.115]. This illustrates the support of remembering factual knowledge (A1), its adaptation (A2) and the combination of components using a given model (B1) in order "to connect two parts of a story" [1, p.116], showing a degree of understanding of the theme (B2).

Zauner and Haller (2004), working on a related project, illustrate four modes related to MR placement: "The first one is the observation mode. This mode allows the author to take a look at the actual placement. Further, the placement marker is visualised by a white... In the second mode, the author is able to modify the size of the object... The third mode is used to move the object...The rotation mode, which is the fourth and the last mode, enables the author to rotate the object." [139, p.6].

In this description of system use we find evidence of the recall and understanding of factual knowledge (A1, A2) as the user can implement affine transformations (translation, rotation and scaling), but without the application to a conceptual area or the integration with procedural knowledge, no further categorisations are assigned. This principle is adhered to throughout the categorisation of the sources, as there is often a demonstration of the functions supported by the user interface without an application to a particular model, principle, or concept (conceptual knowledge).

In 2005, Haller et al. also described the functionality within the AMIRE project, adapting the previous tools to support a simple construction task where "[a]ugmented, highlighted boards help users to find the right boards and to mount them in the right order" [44, p.5]. Given that the toolkit was deployed on a hand-held device (which must be put down to handle the parts), this provides clear evidence in support of procedural knowledge retention (C1). In addition, "users get information on where to place the different components and how to connect them" [44, p.5], exemplifying their procedural knowledge (C2).

Another notable example of pioneering work in the AR authoring field comes from Kirkley and Kirkley in 2005 [58]. The most cited of all the papers in this dataset, this research presents an authoring tool called Information In Place Inc. (IIPI) CREATE. The system, originally prototyped for military use, discusses many of the constructivist learning principles reflected in the RT but provides little evidence of the way in which this tool demonstrates learning affordances in AR. The authors state that the tool "does not intend to replace the existing tools that various designers and developers use but to provide an organising, shared framework for various types of individuals as they create these next-generation learning environments" [58, p.50]. Based on this, the system is shown to support the recall and summarisation of other factual information (A1, A2) and can help to organise the principles and models found in other design environments (B4).

Mobile AR materialises (2009–2012)

While early studies showed promise, a lack of technology readiness limited their utility in the wider market. This began to change with the release of the first iPhone in early 2007 and the first Android phone in November 2008. January 2010 saw the first iPad, and later that same year, both Samsung and Research in Motion (RIM) launched devices with the same form factor.

These devices, although the first in their class, paved the way for more interactive, powerful applications. Given the computational requirements of mobile "pass-through" AR, it was at this point that the tools found a wider audience and application. ARToolkit also evolved, with a Flash version being released in 2008, entitled FLARToolkit. This toolkit was used extensively by creators of AR authoring systems during this period. The lighter implementations, such as those by Sano (2011, 2012) [102] [101] and Jee et al. (2011) [53] used only the aforementioned affine transformations (A1, A2), with others, such as Wang et al. (2010), extending this so that "a user can type in the question and choose the correct answer to the question in the question input section" [130, p.286], demonstrating the evaluation of the facts presented to them (A5). Such lightweight demonstrations were, by this point, generally being overtaken by more feature-rich systems, such as that demonstrated by Mader and Urban (2010), which included hand tools and visual instructions for disassembly or repair operations. Instructions such as "remove the screws that hold the hard drive using a Philips screw driver size PH2" [73, p.12], illustrate knowledge of terminology (A1), factual understanding (A2) and knowledge of subject-specific techniques (C1). Furthermore, they incorporated the use of schematics (B1) to guide the user through actions (B2).

This use of AR authoring to more closely follow a procedure is also reflected in work by Fei et al. (2012), who used an action graph—"an augmented version of the basic finite-state machine" [28, p.12] to allow the design of complex procedures. "[A]ctions [...] in the action graph that represent interactive objects or passive animations of the scenario and can be bound to the virtual objects, transitions (namely edges) represent changes of the storyline and can be bound to certain events" [28, p.22]. The use of familiar objects (A1) and storyline transitions (B1) and their application in a procedural flow (C3) has the ability to approximate real-world complexity and captures the idea of contingency based on both object selection (A2) and the prediction of consequences (B2). In this time period, three papers stand out for their relevance to later work (as evidenced by the number of citations received). These are the AR Gamebuilder implementation (Klopfer & Sheldon, 2010) [59], Dunser et al.'s Augmented Books (2012) [24] and Langlotz et al.'s system for in situ content creation in 2012 [64]. To date, these have been cited almost 500 times, nearly a third of all citations represented by the dataset. Klopfer and Sheldon present authoring software that "includes the ability to bring in new maps, position characters, and items, make logical connections between these items, incorporate data, and insert media" [59, p.89]. Encouraging the user to recall locations (A1) and their relationship to the virtual content assigned there (B1) promotes recall. In later versions of the software, branching logic was also implemented, which could "offer a richer opportunity for students to understand issues in their communities and express their understanding of the relationships that surround them" [59, p.92]. The user is thus encouraged to develop an understanding of both factual (A2) and conceptual knowledge (B2).

Dünser et al. (2012) produced an authoring system that was used to create interactive physics education by augmenting printed books. In a similar fashion to other marker-based tools, "[b]asic interaction is provided by manipulating the book. Rotating or tilting the book pages shows virtual content from different positions and angles, and flipping book pages changes the virtual content displayed" [24, p.107] (A1, A2). However, the designers also added interactive features: "a user could select the desired image and start an animation by placing a finger over the black border ... In addition to this, some AR markers had interactions that allowed the participants to change the direction of particular components of the model, for example, reversing the direction of the current in a wire" [24, p.111]. The selection of markers, either for clarification or exemplification, shows understanding (B2), and the possibility to apply physical principles to an unfamiliar task is categorical evidence of the application of the user's knowledge (A3, B3). The interaction methods in work by Langlotz et al. (2012) were instead performed in the traditional way, using a touch-screen interface, introducing challenges around how to position the device and work on the content simultaneously, a hurdle that many subsequent studies would face. They solved this with a freeze mode, during which content would be added. Two types of features are mentioned: first, that an "object can be translated, scaled, and rotated by selecting the axis of transformation in the GUI" [64, p.628] (A1, A2). Second, in order to "enable the creation of more realistic models, we provide the functionality to assign different colours to the objects as well as to texture the object" [64, p.628], which demonstrates the recognition of the principles of colouring and texturing (B1).

The final grouping of research worthy of note in this time period is that arising from the workshop sessions run during the International Symposium on Mixed and Augmented Reality (ISMAR) during 2011 and 2012, which together added five research items to the final dataset. Grubert et al. (2011) augmented posters, and, while no virtual object transformation or interaction was mentioned, "the user selects regions for triggering the appearance of content as well as regions where to actually display the content" [39, p.2], mapping the content on the page to their design principles (B1, B2). Janer et al. (2011) devised a "tool that allows the creation of an augmented soundscape from a user-contributed sound repository" [52, p.1], requiring memory and interpretation (A1, A2) and an understanding of location-based concepts such as map overviews (B1, B2). Wozniewski et al., in the same conference, showcased their Object Creator, which was capable of cutting elements from the visual feed to use as an augmentation: "A template is chosen from a menu, and depending on the selection, a semi-transparent overlay will identify the part of the image to be mapped onto the geometry" [137, p.3]. This encourages differentiation and discrimination in the local environment (A4) and represents new ground in the way that augmentations could be crafted.

The following year, McClean et al. (2012) put forward an authoring system that “consists of dragging and dropping 3D widgets into the front parallel view of the façade. The user can then set the position, orientation, and scale of these widgets” [76, p.2], covering the transformational affordances (A1, A2) and an appreciation of parallelism and perspective for proper placement (B1, B2). Finally, Jens de Smit (2010) also published a system entitled Layar Creator, which allowed organising sets of related augmentations into ‘campaigns’. “Once at least one page has been added to the campaign, users can drag “buttons” from the right-hand side of the screen onto a page. These buttons... have an associated action to trigger when an end user taps the augment on the mobile client. Possible actions include, but are not limited to, opening a website, playing a video, and initiating a phone call" [113, p.2]. The use of content selection tools indicated support for recall and understanding of the facts present in the campaign (A1, A2) and their embedding in a context where utilising web-based connectivity promotes the use of conceptual knowledge (B1) so that the campaign can have a broader impact (B2).

The appearance of head-mounted displays (2013–2021)

The third phase, the period in which AR also becomes a head-mounted phenomenon, is supported by development in two areas. One is a new generation of depth-aware sensors, led by Microsoft in their release of the Kinect for Xbox One in 2013, which introduced time-of-flight technology into the devices. The Kinect, already a commercial success in the world of interactive gaming, could achieve better resolution, cheaper manufacturing costs (due to electronic calibration) and better resistance to ambient light distortions. This improvement would also feature in the HoloLens, released in 2016, which combined the depth camera with an optical waveguide, a transparent near-eye display, and advanced processing hardware. Another success was the Oculus Rift Kickstarter campaign, run in 2012, that would rekindle interest in the HMD. Examples combining time-of-flight cameras with VR headset visualisation could create an immersive equivalent of the mobile video passthrough. In terms of software development kits, the Vuforia computer vision library would be the next widely successful entry, released in 2011 and improved over a decade. Two dedicated AR development toolkits, ARKit for iOS and ARCore for Android, would follow in 2017 and 2018, respectively. Additional platforms gave rise to a more consistent and varied type of authoring toolkit, and in this period, many new features and approaches have been demonstrated. The singular authoring toolkit identified in 2013, from Ruminski et al., was one of the few to use a quiz format to support learning (A4): “the designer can select the onPresentationBegin event. Then he or she can choose the showQuiz action and can select from a menu an ID of the quiz. The selected quiz will be shown when the virtual object appears on the screen.” [97, p.11]. Yoo and Lee (2014) explored the use of gestures to control authoring functions, demonstrating the apprehension of a conceptual framework (B1) as well as eliciting indicative actions from the user (B2, B3)[45]. While Ty et al. (2014) demonstrated, in detail, the pipeline for building augmentations (including the use of affine transformations)[123], Santos and Luebke (2014) went on to show how these could be used in a learning setting to help recognise and understand the images in context (B1, B2): “teachers can modify the appearance of the image. In this example, it is desirable to scale and position the lungs correctly on the body” [100, p.555].

Noleto et al. (2015) gamified their approach to authoring, illustrating how the construction of conceptual models (B1, B2) can be made more intuitive: “For each mission, the game designer may include game mechanics” [82, p.106]. Furthermore, they used the same framework to encourage the players to remember procedure sequences (C1): “The mission to destroy the barricades will be composed of five mechanics ordered sequentially” [82, p.106]. Their outdoor, location-based mobile game also requires an internet connection, a feature emerging across many modern systems.

Kim and Park (2015) developed a content management system that “stores components such as 3D models, marker information, audio, and video in the database through the CMS Web Interface” [57, p.64], but otherwise only demonstrated visualisation functionality in the software (A1, A2). Similarly, the Argon framework, described by Speiginer et al. (2015), used web-sourced content and focused on the anchoring and visualisation of virtual content (A1, A2)[115]. This iteratively developed system did provide different modes for content viewing; for instance, “[i]f the user is at the locations featured, she can turn off the provided panorama image and have a true augmented reality experience overlaid on the live video feed from their device” [115, p.7]. This practice of considering different modes that consider the benefits and challenges of the target device would also become more frequent.

The first authoring systems to exclusively target a head-mounted device are described by Shim et al. and Yang et al., who, with their colleagues in the Media System Lab, published two papers in 2016 with this ambition. While a hand-held device was used, it was as a controller rather than a central display. Their work focused on gesture training, investigating tangible (i.e. physical) interaction, gesture-based interaction, and a blend of the two, using a marker on a smartphone to provide additional configurability to the marker layout, and using [112]. For hardware, they used an Oculus Rift connected to a SoftKinetic DS325, which is a short-range time-of-flight camera capable of hand tracking. Considering the learning affordances, each of the interaction types presents a somewhat different set of possibilities. The tangible interaction provides the user with a chance to apply procedural knowledge (C3) but without a broader conceptual basis outside normal movement. The touch-screen interface described there manages typical affine transformations (A1, A2). Finally, the blended element encompasses both of the above: “The mobile device interactions are tangible to the user and support discrete control; they also provide continuous interaction for manipulating 3D objects in combination with interactions that use hand gestures” [112, p.1430]. This furnishes learners with clearer examples and uses of the interaction model (B1, B2).

From 2017 onward, there is a greater proliferation of authoring systems, certainly impacted by the excitement surrounding the release of the HoloLens the previous year. The term ‘Augmented Reality’ was becoming pervasive in technology circles and more use cases were being trialled. Procedural knowledge—that which is related to how to go about meeting a learning objective—was more frequently addressed, and the overall complexity of the software architectures was generally increasing. SDKs for both Android and iOS platforms, both formally released in 2017, gave a much larger number of developers access to AR-enabling tools.

One area where a trend of increasing complexity is evident is around the augmentation of text. In 2017, Raso et al. combined data mining operations with AR visualisation, showing how the content could be situated in a physical (A2) as well as semantic context (B2)[92]. Lytridis and Tsinakos present AR-Tutor, which “enables students to enhance their understanding of the teacher’s material by displaying explanatory interactive digital content on top of the traditional book” [72, p.10], again allowing a user to extend their understanding of the material (A2), representing the information in other forms (B2).

Shekhar et al., writing in 2019, describe the connection of AR and natural language processing in such a way that “allows the user to include various commonplace objects and... can describe their relations (e.g., “A chair next to the bed”) and can include humans or animals (e.g., “A man is sitting on a sofa with a dog")." [110, p.64]. This presents learners with the opportunity to re-imagine spoken or printed ideas using visual aids (A2) and also to recall facts (A1) and ideas (B1) related to the content. With more connections being made between aspects of the various components of the authoring tools, it is no surprise that web-based implementations are also becoming more prevalent. In 2017 alone, four of the six authoring tools identified relied on web-based communication (Maia et al., Rodrigues et al., Reynolds, and Raso et al.).

The maturity of the systems being produced is also reflected in their implementation in workplace environments. Many early systems were used in a lab setting or controlled environment, avoiding many of the subtle and unpredictable elements of an active workplace. More robust localisation, scanning, and mapping algorithms meant that quick movements, hard-to-scan surfaces, and light levels were less disruptive, reaching a usability threshold that led to greater commercial interest. The first instance of this was reported by Schleuter in 2018, who used the Unity game engine to provide a remote-assistance service that, in one instance, “guided a user through all disassembly steps, including the removal of the front bolts, plate, gasket, spring, and pistons” of an engine [105, p.47].

A study by Blattgerste et al. (2019) is unique in that it is the only authoring toolkit that directly aims to support those with cognitive impairments. A specific case was presented, “where the workers hove to perform the cleaning of a modern coffee maker... that incorporates 20 distinctive steps (e.g. pressing specific buttons and disassembling, emptying, and cleaning specific parts of the machine” [10, p.7]. There is a clear engagement for the application of knowledge of both factual (A3) and procedural types (C3), supported by the range of typical functions seen in previous authoring systems, including the retention and understanding of both facts and principles (A1, A2, B1, B2).

The use of sensors has also recently been demonstrated as an extension to previous authoring toolkits, where they are either distributed in space (Jin et al., 2018) or around the body (Limbu et al., 2018). In the first, the authors constructed a system called "TanCreator, which allows children to create AR maze games and operate AR characters with paper tokens and sensors in the real world" [54, p.85]. The use of definable inputs allows learners to apply previously understood knowledge (A3, B3) and, given that the system also permits the sequencing of actions to achieve set goals, also the application of procedural knowledge (C3).

The use of body-worn sensors in an authoring toolkit was reported by Limbu et al. (2019). In that framework, which was called Wearable Experience for Knowledge Intensive Training (WEKIT), the use of the HoloLens permitted the tracking of both the head and, when visible, the hands of the user, known as a 'Ghost track', which "allows visualisation of the whole-body movement of the expert or the earlier recording of the trainees themselves for imitation and reflection" [67, p.160]. In this sense, the system offers the learner not only a method to apply procedural knowledge (C3) but also a method to evaluate the action with reference to that of the expert (C5).

T. Wang et al. (2020) produced a similar recording track of the body and hand motion, building it into a context-aware application (CAP): "An instructor wants to demonstrate his or her routine task of repairing a bike, so he or she creates a CAP using CAPturAR with three sequentially connected events: shaking the lubricant, spreading it on the front wheel and then on the back wheel. A novice comes and follows the tutorial [...] CAPturAR detects once the novice completes a step and starts to play the demonstration of the next step." [131, p.333]. Again, the ability to mimic the procedural demonstration of an expert facilitates evaluative learning (C5) while also providing information to support the understanding of the procedure (C2), together with factual knowledge about the parts of the bicycle (A2) and their interconnections (B2).

Perhaps one of the most visually complex, but nonetheless graphical, is BlocklyAR (2020), created and described by Nguyen et al. This tool is a graphical coding framework where elements can be fit together to produce a definition of an AR scene, which is then converted into executable code. "The definition describes the block's appearance and how it behaves, including the text, colour, shape, and connections to other blocks while the generator translates the block to executable code" [81, p.5]. The ability of this framework to let the user manage affine transformations (A1, A2), non-linear sequencing of events and triggers (B1, B2), and apply this knowledge (A3, B3) gives us the first set of general affordances. Since "the visual AR component enables enthusiasts to experience their coding schemes in the mixed 3D space" [81, p.10], we can also attribute a creative affordance to this system, as the user is constructing an entire model based on the factual information provided by the tool (A6).

Chidambaram et al. (2021), in a system targeting head-worn devices, also incorporate object recognition algorithms into their system, though they use registration for the initial location of the virtual content, as performance limits remain a factor when performing complex tasks (feature recognition). In practice, this means that "novices can view videos previously overlaid by the expert users to complete the current procedural task in progress" [17, p.243]. This assists the user in understanding the video, which contains either factual (A2) or conceptual knowledge (B2) but will also support recall (A1, A2). The focus is on retaining and understanding procedural knowledge (C1, C2) and also supporting its application (C3). "ProcessAR also enables the user to perform voice recordings for the purpose of clarifying tasks or to explain possible error cases" [17, p.242].

2.2.6 Findings

A summary of how the categories identified in the dataset are distributed over time can be seen in Figures 5 and 6, showing the knowledge dimension and the cognitive process dimension, respectively.

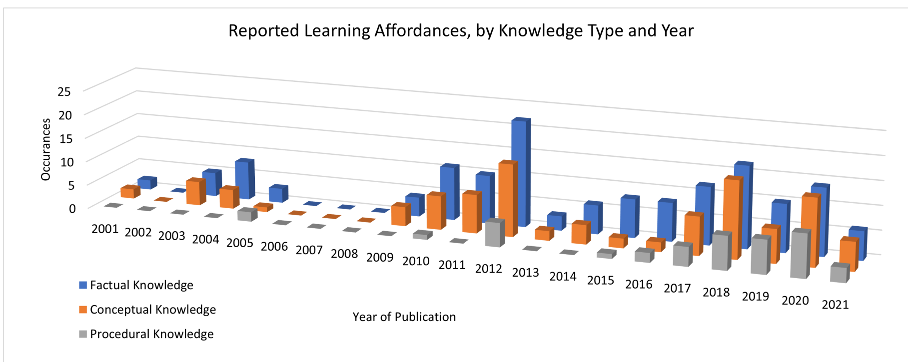


Figure 5: Reported learning affordances, grouped by Knowledge Type

Despite early implementations focusing almost exclusively on factual and conceptual knowledge, newer technologies supporting body tracking (with either machine vision or a HMD) have provided an excellent basis for adapting AR in support of procedural knowledge learning. A steadily increased body of knowledge on this topic will certainly open routes to interesting further study.

In terms of learning actions, the top three levels in the Cognitive Process Dimension—analyse, evaluate, and create—were rarely demonstrated in the literature. It is clear that more complex implementations of AR authoring systems will tend to have more classifications, as each learning affordance requires additional programmed functionality. The complexity of code needed for, say, a geometric comparison of two 3D models (as part of evaluation: A5) will be greater than that needed to only visualise the models (in support of memorisation: A1). As such, the degree of support for increasingly complex learning objectives has steadily increased, though this has been punctuated by the development of improved technology or entirely new devices. There now exist,

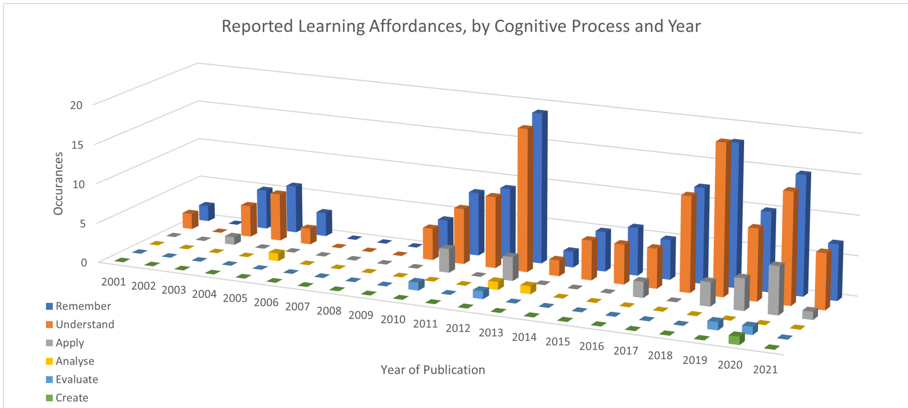


Figure 6: Reported learning affordances, grouped by Cognitive Process

however, examples demonstrating AR-driven support for all major categories; the research arena is primed to produce additional insights built on these and will certainly deliver further advances in all areas.

The consolidation of early efforts to accurately and easily place, sort, and manage visual elements is more or less complete; new applications can quickly adopt this functionality, and there is little novelty left in working only for this purpose. Newer manifestations of AR authoring systems are thus much more focused on the combinatorial gains of AR platforms with other (often web-based) tools or the extension of functionality into specialist arenas, such as industry-specific training.

In summary, what is apparent is the maturation of a new type of application that is constrained by technology and the environment and encourages both participation and creativity from the user. From Tiles in 2001 to BlocklyAR in 2020, this paper has mapped the first learning affordances to the creation of an entire, testable framework. It spans the entire range of the Cognitive Process Dimension, showing how software developers, just as any other learners, build upon the work of others and make their way through the levels of constructivist learning.

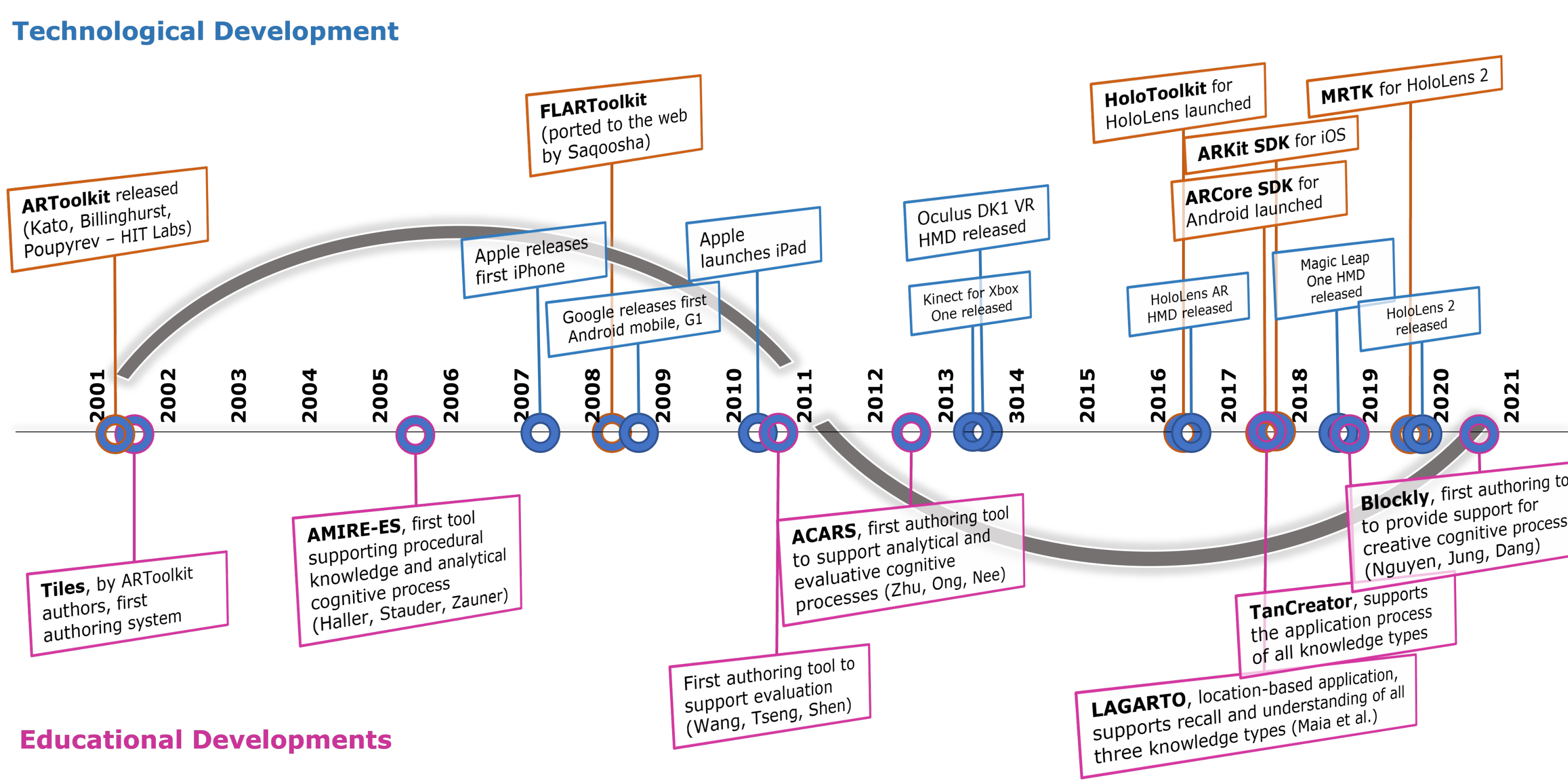


Figure 7: A summary of notable technological and educational developments in AR authoring, 2001–2021.

2.3 The WEKIT Project

The work presented in this thesis was part of the project Wearable Experience for Knowledge Intensive Training (WEKIT), which was a consortium of organisations and academic institutions from across Europe. Oxford Brookes University was a technical partner, and the work was conducted alongside developers from the Open University of the Netherlands (OUNL), VTT, and Ravensbourne University London, who also designed and produced the smart garment into which sensors were integrated.

Members of the WEKIT consortium have published closely related research. Limbu et al., writing in 2019, discuss the use of instructional design methods when using AR and wearable sensors [68][66]. These papers offer a useful complement to the software architecture described here, though they build on an alternative theoretical framework. Vovk et al. ran a special interest group on spatial interfaces for AR in order to uncover insights on this topic through the practical design and use of cutting-edge HCI tools [128]. Participants in the use cases described here also contributed their perspective to aid other research, supported by metrics that assess system usability [47] and simulator sickness [129].

During the development of the first prototypes of the ECS, affordances were identified as having descriptive power. A short paper was published by the author, in collaboration with several others, in 2017 about several of the devices that were candidates for providing data [40] to the system. Also already published by the author is a paper describing the first iteration of the structural equation model used here. This paper provides more details on the validation of the initial questionnaire and some of the metrics used in the assessment of the model [43]. The author also published work in 2019 that gave an overview of the software architecture that is presented in more detail below [41]. That paper also included a pattern language that connected the wearable, social, and programmatic layers of the experience capture system. A subsequent, open-source implementation of the ECS is also briefly summarised in a chapter published in 2021 [42].

2.4 Wearable Technology to Support Augmented Reality

We often use movement, in the form of a gesture or action, to illustrate our words. This is especially true when we wish to express concepts that are themselves not learnt through verbal instruction, such as how to use a certain tool, instrument, or piece of equipment. Hand movements and gestures are closely intertwined with communication; we use our hands to aid memory recall or to support or replace speech. McNeill even goes so far as to suggest that we must use both verbal and non-verbal communication in order to fully conceptualise and communicate our thoughts [77].

Considering the role of movement-based communication in the use of AR authoring tools can be interpreted in two ways. Either as a means to instruct the AR system, drawing on insights from the field of human-computer interaction (HCI), or as a data source that the AR system can capture and replay, providing a visual record of physical action, perhaps to help a learner construct new knowledge.

Capturing the range and subtlety of such movement, however, presents several challenges. The human arm contains around nine degrees of freedom (DOF). The highly complex shoulder joint, joined to the rest of the skeleton by only the clavicle, allows for humeral movement about all axes and, with careful orchestration of several additional muscles, allows for the movement of both the scapula and clavicle [104]. Besides flexion and extension of the elbow, the forearm also has around 180 degrees of possible rotation, adding close to two degrees of freedom. The wrist may flex, extend and deviate in either the ulnar or radial directions, adding another two. Furthermore, it has been shown that, in purely virtual environments, we may permit some error and still perform the required tasks [6]. If the overlay in AR is inaccurate, however, the mismatch between the expected result and the one achieved will be stark.

An early example of wearable HCI was the Power Glove, produced by Nintendo in 1989, which was capable of sensing movement and the position of the hands, together with information on finger position for each finger. It did this using a combination of conductive ink, a gyroscope, and two glove-based ultrasonic transmitters, whose positions were triangulated by static receivers. These also estimate yaw and roll. Only two games, Super Glove Ball and Bad Street Brawler were released for use with the glove, and these were hampered with complexity and performance issues and the glove ceased production after a few years.

Since this time, the explosion in the power of microprocessor architectures has paved the way for more advanced systems, where a variety of hand-held controllers provide exceptional sub-millimetre positional tracking accuracy with instantaneous feedback to head-mounted displays. However, since they are hand-held, our normal interactivity with the world around us is severely limited, removing a key immersive quality from an AR experience.

Inertial Measurement Systems

Body sensor networks (BSNs), specifically those equipped with IMUs have become more widespread due to advances in production, modularity, and the open-source movement. Filippeschi et al. [29] review many of the previous efforts in IMU-based human motion tracking (IHMT) and present five distinct methods for characterising human movement at different levels of complexity, both in algorithmic and architectural terms. Work has been done on gait analysis for the purpose of supporting the diagnosis of Parkinson's disease [21]. Other approaches have used ultrasonic position measurements to support the IMUs own frames of reference [126] and there have been good strides made into full-body motion capture solutions, especially by Xsens in 2009 [95] and Vlasic et al. in 2007 [126]. The latter, in conclusion, described the need for solutions to both the drift and axis-mapping problems.

Myoelectric Control Systems

The use of myo-electric devices to support prosthesis control has been around since the middle of the 20th Century [7], more recently extending the interface design through gamification [119], with others incorporating signals from multiple types of sensors [108]. Due to its low cost, one of the more widespread interaction tools was the Myo armband, produced by Thalmic Labs Inc. [51]. It contains an ARM Cortex-M4 microprocessor that handles data from eight skin-conducting electrodes, evenly spaced and housed in a wearable unit capable of adapting to various arm sizes. On-board processing provides a 50 and 60 Hz notch filter to remove power line interference. The Myo armband also contains an MPU-9150 IMU, consisting of an accelerometer, gyroscope, and magnetometer.

The device's limitations have been noted to be around sensor positioning and the rate of data collection, which is 200 Hz. This is significantly less than devices used in clinical settings (typically 1000 Hz), though there have been efforts to determine the usefulness and recognisability of the pattern shown in this data. The application programming interface (API) provided with the hardware also only permits communication with a single device at once. Despite these drawbacks, the Myo has contributed to several implementations.

Phinyomark et al. (2018) studied reduced accuracy based on a 1000–200 Hz change. Previous work by the same authors has focused on identifying useful feature sets for the classification of hand and finger gestures. The Prosthetic Hand Assessment Measure (PHAM) training protocol was used alongside the Myo armband and inertial measurement units for display in the HoloLens in a work by Basker [6].

2.5 Unified Theory of Acceptance and Use of Technology

The structure and content of the Unified Theory of Acceptance and Use of Technology (UTAUT) model have their roots in the Theory of Planned Behaviour (TPB) [2], itself stemming from the Theory of Reasoned Action [30], a theory that holds that we 1) make logical, reasoned decisions about the information around us and 2) engage in behaviours that originate with an intention to act. While we are considered rational actors, the influences that underpin this rationality are multivariate and may act to reinforce or diminish one another. Highlighting this kind of internal dynamic, TPB added the notion of perceived behavioural control and also provided a clear basis for constructs such as (computer) self-efficacy, where a person's reflective belief about their own degree of control or competence is the subject of investigation.

In 2003, Venkatesh et al. published the UTAUT model [125] in an attempt to bring together many of the approaches that were in use at the start of the millennium. Their model initially saw four constructs: performance expectancy (PE), effort expectancy (EE), social influence (SI), facilitating conditions (FC), loading onto (influencing) two others, behavioural intention (BI), and use behaviour, a measure of the actual usage frequency (UF) of the technology. BI is a central concept: a stronger intention to use a technology will, generally but not without exception, lead to more frequent use of it. BI has been shown to be a reliable predictor of eventual use frequency (UF) [25].

An updated, extended version of this model (UTAUT2) was published in 2012 by Venkatesh, Thong, and Xu [124]. Here, three other constructs were added, each thought to also predict variation in the behavioural intention to use a technology. They are hedonic motivation (HM), price value (PV), and habit (HT). In this updated model, both habit and facilitating conditions are thought to load onto both behavioural intention as well as use behaviour directly. In a recent meta-analysis, Dwivedi et al. [25] found strong support in many of the constructs impacting both behaviour and use and reported on common extensions to the core model.

Figure 8 shows the UTAUT2 model structure, together with the mediating factors of age, gender, and experience acting on the various relations.

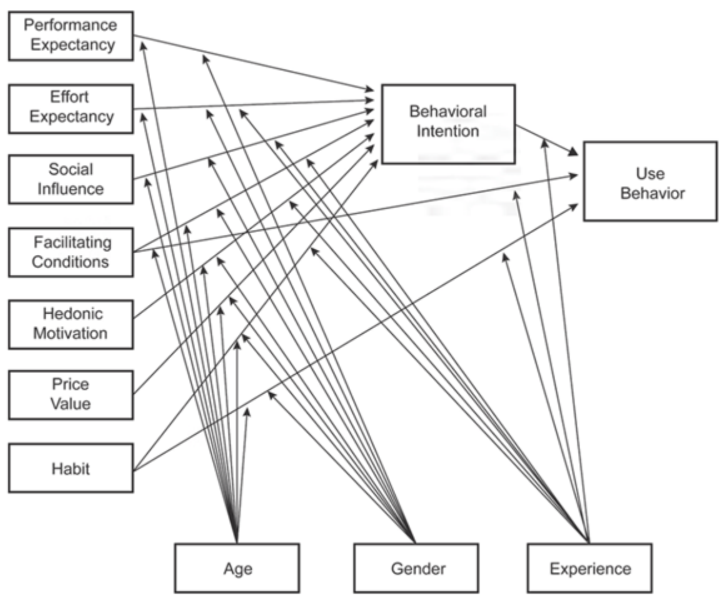


Figure 8: Unified Theory of Acceptance and Use of Technology 2 Model [124]

2.6 Summary

This chapter has investigated the development and implementation of systems that support users in constructing their own versions of AR. The applications all feature world-locked visualisations presented in a way that allows the user to design, configure, or otherwise manage their presence and functions in this new reality. They vary considerably in their complexity and ambition, but all offer an opportunity for a user to learn something new.

A notable limitation of the review was the absence of consideration for meta-cognitive knowledge in the papers identified. Investigating the impact of the AR experience on meta-cognitive knowledge in general may provide clues on how to enhance this in more structured learning environments. Given the wider theoretical basis of this work, however, such meta-cognitive knowledge needs to be assessed alongside meta-constructive practice. In this framing, rather than simply aiming for better mental models, we work towards creating better opportunities for learning, a process that includes both self-referential learning and a deepening connection between the learner and their environment. This topic is revisited in Chapter 7.

There are already signs that immersion can also impact learning outcomes, and vice versa [34]. In the course of this review, three other areas were identified as being ripe for further study: standardised testing, a common language for augmented experiences, and ethics.

Despite the prevalence of common features among the authoring toolkits described here, there is little evidence of standardisation around how to test these systems. While many opt for the "quick and dirty" System Usability Scale put forward by Brooke, the manner in which the questions are framed is usually not reported, and there is little in the way of statistical analysis, in part because the sample sizes are often small (10–30 people). The risk here is that a subjective notion of usability can be interpreted as an objective measure of value. More than an omission or fault, the issue is that SUS is simply not designed to provide this kind of measure. It is, however, designed to determine how well a system meets certain criteria, so it is well-placed to consider competing iterations and supplier methods where these have the same desired outcome [23].

Standardisation is useful when discussing the more general and abstract areas of study of both augmented experience capture and human performance augmentation, and there is a need for common terminology to describe the immersive experience. Amongst developers of AR authoring systems, there is not yet any such framework, meaning that knowledge constructed in one context is subject to interpretation when brought to another. Efforts have been made in many of the review papers to provide rationale and structure to our interactions, but none have found universal support.

Finally, there are ethical questions that often arise with the use of AR technology. Wearable or hand-held systems both demonstrate the ability to track the movement, state, and communications of both the user and those around them and, with a connection to the internet, can be used with other big data stores to generate correlations, insights, and other metrics relating to the physical or psychological states, with potentially invasive consequences. While research such as this is carried out under ethical oversight, a shift to greater commercial application brings new challenges when ensuring that privacy, both for the individual and the organisation, is properly maintained.

Chapter 3: Methodology

The previous chapter observed a trend of increasing complexity in AR training solutions. The step change encountered with the advent of the head-mounted form factor opens many avenues for closer inspection of the training process. An equivalent shift in the hardware's capabilities, demonstrated first by Microsoft's HoloLens, connects these to software design tools that are rapidly accelerating the options available to developers. With such a wealth of opportunity, identifying structured, reliable approaches to system design is of the utmost importance.

This thesis posits that there are a range of affordances that can be made available to a user of AR and WT that can improve their performance of complex tasks. It suggests that, with careful and systematic implementation, these affordances will be identified, acted on and, besides boosting their ability to perform the task at hand, will positively impact their attitude towards the technology that is mediating their experience. The process by which these assertions are developed and tested is the subject of this chapter.

The general theoretical framework is constructivist and draws on insights from Luhmann's radical, operative systems theory, described in more detail in Chapter 7, as well as the notion of affordances as meaningful structures in the operationalisation of pedagogical tools, introduced in Chapter 1. While there is not a complete departure from ontology, which Luhmann would insist upon, this thesis explores how the overlapping biological, cognitive, and communicative systems are affected by the inclusion of immersive technology.

Taking such a systems-oriented approach necessarily entails cross-disciplinary work and, as a result, requires the appreciation of methodologies and methods in different domains. While much of the technical work involved a pragmatic, iterative approach to problem-solving, the theoretical models represent a slow-changing basis for understanding, developed through careful analysis and verification.

For illustration, the ECS software, described in Chapter 4, was built following the standard software engineering design-develop-test cycle, where the testing group evolved from developers to project partners to end users and feedback was incorporated into the system at every stage. The wearable training system as a whole, however, was subject to a broader, five-phase instructional design model, incorporating activities related to analysis, design, development, implementation, and evaluation, each of which is deserving of some introduction.

The analysis phase was made easier by the fact that the work was carried out as part of the EU-funded WEKIT project, which had a set of broad, predetermined objectives [60], all in support of a central goal, which was to construct a system that could capture expert knowledge and share it with trainees. These project requirements, together with inputs from literature and the Requirements Bazaar situated the work that came after.

There was involvement of end users in the design phase, through a participatory design process that gathered insights from a range of user groups, and in the evaluation phase, which informed successive iterations of both technology and models. When developing wearable components and their interface with the ECS, the methodology used followed prototyping principles. Sensor selection was done early on, together with key architectural and protocol choices. The exact capabilities and performance of the communication channels for AR were, however, less well-known, meaning the development itself needed to be adaptive to both technological constraints. This is exemplified in the shift to direct communication between the Myo armband and HoloLens.

The implementation strategy was based on industrial use cases. The requirements were set by the needs and goals of each of the situated work flows. This meant complete adaptation to the workplace, together with any facilities and constraints it offered. Finally, the evaluative process was guided by a structural equation modeling (SEM) research design, which follows a system of hypothesis generation and testing, making use of a range of analytical techniques, such as confirmatory factor analysis and measures of model misspecification and fit. Rather than draw out these individual categories further, the next pages describe the research in terms of its key high-level activities: those of software engineering, wearable sensor integration, and trainee evaluation. Each will be presented in more detail, showing how it contributed to the validity and reliability of the work.

3.1 Software Engineering

The project work surrounding the construction of the software provided many subtle and intangible benefits, up to and including its use by hundreds of people in user trials. The structure of the methodology presented here should be prefaced with this acknowledgement, perhaps more so because this thesis explicitly recognises the social environment that all work is contingent on. It is only through the cooperative construction of pyramids of contingent, aligned activities that such complex efforts can be undertaken at all.

Requirements Engineering

Each of the topics of requirements engineering was addressed through interaction with expert trainers as well as other members of the industry partners and collaboration within the technical teams dedicated to the delivery of the ECS, as identified by Zave [140]. These include the identification of learning objectives, the identification of problems to solve in reaching these goals (from an implementation perspective), the development of technical solutions addressing these problems, and the formalisation of these strategies into knowledge contributions, all within an immersive technology framework.

The first phases used a requirements bazaar [93], which is an open structure where experts can leave information about given topics, which is then discussed and condensed into a set of specific needs. Informally, the requirements bazaar was used to stimulate further discussions with the trainers, exploring the details of a particular teaching practice. It was important to receive information about the didactic framework of the trainers in each of the areas of study: sonography, aeronautical engineering, and astronautic maintenance.

Another input came from the Augmented Reality Learning Experience Model (ARLEM) [134], since this was the format of the output from the ECS. Every step in an activity needed to be represented fully in this way, which consisted of storing the information in either an activity or workplace data structure. In this way, the training procedures were easily transferable between devices and, where appropriate, between locations. The final input sat between the first two and was an assessment of the technological capabilities of the AR devices, in particular the head-mounted display—a first-generation HoloLens. While this was a more fluid set of constraints, it was nonetheless essential to have a clear understanding of which features could be authentically recreated for the trainee, considering factors such as realism, fidelity, and performance.

Software Patterns

When designing software, patterns are the building blocks and joints that constitute its architecture [15]. In this case, they were assembled so that they could serve three purposes: allowing trainers to augment their work flow, managing and maintaining the compositional layer, and compiling this layer as an ARLEM record of both activity and workplace (for more details on this standard, see Appendix A).

Two approaches informed the general software architecture. The first is a central model-view-controller (MVC) pattern combined with feature-driven development for the peripheral functions. The MVC structure separates the code according to whether it handles communication with the data mode, in this case ARLEM, the visual elements, which were the collection of digital media in the workplace, and the user interface (UI), built in three dimensions. Rather than a single controller, the software hosted a nested structure—one activity, many action steps, each with many instructions—that was reflected in the UI's menu structure and the number of controllers present.

Various tools and techniques were used in the construction of the experience capture system (ECS) described in Chapter 4. Diagrams depicting information flows, such as UML diagrams, class diagrams, or call graphs, were used in both the specification of new features and the inspection of those that were already integrated into the system.

User interactions were passed from activity to action step to instruction, picking up context along the way, resulting in the desired visual or sensory feedback (e.g. "recording has ended") and a fully formed ARLEM statement that could be added to the learning record. An event routing pattern was used for the former, while the data model would be populated in an index-referenced way, allowing steps to be edited. While the controllers would route the user interaction, the effects were driven by feature-specific scripts connected to the relevant teaching affordance.

3.2 Wearable Sensor Integration

Where the construction of custom hardware was involved, as seen in Appendix B, a prototyping methodology was used. This was informed by the requirements bazaar as well as focus groups, following a user-centred design approach. Other work, more concerned with the connections and integrations of data channels from off-the-shelf devices (see Section 5.2) with the AR headset, meant that, from a methodological perspective, the approach was similar to that of the ECS, albeit with additional constraints and needs arising from the wearable nature of the technology.

Overcoming challenges associated with robust, repeatable sensor placement is essential when using wearable devices [62]. Wearable motion tracking and pose estimation are achieved using direct (or forward) rather than inverse kinematics models, and so, while significant advances have been made in deducing pose from optically tracked elements, this work built models that directly transpose sensor data onto a kinematic chain, leading to a greater emphasis being placed on data throughput and validity. To support this, the software development process used several iteration cycles, involving regular, repeated testing often at the limits of the devices' capabilities.

The Myo armbands, which reported movement and muscle activation data, were subject to both device-level and user-level calibration, both of which were incorporated and, to some degree, automated by the software that connected them to the AR headset. Putting on the armbands or repositioning them during operation would require re-calibration, which involved the user standing in the "A-pose" for a few seconds, during which the reference position was set.

Given the need for regular calibration as well as robust, high-bandwidth data communication, the design methodology of the wearable sensor network was centred around the frequency and capacity of the data transmission requirements. This meant that the communication protocols used, consisting of connect-and-stream as well as publish-subscribe paradigms, drove the design of the supporting software. To illustrate, much of the logic supporting the capture of movement data was asynchronous, whereas the functionality for monitoring temperatures was event-driven.

3.3 User Trials and Evaluation

The user trials, in which participants would don the headset and sensor vest, were conducted at three locations, each run under the auspices of the industry partner. Participants were introduced to the technology and the study and given an overview of them in written form. They were also provided with information about this research and acknowledged their voluntary participation and freedom to revoke it at any time without needing to give a reason. Since all the participants were first-time AR users, they began with an interactive guide on adjusting the headset and the interaction methods ('air-tap' and 'bloom' gestures).

Now prepared, each person was instructed to load the WEKIT training program and begin the task. They were generally accompanied, but not guided, during the task, and support was available for technical malfunctions or problems that they encountered. Having completed the procedure, they were asked to give their perspective on several topics, among them their view of the technology they used. A questionnaire consisting of roughly 20 items, each on a scale from 1 to 7, provided data for the study, which was ultimately conducted with over 400 people. Each participant's session was assigned a code, a copy of which was given to them, allowing all recorded data to be obfuscated on the day. The data was kept anonymously and securely from then on.

In studying the use of new technology, several established methods are available to the researcher, the most common being technology acceptance modelling, usability evaluation, and user experience evaluations, each also possessing a number of sub-methods. The choice of technology acceptance modeling was made for two reasons: that the basis for investigation came from empirical psychological research and that the process of extending the core model was both encouraged and well-defined.

The questions informed a technology acceptance model, which was an extension of the Unified Theory of Acceptance and Use of Technology, in its second iteration (UTAUT2) [124], adding constructs specifically geared towards the use of AR and WT. Over two iterations of modelling, analysis, optimisation, and inference, the insights from those data came to indicate general patterns and connections exhibited when these technologies were put to use. Convergent patterns, determined using confirmatory-factor analysis, were examined with various statistical tools. Those using perturbations, trends, and anomalies gave a picture of where the model fit well and where it did not. Modification indices, measures of model misspecification, and goodness-of-fit metrics were used, always in combination with theoretical rigour and clear reference to the variance that each modelling metric represents. Corrections for non-centrality and Monte Carlo cross-validation were among the tools used to ensure statistical validity.

Studies on technology acceptance require some adaptation, depending on the field of research, to account for a specific situation or application of knowledge. Augmented reality systems, which intermediate our natural experience of the world, are unique in both respects. This technology comprises a unique experience of blended real and virtual affordances and, in doing so, opens to us a world of undiscovered competence and investigation. Even from the perspective of computer science, the interaction modes and modalities are largely new, and the development of peripherals and design systems for interfaces is relatively nascent.

In order to effectively investigate attitudes and intentions towards AR and WT, several technology-specific constructs were proposed and tested. Some lines of inquiry offered little in the way of supporting evidence; others offered clearer relevance to core constructs and showed strong degrees of covariance with them. At each stage, the SEM models were used as informational tools rather than objective measures of success. After all, a better fit between model and observation may allow certain claims to be made with greater certainty, but unless those claims are useful to practitioners, they hold little value. Thus, the aim of this chapter is to present those models and the elements within them, as well as demonstrate why these results have practical utility in terms of sense- and decision-making, in line with a broader constructivist view of knowledge-building.

The findings from the acceptance model, together with insights from the literature review and the experimental work, can then be shown alongside one another in light of the theoretical frameworks for sense-making, communication, and learning with AR and WT. At all stages, there were both new and common hurdles, perhaps derived from the juxtaposition of new and common activities: using immersive technology to help us learn from each other.

Several R packages were used in the collation of results within the RStudio environment (version 2021.09). For modelling and analysis, the *lavaan* (0.6–9) and *psych* (2.1.9) packages were used; *knitr* (1.36) helped in tabling the data; and *semPlot* (1.1.2) and *ggplot2* (3.3.5) provided additional graphing functions.

Two Omissions

There are two elements that are present in the core UTAUT2 model that are not investigated here. The first is the 'Price Value' construct, which has not been included in the questionnaire. The construct represents the balance between perceived benefits of the system and the costs of using them, from a consumer's point of view [120]. In this context, this is problematic because, at the time of the study, the devices used were not available to consumers, only organisations. It was anticipated and confirmed by experience that the cost of the head-mounted device (Microsoft's HoloLens) was not generally known, and so such a cost-benefit estimation could not be reliably made.

The second omission is the use of mediating factors, such as age, experience, and gender. This is following the tradition of models prior to UTAUT2, such as TRA, TPB, and TAM, which consider such mediators to be relevant only when individuals share a context [26]. Since the focus of this work is to find cross-contextual benefits of the technology, their inclusion in the model is not thought to provide useful insight, appearing at the cost of increased complexity.

3.4 Summary

This chapter defines the types of research activities that inform each area of development. An overall picture is an engineering strategy involving the specification, construction, and evaluation of a technical system. Within an overarching instructional design approach, there was the implementation of requirement engineering, prototype design, and statistical evaluation. Learning objectives, specified early on, were translated to technical requirements and, later, software features. Evaluation of the outcomes relies on structural equation modelling of technology acceptance data collected from trials in three industrial use cases.

Chapter 4: Capturing Procedural Knowledge

Having outlined a methodology for building an immersive technology system that can take in expert knowledge and turn out competent trainees, we now look to the practices and processes that led to its construction. The Experience Capture System (ECS) described here allows a trainer to construct a digital work flow within a physical one. The tool allows trainers to generate a standardised description of both the procedure and the workplace while performing the task at hand. These descriptions are coherent with the Augmented Reality Learning Experience Model (ARLEM) standard [134], making the captured content transferable between devices and, where possible, between workplaces.

The activity model describes when and how the information is shown to the learner; the workplace model describes where it is placed and how it connects to other elements in the space. Within both models, there is a nested structure of attributes, breaking the activity into steps, each of which contains augmentation instructions. A detailed description of this model can be found in Appendix A.

This chapter focuses on propositional knowledge—often referred to as the "knowledge-that" aspect of learning—and the way in which the trainer can interact with AR technology to provide affordances to the learner that represent this knowledge. It describes how this knowledge, together with ARLEM, can then be used to specify the necessary functions of a system that can take advantage of the trainer's expertise as well as provide a description of their didactic choices in a standardised form. This was accomplished in three steps. First by establishing which instructions and affordances were needed, then by generalising these ideas, and finally by formulating technical solutions that would satisfy the general criteria for instruction provision, as described in Chapter 3.

For example, if an engineer says, "I would like to put a note on this panel, describing how to open it", this can be interpreted as 'attach text to surfaces', and the subsequent system function would allow the engineer to select a location and then create a note that would be displayed there. To the learner, the panel is now readable and, assuming the text is understood, manipulatable. Pointing out that the panel is open-able is, of course, only useful in certain situations and when the note is observed and understood (as relevant) by the learner; it is contingent on the technological affordances (e.g. the 3D model is displayed correctly), socio-cultural affordances (the note is in a readable language), and educational affordances (the wording is of an appropriate level of complexity).

In the process of developing the system, all of these perspectives had to be considered. Technological constraints may also be placed on the trainer when working with the tools, such as the ease with which either on-screen keyboard or dictation tools are used, which may influence the actual use of this teaching affordance - a noisy environment presents a significant challenge for speech recognition. The central device used was the HoloLens, though the principles described below could be applied to any head-mounted system that has SLAM functionality, wave-guide displays for the visual elements or a microphone and speaker system for the audio channels.

4.1 Informing the System

Throughout the duration of the wider WEKIT project, there were several contributors to the overall software system, which consisted of an experience recorder, a 'player' that would turn these recordings back into immersive activity flows, and a wearable component that not only tracked the body's movement and physiological state but also went through several design iterations, optimising for comfort and utility. The experience capture system and, in part, the construction of the wearable prototype were ultimately the responsibility of the author, but all the work was conducted in discussion with, and support from, all the technical partners involved in the WEKIT project, as mentioned in the acknowledgement section and in Section 2.3.

Earlier versions also demonstrated feasibility for many of the key aspects of information visualisation and sensor integration. Following the recommendations of Limbu et al. [67], this work develops the instructional design methods (IDMs) used, taking them as groupings of teaching affordances, and makes improvements in interaction design. This work presents both new and updated representations of the IDMs and demonstrates a more parsimonious code structure.

4.1.1 Workplace Procedures

The test beds for the ECS were in the fields of healthcare, aeronautics, and astronautics. As part of the WEKIT project, each partner designed a procedure that would be supported by immersive learning technology. Experts in each field were asked to design, with the use of the system, a work flow in which trainees would develop competence. Earlier trials had also been conducted with an earlier version of the software, prior to the construction of the recording functionality and modularisation of the task instructions. Limbu et al. [67] describe these procedures, which included pre-flight safety checks, echogram capture, and the installation of a stowage rack similar to those used on the International Space Station. Aside from the collaborative process of setting up their procedures, the results of which are given in Section 8.4, the industrial partners also gave feedback on the performance and usability of the system and assumed the role of beta-tester in later stages of development.

Despite the trials operating in different industries and, as a result, different surroundings and settings, there were several commonalities in the way the technology was introduced and tested. The fact that it was almost always the first time that people had experienced AR meant that the introductions to the software and the methods of interaction were of particular importance. Structured information was available in the form of the HoloLens' own guide to interaction, and we also provided simple goal-oriented challenges (such as operating menus) to boost the participants' familiarity with both the visual and interactive aspects of the technology.

Once there was a reasonable degree of competence using the device, the demonstration shifted to a presentation of the WEKIT software, which would guide them through the procedure. They were advised to pay attention to any and all augmentations and encouraged to explore the augmented space for clues, should they need guidance. Having donned the sensor vest and AR headset, the participants were instructed to begin the AR activity and follow the instructions given. Once the procedure was completed, the participants would remove the wearable devices and be guided to the feedback stage.

At every stage, from initial communication to the storage of data records, the activities were framed by ethical considerations. The participants' voluntary inclusion and ability to withdraw at any point was made clear on their arrival at the trial location and prior to the procedure. Their comfort when wearing either the vest or headset was also a central concern, and the teams looking after the participants were aware of common sources of discomfort or difficulty when using the system.

EBIT: An Ultrasound Scan of the Carotid Artery

In the first case, the participants were asked to perform an examination of a patient's carotid artery, also known as a Quality Intima-Media Thickness (QIMT) examination. In this procedure, a sonographer images the patient's artery using a hand-held probe, which can precisely measure the structure of the subcutaneous anatomy. Trainees in this trial would take a measurement of the diameter and wall thickness of the carotid artery. To do this, they must interact with various parts of the ultrasound (US) equipment and use the probe correctly on the patient.

Figure 9: Esaote MyLab Ultrasound Equipment

The ECS was to be used on three different models of US equipment (Esaote MyLab8 and MyLab9, and Philips). Each has a keyboard, trackball mouse, screen, and probe. The display is on a movable arm and has a range of movement of around a meter. The probe is plugged into the machine and has a cable length of approximately 2 meters. The operator sits in front of the machine, and the patient is lying (supine) on a bed or trolley on the right-hand side of the machine, within easy reach of the sonographer. This arrangement is shown in Figure 9. In considering a range of possible procedures, a set of five common requirements was provided by Ebit. When using the system, these general competencies include how to: 1) Position the patient correctly, 2) Choose the right probe, 3) Configure the US equipment, 4) Correctly position the probe on the patient, 5) Move the probe so as to properly view the anatomical structure (of the artery) and 6) Take measurements using the equipment based on the displayed images. For the current procedure, the ways in which learners were guided to achieve success were discussed, resulting in the set of corresponding augmentations shown in Table 7, along with a summary of the action steps.

Table 7: Medical Sonography Trial: Actions and Augmentations

Interaction	**Action required**	**Possible augmentations**
Self	Position yourself in a certain location relative to the patient and machine	Audio instruction, location marker, identify key locations
Probe	Identify correct probe	Audio instruction, 3D model of probe
	Proper grip and arm position	Audio instruction, 3D hand model
Patient	Position them in a certain way	Audio instruction, (in)correct images, 3D patient model
Screen, panel	Configure and start session	text note, audio instruction, keyboard location marker
Probe, patient	Position probe correctly on patient	location marker on patient, video instruction, audio instruction
Probe, patient, screen	Locate and centre carotid artery on screen	audio instruction, images with common mistakes
	Rotate probe to display artery longitudinally	audio, video, images, text overlay
Probe, patient, screen, panel	When screen overlays match the artery walls, acquire an image	audio instruction, image goal, images of common mistakes, location marker on image acquisition button,
	When QIMT's SD $<$ 20, freeze the image	audio instruction, location marker on FREEZE button

LuftTransport: Replacing Nose Wheel on Landing Gear

The next case relates to aeronautics and a procedure to change the nose wheel of a small aircraft—a Beechcraft B200 used as an air ambulance across the north of Norway. As part of the regular maintenance procedures, the nose wheel must be removed from the axle and a new one installed.

Figure 10: LuftTransport hangar and air ambulance aircraft.

The procedure took place in a hangar (seen in Figure 10, with the B200 in the lower right corner) that services a number of aircraft and was done under normal working conditions, meaning the frequent arrival and departure of vehicles from the hanger. The nose wheel assembly had several features or parts with which the trainee would need to interact. Removal of the cotter pin, axle nut, tabbed washer and spacer would allow the wheel to be removed, exposing the axle, bearing cups and bearing seals. There is an inspection of all parts for damage and wear and the application of aviation grease to some. Installation of the new wheel also made use of a (calibrated) torque wrench. The key competences expressed in this procedure were: 1) an understanding of the assembly, including its parts and their communication; 2) greasing certain parts correctly; and 3) inspection of parts for damage and wear. The augmentations related to this procedure's actions are shown in Table 8.

Table 8: Aeronautic Engineering Trial: Actions and Augmentations

Interaction	Action required	Possible augmentations
Removal Procedure		
Self	Walk to the nose landing gear area, avoiding the wing and any obstacles	Audio instruction, location target, visualisation of pathway
Wheel assembly	Identify (nose landing gear) wheel assembly	Video instruction
	Familiarise yourself with the parts	image (schematic), 3D model of assembly, with text annotated parts
Cotter pin, axle nut, tabbed washer, spacer	In the correct order, carefully remove these parts	audio instruction, video instruction from the trainer's perspective, 3D models of individual parts
Outer grease seal, bearing cone, wheel, axle	Remove parts from axle	audio instruction, warning - caution for wheel removal, physical demonstration of handling technique
Inner bearing cone, grease seal, axle	Remove remaining parts	Image of result, video of action, text information to show completion
Installation Procedure		
Axle, bearing cup, grease	Lubricate axle and bearing cup	location marker, image of grease quantity
	Check for damage/wear on axle and nut	audio instruction, images of common faults
grease seal, bearing cup, axle,	Properly seat the parts on the axle	audio instruction, image of correct outcome
Wheel, axle	Install wheel	audio instruction, video instruction from the trainer's perspective, animated models of parts moving into place, advice on supporting the wheel
Outboard bearing and grease seal	Fit parts onto axle	audio instruction, 3D models of parts, highlighting text feature, image of correct position, text description of rationale
Spacer and tabbed washer	Fit parts onto axle, aligning tabs	audio instruction, images of correct position, warning/caution: check bearing seal
Nut, axle	wind nut onto axle, finger tighten	clockwise motion indicator
Nut, torque wrench, wheel	Tighten to correct force while rotating wheel	Video of the torque wrench being used correctly to tighten the axle nut, mark hand positions, clockwise motion indicator
Cotter pin, nut, torque wrench	align parts, fine tune tightening if required	video of small adjustment, video of pin installation, 3D models of parts

Altec: Inspecting Mars Rover Solar Panels for Damage and Connect Battery

Interest in planetary exploration has seen renewed interest in recent years. Immersive technology's potential for just-in-time delivery of information is of particular interest in astronaut training, a process that takes many years and requires highly specialised problem solving. In anticipation of a human presence on Mars, the ExoMars Rover was selected as the focus of the third workplace, pictured in Figure 11.

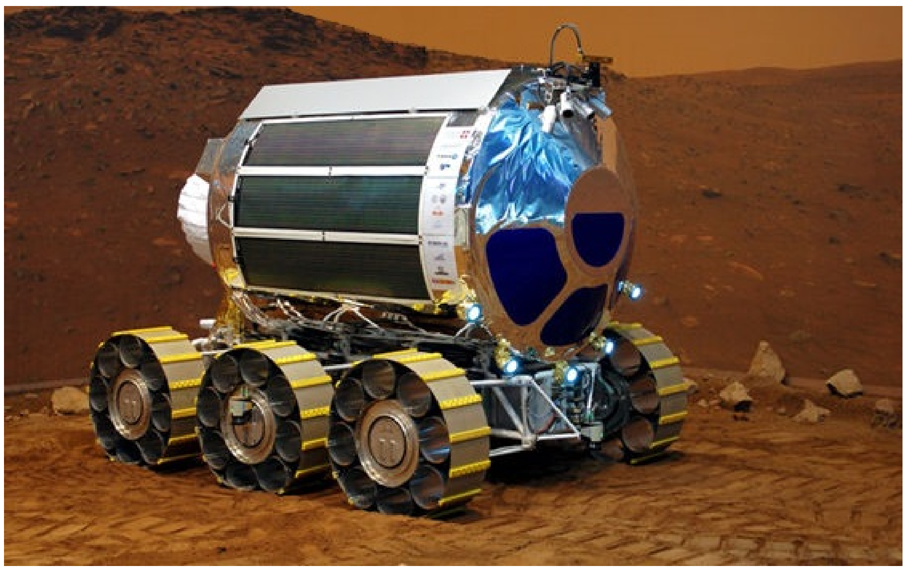


Figure 11: Altec Workplace

In this trial, the training environment is the Altec's Mars Terrain Simulator, which is a rocky surface modelled, in both geometry and composition, after the surface of the Red Planet. A scaled-down, though functional, version of the ExoMars Rover is the focus of the activity, and the trainee is expected to interact with various parts of it, as well as a portable battery charger housed in a protective case. This procedure demonstrates a relatively straight-forward but nonetheless essential example of extraterrestrial guidance: a visual inspection of the state of the solar panels and the connection of a battery.

Three main competences arose from the formation of this procedure: 1) careful and deliberate interaction with the rover; 2) proper visual inspection of solar panels; and 3) the correct use of the battery charger. As with the previous situations, the environmental feature, action step, and related augmentations are shown in Table 9.

Table 9: Astronautic Engineering Trial: Actions and Augmentations

Interaction	**Action required**	**Possible augmentations**
Self	Walk to the front-right side of the rover, avoiding any obstacles	Audio instruction, location target, visualisation of pathway, image of coordinate system
Control panel	Locate control panel and remove red key	Audio instruction, location marker for panel, show trainer arm position for removing red key, image of key
Emergency stop button	Find button and press it	Audio instruction, location marker for button, view of trainer's hand position
Solar panels	Perform visual inspection of both left and right side panels	Audio instruction, images of common faults
Battery charger device box	Find a retrieve box	Audio instruction, location of box, path to box
Charger, battery charge connector	Carrying charger, locate connector on rover	location target, location marker, guiding path
Charger, rover wheel	Place charger on nearby wheel tread	Location marker, video showing charger position on wheel
Charger, rover connector	(Dis)connect charger to rover	Audio instruction, connection indicator, recording of trainer's (arm) position
Charger	Press MODE to start/stop charging	Text note indicating state
	(Un)pack charger	(Un)packing indicator icon, video of battery charger handling
Emergency stop button	Release button	Audio instruction, location marker for button, indicator for pressing and clockwise rotation of button, view of trainer's hand position
Control panel	Locate control panel, insert red key	Audio instruction, location marker for panel, show trainer arm position, image of key position

In addition to describing the procedures, the trial partners were also involved in a participatory design process focusing on the tasks to be trained, the inclusion of several wearable sensors, and the set-up of trials that would evaluate many aspects of the system, including technology acceptance. In these discussions, three top-level goals stood out as necessary features of an ECS. The system should be able to: 1) present all types of identified affordances, 2) give the trainer as much freedom and control as possible in building the activity, and 3) save the procedure and share it between devices and workplaces.

Having agreed on the various augmentations the procedures would require, the next task was to understand how they could most effectively be represented as actionable affordances. The process of drawing these out from the suggested augmentations was guided by the taxonomy discussed in Chapter 2—both the type of knowledge and the cognitive process were considered—with the goal of assembling an “affordance vocabulary”, describing a variety of routes to different learning objectives.

4.1.2 Identifying and Grouping Affordances

When providing supporting information for learners engaged in the tasks described, seven areas were identified that could be realised through the coordination of technical development and trainer input:

Augmented Path This is the visualisation of a marker that directs the learner to move to a particular location or to follow a path to reach a location. The visual marker or pathway (or a combination of both) can be placed in the air or on the ground. Many tasks and activities begin with such a step, and its utility was quickly identified. From the trainer's perspective, this orients the learner in anticipation of interacting with tools or equipment; the learning affordance here is in making things accessible, though the purpose may remain hidden.

Directed Focus This IDM causes the learner to give their focus to a specific location. In an AR context, this location may or may not be in the field of view of the learner, so the visual element must take this into account. Near (easily spotted) objects must not be obscured by the visualisation intended to highlight them. Distant objects (or those well outside the view of the learner) require that the user turn or move to discover them, and so an indicator (visual or audible) can be used to support this.

Point-of-view Video As an extension to the standard practice of recording tasks with video, the perspective of the camera can play an important role in guiding the learner. When they see an action from this perspective, it tends to promote imitation and carries subtle information such as posture, speed of movement, and area of focus. To illustrate, if a video shows the trainer looking first at the patient, then towards the screen, then again at the patient, the learner will be inclined to follow the same sequence.

Think-aloud protocol Another IDM laden with potential is the think-aloud protocol. This is the process by which a trainer will record an audio clip while performing the task. The tendency to vocalise actions while they happen has been suggested as an aid in both pedagogy and teacher training and, in the context of procedural learning, is thought to have broad use. In many of the practical and theoretical discussions, it was identified as an IDM that would be used alongside others. For example, a picture may be shown and a voice heard saying, “I'm turning the part through 90 degrees, so that it ends up like this “.

Annotations These are pieces of media content (text, labels, graphics, audio, and video) that are anchored to the physical environment (a spatial map). Their affordances are offered to the learner by the content of the message. Given that the annotations are world-locked (though this may apply to something that is movable, such as an aeroplane), we can say that they lead to stronger contextualisation of things and places, which we could describe as interpretable to the learner.

3D Models Although technically a type of annotation, models can contain such a wealth of information that they are given their own category. Special features include: - models can be seen from different directions, and the directionality is important. - they possess a greater inherent quality of interactivity because of their volume. - when they are animated, they can provide temporal information. - 3D modelling has long been used to create digital twins, but when they can be overlaid or used in context, they enrich physical artefacts.

Ghost Track A record of the pose (position and orientation) of the trainer's hands and/or body in a workplace reference frame. This IDM allows the movement of the trainer to be stored as a sequence of coordinates, which can then be replayed, in some visual form, to the learner. As with directed focus, the trainer can use the expressions 'here' and 'there' rather than attempt to describe a particular location. Instructions such as "from the front of the vehicle, look along the edge to make sure the wires are aligned" are more quickly understood when the learner can identify which edge and wires are involved.

Table 10: Summary of instruction design methods.

IDM	Description	Learning Affordances
Augmented Path	A location or route to a location is marked out	From this point certain things are reachable, accessible or visible.
Directed Focus	A trainer can indicate a location of interest	Highlighted locations are more noticeable
Point-of-View Video	A video is recorder during task practice, from the HMD	Hand actions are imitable, outcomes are reproducible.
Think-aloud Protocol	The trainer reports (orally) what they are doing while they are doing it	Certain elements of practice are succinctly (or, sometimes, exclusively) communicable in words, these are now expressible.
Annotations	Visual elements (text, images, icons, glyphs) are anchored to the environment	Annotations convey additional meaning about something in the environment. The affordance leveraged by using this IDM will depend on the annotation used. The location can be described, in any case, as interpretable.

(Animated) 3D Models	A special class of annotations that have volume and may be animated.	These objects are all, to some degree, interactable, as the learner can walk around them. They have the possibility for further enhancement if they are animated, have an observable structure or direct user interaction.
Ghost Track	Visualisation of gaze direction and body movement, based on device localisation	Gaze (head direction) and, to some extent, body position are imitable by the learner.

The instructional design methods used in the system, summarised in Table 16, provide helpful guidance in terms of functionality and code design but, in order to be useful as a data model, also require a formal and standardised description of the actions that take place in the activity. Furthermore, the procedure's functional environment (the workplace) needs a corresponding description in order to connect the actions of the delivery of the relevant digital procedural content.

4.2 An Experience Capture System

Earlier work by various members of the WEKIT consortium, especially that done at the Open University of the Netherlands (OUNL), was invaluable in highlighting some of the immediate challenges associated with the construction of an ECS. Indeed, the software under construction was informed by several successful implementations of training systems. The most notable of these are the TELLME project [133], which looked at practical implementations of AR in the workplace, and ARgh! [48], which sought new forms of human-machine systems.

The IDMs presented describe, in large part, what needs to be communicated. The ARLEM activity and workplace models provide the answer to where, in terms of a data model, to store this data. The purpose of this section is to describe the various structured and information flows that allow the first to be translated into the second and, in doing so, show how the programmatic links between them can be exposed, to provide clarity and opportunity for researchers wishing to take advantage of the system, highlighting the design choices and showing how the system design was supported and constrained.

4.2.1 System Functionality

An overview of the software functionality is seen in Figure 12, which shows how each selection is dependent on previously, correctly performed selections. For example, if a trainer wishes to add a new action within a procedure, they can achieve this by logging in, then starting a new activity, then adding an action step, then creating an instruction, and finally recording or attaching digital content to this instruction.

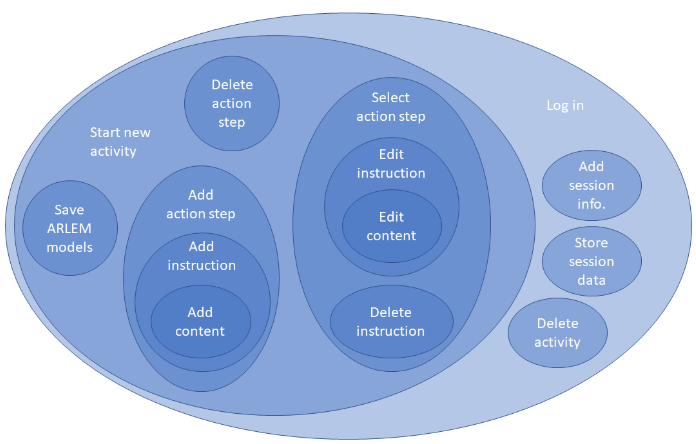


Figure 12: Sofware functionality overview: darker areas require more dependent steps and are potentially more difficult and prone to errors.

The result of error-free operation is that content is added to the procedure through a nested structure, with layers represented, in descending order, by activity, then action step, and finally instructions, which contain the raw data. Each procedure (activity) contains one or more action steps, each having one or more instructions (equivalent to IDMs), each of which is given media content, a pose (position and orientation), and other meta-data.

Given the specificity of the data model and the visual significance of an AR training system, a model-view-controller (MVC) architecture was found to offer the most useful overall guide for development, as it allowed the construction of visual components and the underlying code framework to be tackled in parallel. A generic representation of container objects within the data model could be used and their visualisations could be updated without the need for changes to the code. Since this work was part of a larger project, this also helped to streamline and scope the development. The standard MVC structure was, however, extended to reflect the different levels of interaction and data representation in the ARLEM data model. This layout is shown in Figure 13.

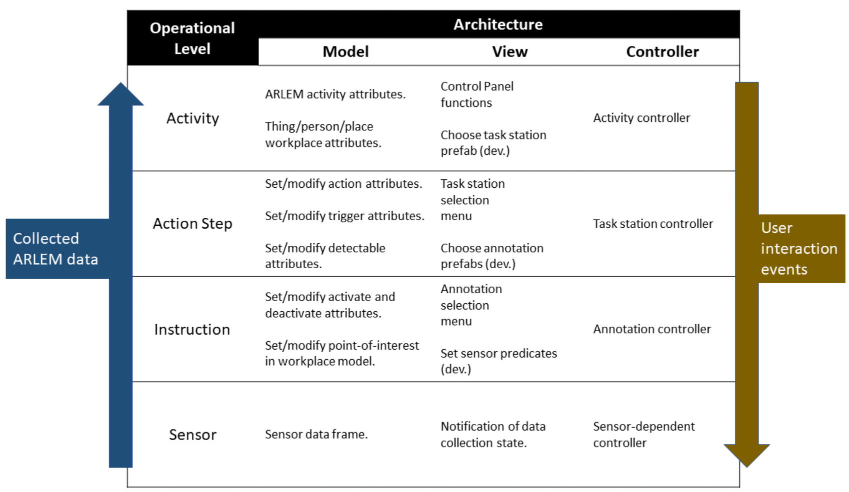


Figure 13: Model-view-controller operational levels.

Generally speaking, the software worked to support two main processes: a downward cascade of user events through the operational levels and the upward propagation of sensor-driven information, populating a record of each work flow as it is captured.

The core data model comprised the ARLEM activity and workplace descriptions, though additional data frames were used to capture sensor data, either from the HoloLens, where position and orientation of the head and hand were stored, or from other wearable sensors connected to the system. These are described in detail in the next chapter.

The controllers of each operational level would, at each level, do three things:

1) configure the elements on that level; 2) manage the creation and order of elements on the level below; and 3) serialise its data model representation and pass this to the level above. There was a single activity controller that coordinated the entire workflow, one task station controller for each action step, and an instruction controller for each augmentation connected to the action step. One or more sensor streams could be attached to an augmentation.

The view component addressed the presentation of information. Given the complexity of the system and the need for scalability, each user interface (UI) element also had a corresponding interface within Unity, so that the reconfiguration of each level could be achieved using the game engine's UI.

Activity Level

Upon loading the application, a control panel was shown that could be used to name, save or exit the activity (pictured in Figure 14). In addition, the status of any connected wearable components was shown. Their connection and utility are the subject of Chapter 5. The control panel could be moved around, pinned to a wall in the workplace, or summoned with the voice command "move panel", which would position the control panel in front of the user.

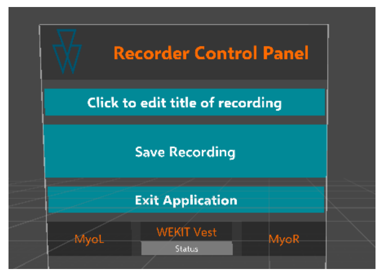


Figure 14: ECS control panel.

This control panel represents the visual component of the activity and was separate from its controller which, on loading, simply awaited user interaction with the spatial map. On interaction, a new task station would be created, both in the visual space and the data model. Following ARLEM, each new task station was furnished with a 'place', 'action', and 'detectable' data element that would track corresponding changes.

Action Step Level

The visual representation of an action step, a 'task station', was a sphere of approximately 5 centimetres in diameter. This was chosen for its geometric symmetry and the consequent ease with which it could be placed sensibly in the environment. Attention was given to the way in which information related to each task station was shown and hidden. A single task station could be active, displaying its contents, while others would be hidden, allowing for better focus on the task at hand and preventing the workplace from

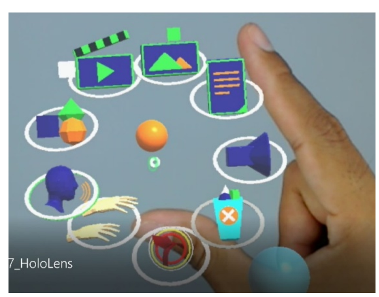


Figure 15: Task station menu.

becoming cluttered, which was especially important when working in more confined spaces.

On interaction, a task station would present the user with a set of available instructions, as shown in Figure 15. In the figure, eight types of instructions are shown, together with a delete function that could be used to remove the action step from the procedure. The placement of the task station menu was the result of an initial target for radial distance, combined with constraints given by the environment. As is the case for any AR UI, the exact surface geometry on which the object is anchored is unknown. This uncertainty was remedied by first allowing the menu to move outwards from the centre point, which was already offset from a surface, and then allowing the objects in the menu to change direction if they came into contact with a surface during their motion.

The target menu diameter was configurable in Unity, along with its colour and magnification when highlighted. These variables are held by the 'task station controller', which also configured the previous generated 'place' and 'detectable' ARLEM elements. Changes to these elements would be monitored and, when the task station was deactivated, would be passed to the activity controller to update the main data model. A separate script, the 'task station context menu' allowed reconfiguration of the menu items in a simple and extendible way, using a list of ready-made prototypes, of which the menu was one instantiation. The options for both the controller and menu can be seen in Figure 16. Every task station created by the trainer has its own instance of controller and associated context menu.

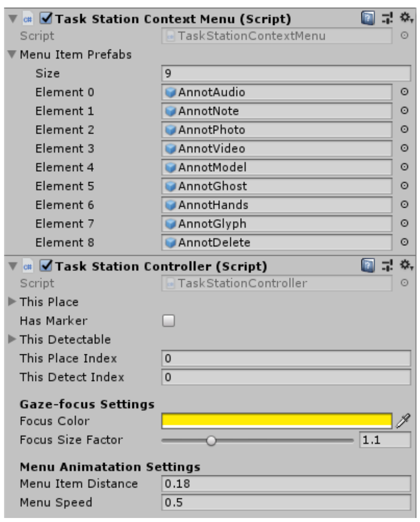


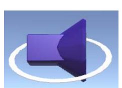
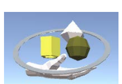
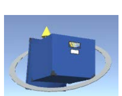
Figure 16: Task station controller and context menu options in Unity.

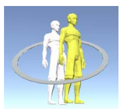
Instruction Level

Unlike the task stations, the instructions are draggable elements, meaning that, once created, they can be moved around by the trainer. Every object in this level shares a common data controller that allows for the creation of menus in the user environment. Aside from the instruction identifier, two attributes are common to all the instructions and set by the controller, namely position and rotation, both relative to the parent task station. Whenever instructions are created or moved, their pose (position and rotation) is reported to that task station.

Instructions are also prototypical elements, and each has a visual component, a set of attributes mirroring those in the ARLEM standard, and a dedicated class containing functions that communicate with the sensor level. These elements are listed in Table 11, showing the attributes that each contributes to the overall data model and the key functions to which the controller points. Attribute values that are set during operation are identified with square brackets and default values are given, where applicable.

Table 11: Summary of instructions.

Totem	**ARLEM Attributes**	**Functions**
	Text Note .type= "tangible" .predicate= "label" .text=[user-defined value] .scale=[0.05]	Edit note Clear text
	Image Annotation .type= "tangible" .predicate= "image" .scale=[0.5]	Take a Photo Select image on disk
	Point-of-View Video .type= "tangible" .predicate= "label" .url=[link to recorded file] .scale=[0.5]	Start recording Stop recording Play video Pause video
	Think Aloud .type= "tangible" .predicate= "audio" .url=[link to recorded file] .scale=[0.5]	Start recording Stop recording Play audio Pause audio
	3D Model .type= "tangible" .predicate= "model" .url=[user-defined value] .option=[user-defined model name]	Load model from disk
	Glyph Annotation .type= "tangible" .predicate= [user defined e.g. "locate", "rotate"] .scale=[0.05]	Locate (target) Locate (pointer) Rotate clockwise Rotate anticlockwise ...

	Hand Recording .type= "tangible" .predicate= "hands" .sensor= "myo" .url=[link to recorded data]	Start recording hand position Stop recording hand position Play stored recording Pause playback
	Ghost Track Recording .type= "tangible" .predicate= "ghosttracks" .option=[name of ghost visualisation prefab] .url=[link to recorded data]	Start recording body pose Stop recording body pose Play stored recording Pause playback

An 'instruction controller', unique to each type of augmentation, had the purpose of adding content to the instruction, either by loading a media element or controlling the capture of a data stream or other user input event. While there was significant variation in the incoming data types, it was nonetheless possible to create a generic framework in the Unity environment that allowed the connection of new media types to the framework without the need for additional software development.

4.2.2 Component Architecture

The architecture supporting this functionality can also be characterised by its interfaces. Since the software meets requirements coming from various sources, it is helpful to see in which areas of the software these requirements are handled. Uploading and transferring activities to other devices clearly requires compatibility with web standards, and ARLEM provides the specification for these activities. Developers were able to specify and address specific instructions, each of which had a common link to the core controllers. The model also provides well-defined scope for UI designers when creating or updating the visual elements. These connections are shown in Figure 17.

The left column of the class diagram shows behaviours connected to each level (activity, action, and instruction). The functions they contain connect to user interaction events, such as saving the activity, using voice commands, or selecting menu options. Shown on the right are the classes used to provide instruction-specific functionality, some of which have manager classes handling the storage and display of content.

To visualise the menu selection process, the task station and instruction controllers inherit a menu class containing the relevant UI handler functions, the 'task station context menu' and 'instruction menu', respectively. Collaboration graphs of these components are shown in Figure 18 and 19. The task station level is more structural, as it orchestrates the various instructions, tracking and managing their changes, whereas each instruction has only a single object to work with. Instruction menu items are largely media functions, handling the recording and playback of data streams.

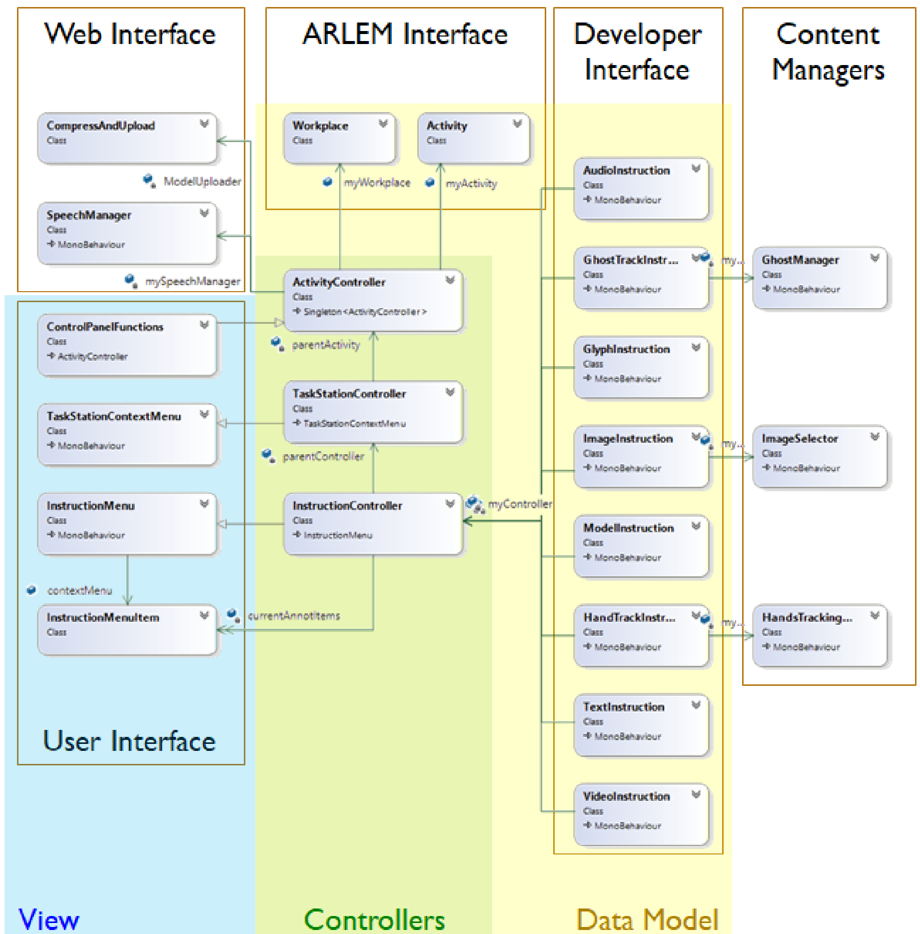


Figure 17: ECS class overview.

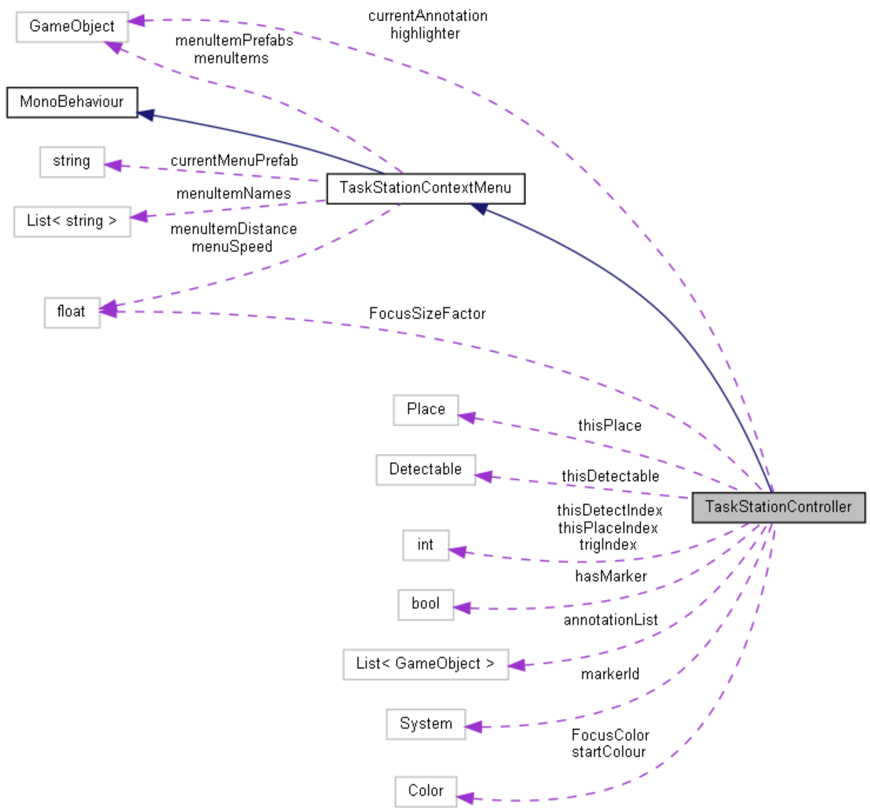


Figure 18: Task station menu collaboration graph.

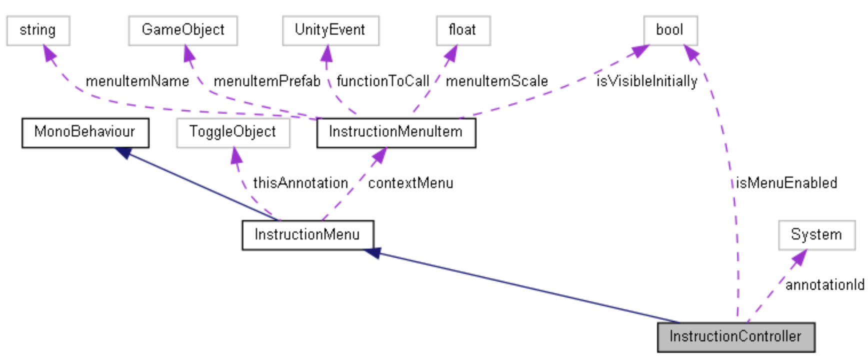


Figure 19: Instruction menu collaboration graph.

Each task station controller stores information drawn from the properties and relations seen in Figure 19, which is where the references to real-world positions are kept, along with all ARLEM-specific data and those supporting the smooth operation of the UI system. The 'Place' and 'Detectable' connected the task station to the workplace model and sensor data layer, respectively.

The annotation ID, held by the instruction controller, is the main identifier that connects the activity and workplace models. The 'ToggleObject', which represents ARLEM activate or deactivate statements, references this value, connecting it to the 'place' elements held by the task station controller.

4.2.3 Data Flow

The gestural interface of the HoloLens involved the use of an 'air-tap'—a d downward movement of the forefinger while in the visual field of the front-facing cameras of the device. This gesture could be quickly repeated (a kind of 'double-click'), distinguishing it from interaction using a single action. The double tap was used to create task stations on the spatial map, after which only single air taps were used.

The entry point for user interaction is in the activity controller, and events are triggered through the gaze pointer of the HoloLens. This is a projected ray from the centre of the device's display, tied to the movement of the head, which is visualised with a cursor that is projected onto the surface of the virtual content, including that of the spatial map. Events take into account both the surface on which the cursor sits as well as the type of interaction gesture performed by the user.

A call graph of the initial interaction logic is shown in Figure 20. Those functions marked in red indicate where further structure is not shown.

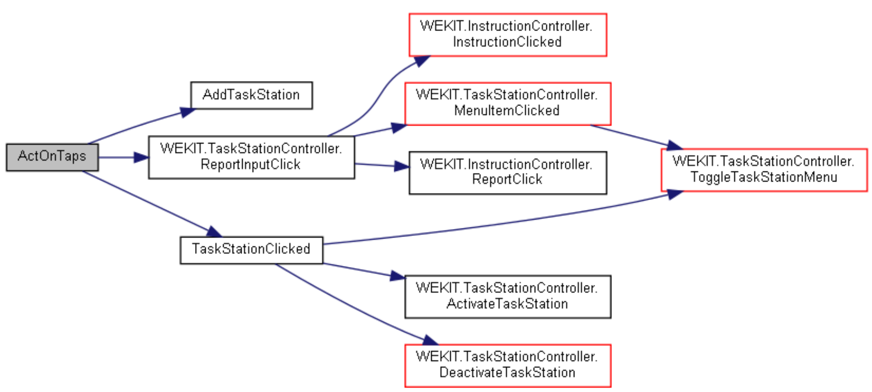


Figure 20: Interaction call graph

This is the simplified decision tree for the nested menu system; each menu item refers to a three-dimensional object aligned with its environment. The first decision point queries the location of the tap (user interaction), changing the activation state, toggling the menus, or adding task stations to the space. Where there is user interaction with instructions—the content containers—the interactions are routed through their respective task stations, allowing changes to be detected in an event-driven manner.

As the interaction events were serialised and passed between controllers, the management of the ARLEM data model happened as required. Where there were changes to the activity, such as renaming it, there was an activity-level update. Where an instruction was edited, the task station controller would have visibility of the changes present in each instruction while permitting parallel changes, such as recording audio and body position simultaneously.

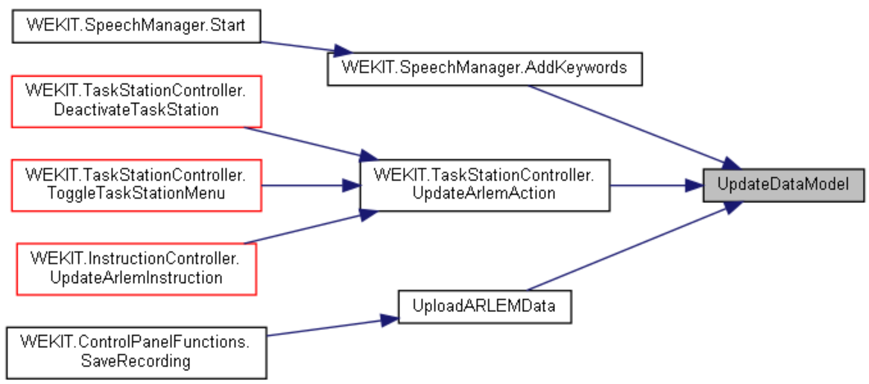


Figure 21: Caller graph for data model updates.

Figure 21 shows the event tree that leads to an update of the ARLEM workflow. The first, a voice command ("save recording"), would capture any changes that had already been reported to the activity controller. Deactivating the task station (by selecting another or deleting it) would also lead to an update, as would closing the task station's menu. Since the creation of a new instruction also prompts you to close the existing task station menu, this would also trigger an update to the data model. Updates to the instructions (by editing their content or removing them) would also update the task station's data model, in turn registering changes in the activity and workplace data objects. An update can then be triggered explicitly with a function linked to a button on the activity's control panel.

4.3 Summary

This chapter has described how propositional knowledge can be made available to learners using immersive technology. It has shown how procedural tasks can be broken down into meaningful units that can be operationalised with the use of IDMs, and how these can be formulated in a language that can be operationalised within a software architecture. There was consideration of web, user, ARLEM, and developer interfaces to this software, as was the flow of both interaction commands and information about the task and the augmented environment.

Competence, however, is not only built on propositional knowledge. The fact that an expert can perform a task to a higher standard is often the result of the time-dependent processes of repetition and reflection. Repeating tasks and being able to understand how and why the outcomes arose requires deliberate practice and the ability to infer how to better perform the task. To properly understand embodied procedural competence (or, more simply, performance), we must look at that practice not only from the point of view of factual acquisition but also from the point of view of physical involvement in the task.

Chapter 5: Measuring Embodied Performance

The experience capture system, in its ability to capture factual knowledge and use it to generate an augmented work flow, is focused on communicating *what* to do and *when* to do it. Procedural knowledge—information about *how* to perform a task—can also be presented in such a way, but a purely semantic description of this type of knowledge lacks the tangible, mimickable quality of watching someone complete that task; put another way, it ignores psycho-motor competence. This has been identified as a limitation in many AR authoring systems (see Chapter 2) and this chapter addresses this need by demonstrating how movement of the head, body, and arms can be measured, organised and used to inform an AR scene. It concludes with a summary of the development of the body sensor network and its connections to the ECS.

That we perform intricate or subtle tasks with our hands and make this skill reproducible in different contexts, while long understood to be part of what it means to demonstrate expertise, has not seen the same degree of investigation as propositional knowledge capture. This is largely the result of two factors. The first is that movement data, taken on its own, is not an especially good source of information; perhaps we can say with certainty that a person is leaning forward at an angle of twenty-one degrees, but unless we understand something about the task at hand, the physical environment, and also the outcome of the task, there is little that can be gleaned from this number. If, however, we notice that the person is leaning to avoid the wing or an aircraft or that, in this position, it is possible to align an ultrasound probe with the correct anatomical reference, the number has some meaning. The second reason why movement data (or rather, its analysis) is not more prevalent in the literature is that it is hard to capture experimentally, leading to unreliable results, throughput limitations or a low signal-to-noise ratio. Each of these challenges is addressed below, and strategies to overcome or mitigate these issues are described.

AR has extended the possible range of use cases for motion capture data. The mapping and localisation capabilities of AR HMDs are significant both technically and conceptually, not only in providing the basis for a semantic description of the workplace and the activities that take place there but also in setting a reliable anchor for localising and orienting other coordinate frames based on static or known features of the environment.

Triggers that occur in the work flow, such as interacting with the interface to determine the start or end of a particular action, can act as timestamps to segment otherwise continuous data. Similarly, notifications within an AR work flow may also be informed by movement data, for instance, to remind someone not to walk under the wing of an aircraft. Other triggers may be as simple as reminding someone to maintain a particular posture to minimise risk (e.g. safe lifting procedures) or as complex as guiding the person through a surgical procedure. Each has requirements in terms of reporting, sensitivity, and risk and would require careful treatment, but it is thought that, if used appropriately, all could benefit from motion- or position-based feedback.

Measuring movements from the head and torso is the topic of the first part of this chapter, and there is a description of the technique used to capture positional data from the HMD and an e-textile component built on a distributed microprocessor architecture. This body sensor network, described in more detail in Appendix B, uses the MQTT messaging protocol to communicate with the relevant part of the ECS, the so-called 'Ghost Track Instruction'. The second part of the chapter deals with the capture of data from the arms, which was captured from a body-oriented reference frame and used alongside the HMD's optical tracking system in order to maintain the persistence of the augmented appendages. The chapter also describes the collection of data via the Bluetooth Low Energy (BLE) protocol. In order to stream from several devices at once, a custom-built BLE adapter for the Universal Windows Platform (UWP) was written and, with this, both inertial measurement and surface electromyographic (sEMG) data were captured. Recorded data were visualised using the Hand Track instruction of the ECS.

5.1 Body Tracking

When authoring AR experiences, two types of reference point exist: one that is locked to the world and one that follows the person. Chapter 4 dealt with the first type. Here we are concerned with the second: visualising a virtual representation of the body's position in space with precision and reliability.

The first point of interest is the HMD, which is capable of tracking its position based on visual-inertial simultaneous localisation and mapping (VISLAM) algorithms [55]. These allow for a 3D position to be derived from features in the environment and, to some degree, internal sensors. The HoloLens was the first untethered device to achieve high precision in this regard, demonstrating approximately centimetre-scale accuracy across a room-scale area.

The spatial mapping produced by the HoloLens is a 3D point cloud, which is then run through a meshing algorithm that creates edges, triangles, and polygons, imitating the physical geometry of the environment. These maps are stored in the device's memory and can be re-loaded and oriented onto a space to support faster re-mapping and localisation. One limitation of the system is that, on starting, it would use the first available pose measurement as the origin for a local coordinate system and report new positions relative to this. For world-locked augmentations this can be problematic, though it can be mitigated by using a marker to realign the workspace, which depends only on the accuracy of marker placement. For body-relative augmentations we can work relative to the head-mounted camera position itself, somewhat simplifying things.

Having determined a workplace frame of reference, the Unity game engine easily allows us to extract device pose information and, by making use of the Mixed Reality Toolkit (MRTK), a software development kit produced as a companion to the HoloLens, we can track the location of the HMD.

With the device pose acquired, we looked at the orientation of the torso. Due to the flexibility of the spine, at least two measurement points are needed to approximate posture. Nine-axis IMUs were used, together with others that measure the physiological state of the wearer. A description of the hardware systems, built by WEKIT collaborators, can be found in Appendix B, which details the sensors used and describes the data channels that were made available. It also describes the MQTT communication protocol that these sensors use to communicate with the ECS.

5.1.1 Building the Ghost Track Instruction

The main requirement for integrating data from the body sensor network into the ECS is the creation of an instruction that records, stores and visualises the relevant information. In building this component, two common design problems were addressed. The first is the occlusion problem—that using holograms to indicate things can visually obstruct those same things one wishes to highlight. The solution opted for was to make the avatar semi-transparent. The second is the uncanny valley problem, where avatars who are simultaneously realistic and unrealistic conjure a feeling of "creepiness", which was avoided by using an unrealistic totem (3D icon). The absence of detail and translucent nature of the avatar soon made it known as the "ghost".

While most of the system described so far is agnostic to the type of HMD used, when retrieving specific sensor information, some device-specific code is needed. The HoloLens (first generation) device was supported by the Mixed Reality Toolkit and the Unity game engine. Using these tools, the 'HoloLens Sensor' class could easily record the camera position and gaze direction. The 'Hands Tracking Manager' detected the presence of the hands and provided a list of available positions, which were recorded alongside the head pose. Listing 1 shows the C# code used to organise this information.

```
public void UpdateSensorData()
{
    // Create new, empty data frame.
    currentHoloDataFrame = new HoloLensDataFrame();
    // Get wearer's head position
    currentHoloDataFrame.HeadPosition = Camera.main.transform.position;
    // Get wearer's gaze direction
    currentHoloDataFrame.GazeDirection = Camera.main.transform.forward;
    // Get hands' positions (if visible)
    var hands = myHandsTracker.handedHandPositions;
    foreach (var hand in hands) {
        if (hand.Key == UnityEngine.XR.WSA.Input.InteractionSourceHandedness.Left
             ) {
            currentHoloDataFrame.LeftHandPos = hand.Value;
        }
        if (hand.Key == UnityEngine.XR.WSA.Input.InteractionSourceHandedness.
             Right) {
            currentHoloDataFrame.RightHandPos = hand.Value;
        }
    }
    // Call base class to execute callback stack for listeners
    base.FixedFrameUpdate();
```

```
}
```

Listing 1: Updating the current HoloLens sensor record.

A sensor base class, inherited by each channel's content manager, ensured that the data throughput was constant using a fixed interval update pattern. It also specified a set of event delegates and provided current data to any event listeners. A graphical representation of this description is shown in Figure 22.

The head and hand positions of the ghost came from the sensor data of the HMD (the HoloLens) and the body posture from the e-textile. These data pipelines meet in the Ghost Track instruction, where they are stored as a set of records. On playback, these records are passed to the 'Ghost Manager' class, which handles the visualisation of the avatar.

The 'Vest Sensor' class creates and configures an instance of the MQTT connection. It also inherits the relevant data frame structure and callbacks that signify the arrival of new data. Just as with the HoloLens data frame, the e-textile sensor frame is updated and invoked at regular intervals, passing current data to the listener. Due to this class' purely event-driven functionality and single entry point, its call graph represents all influences within the system. This is shown in Figure 52.

The 'Ghost Track Instruction' class, accessible from the selection menu of the ECS, is registered to record both of the previously described streams of data. Functions to start and stop recording data are called explicitly by the expert during use of the ECS. When ceasing the recording, the stream listeners are unregistered, and the file is saved. The update functions in Figure 23 are loops.

The ghost track is an example of how new types of instruction can be added to the ECS. By providing a developer interface that can be called by the Instruction Controller (see Figure 17), the system can be extended in a modular way, allowing for a more extendible approach to development.

Aside from the ability to record and replay movement around the workplace, the ability to continuously localise actions provides an additional, body-oriented reference frame from which to view action and augmentation. The next section looks at how, from this frame, we may extend this system by including a person's arm movement as well.

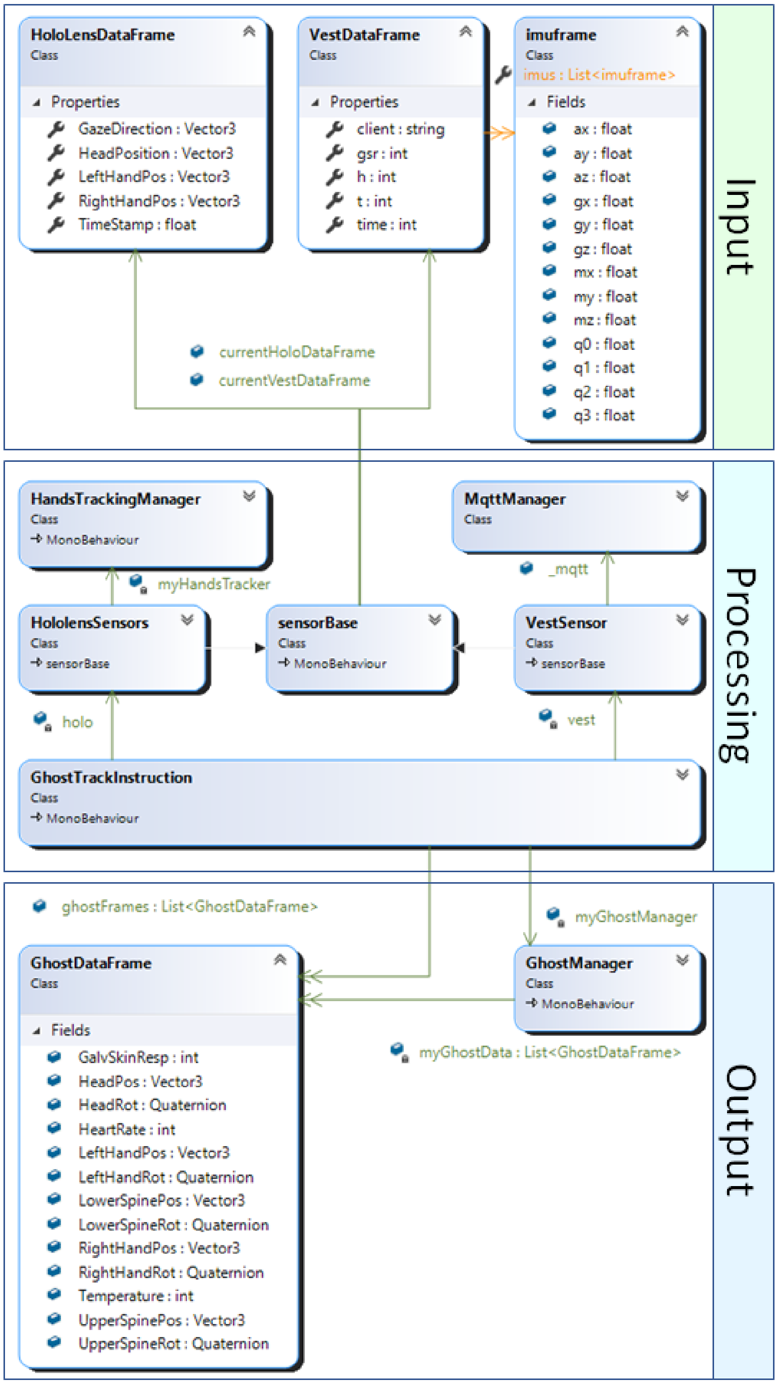


Figure 22: Ghost Track class diagram.

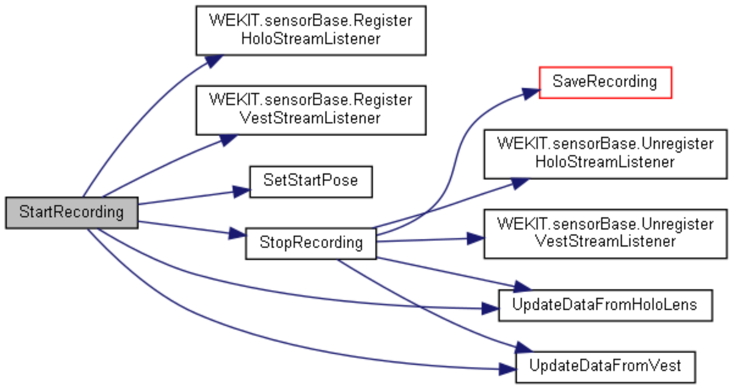


Figure 23: Call graph for beginning a recording of the Ghost Track.

5.2 Arm Tracking

This section presents details about the development of a system to capture arm activity so that it can be used in AR. Two armbands, capable of measuring movement and muscle activation were used—each a 'Myo', from Thalmic Labs. This part of the project saw the connection of the armband, via a bespoke adapter, to the HoloLens, informing both the visual output and the data model. There is a description of the multi-channel asynchronous system that allows for multiple connections and high throughput, which are normally significant limitations when using the communication protocol. Finally, there is a description of how the dataset, consisting of inputs from all channels as well as some low-level features, is stored and utilised by the ECS as a 'Hands Annotation'.

5.2.1 Integrating Myo Armbands

Of the devices available (see Section 2.4), the Myo armband was one of the few capable of measuring both IMU and sEMG data. This capability, as well as the ergonomic fit, wireless operation, and affordable price point were the main motivations for its selection.

The connection protocol used by the system for arm tracking is Bluetooth Low Energy. Its use was mandated by the choice of the Myo armband for data capture. While the protocol provides a clear and manageable structure for data transmission, the connection to the HMD proved to be more challenging due to a Windows-based limitation that only a single active BLE connection can be maintained at any one time. To overcome this, a new adapter was built that, while using the underlying UWP BLE framework, man-

Figure 24: Thalmic Labs' Myo Armband

aged all connections and data transfers. The other key challenge faced when using this device was the drift present in the IMU, which was partially mitigated through the use of opto-kinetic sensor fusion.

Inside each armband is a Cortex M4 120MHz processor, two 260mAh batteries, a NRF51822 Bluetooth Low Energy module, an inertial measurement unit (IMU), a vibration motor and eight sEMG electrodes, each with an accompanying circuit board. The IMU used is the Invensense MPU-9150, which can deliver 9-axis data at 100 Hz, limited by the magnetometer reading (the accelerometer and gyroscope can achieve 1 kHz). The device weighs around 250g and can fit arms with a diameter of 19–34 mm. [122].

Muscles recorded by the armband

There are ten superficial skeletal muscles that provide the bulk of the EMG signal. In the standard anatomical position, half of these lie in the anterior (front) compartment and half in the posterior.

Flexion of the forearm at the elbow is principally driven by the brachioradialis, shown in Figure 25. This muscle is served by the radial nerve, along with the extensor carpi radialis muscles on the posterior side. The remaining muscles of the anterior compartment work to flex the wrist (*carpi*) forwards with the palmus longus or in radial (away from the body) or ulnar directions, using the flexor carpi radialis or flexor carpi ulnaris, respectively. Inward rotational movement of the forearm (*brachium*) is done with the pronator teres. The flexor carpi ulnaris is supplied by the ulnar nerve, while the others receive signals from the median nerve.

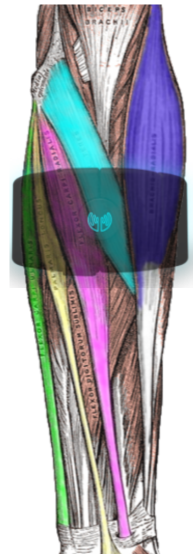


Figure 25: Superficial muscles of the anterior compartment.

The posterior compartment, shown and coloured in Figure 26, has a further five superficial muscles that contribute to the EMG signal: the small anconeus muscle that permits extension of the arm and provides support to the elbow joint; the extensor carpi ulnaris contracts to bend the wrist towards the body and works to extend the wrist, in concert with the extensor digitorum, which also works to extend the fingers. Radial abduction of the wrist (movement towards the thumb) is governed by the extensor carpi radialis (brevis and longus) muscles. Figure 26 also depicts the brachioradialis (far left).

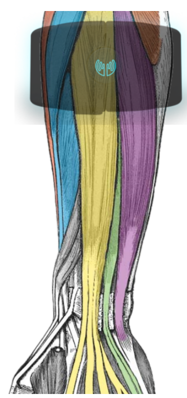


Figure 26: Superficial muscles of the posterior compartment (modified from [38]).

Though the specific movements mentioned above may require the activation of individual muscles, themselves aggregations of hundreds of muscle axons, the vast majority of arm actions will engage several simultaneously. The sequence in which they are activated and, to some degree, the amount of muscle force used can be determined through the use of sEMG. The typical placement of the armband is shown overlaid. It can be seen that, on the anterior side, variation in both rotational and distal placement would cause measurement error, while on the posterior side, changes in distal placement (along the arm) would not lead to incorrect muscles being measured.

5.2.2 Gathering Arm Data

Bluetooth Low Energy (BLE) is a communication protocol built around the advertisement of packets that identify and connect a server to a client. Using this protocol, the peripheral device can connect to a single central device, and the central device may maintain up to seven peripheral connections for a total network size of eight. BLE devices operating this protocol use the Generic Attribute Profile (GATT), which groups information into services, each holding one or more characteristics. There are three types of characteristics, distinguished by their communication properties. *Read* characteristics hold a value that can be accessed from the client; *write* characteristics receive information that can be used to change the state of the peripheral device; and *notify* characteristics will transmit data to a central device when it becomes available. The last type is also referred to as *indicate* characteristics when a response from the client is required. Notifications are subscribed to by the central device, which first requires a connection. Listing 2 shows the endpoints that were used.

```
private void Setup_Data_Channels()
{
    myoGuids = new Dictionary<string, Guid>();
    myoGuids.Add("MYO_DEVICE_NAME", new Guid("D5...42")); // Device Name
    myoGuids.Add("BATTERY_SERVICE", new Guid("00...fb")); // Battery Service
    myoGuids.Add("BATTERY_LEVLL_C", new Guid("00...fb")); // Battery Level
    myoGuids.Add("MYO_SERVICE_GCS", new Guid("D5...42")); // Control Service
    myoGuids.Add("MYO_FIRMWARE_CH", new Guid("D5...42")); // Firmware ver. (read)
    myoGuids.Add("COMMAND_CHARACT", new Guid("D5...42")); // Commands (write)
    myoGuids.Add("MYO_EMG_SERVICE", new Guid("D5...42")); // Raw EMG data service
    myoGuids.Add("EMG_DATA_CHAR_0", new Guid("D5...42")); // EMG ch0 data (notify
        )
    ...
    myoGuids.Add("IMU_DATA_SERVIC", new Guid("D5...42")); // IMU service
    myoGuids.Add("IMU_DATA_CHARAC", new Guid("D5...42")); // IMU characteristic
}
```

Listing 2: A code snippet showing the Myo's GATT endpoints

Once connected, the armband communicates two data streams as periodically updated values for Bluetooth GATT characteristics, which change at around 60 Hz during operation. The IMU uses a single channel, while the EMG data occupies four channels containing sequential data measurements, allowing data to be reported at more than 200 Hz. For muscle data, the main challenges related to data loss in transmission since BLE packets can become stalled or go missing during operation. For inertial data, there was a need to coordinate the movements with the headset, known as an axis mapping problem. In this case, the frame of reference to which the armband should be mapped was also a moving target. To tackle these issues, two pieces of software were built; the first was a BLE adapter capable of connecting with and receiving data from two armbands simultaneously, allowing the data pipelines to be configured more precisely. The second was a data recorder, which worked to transform raw EMG data into structures that could be subject to analysis and would send inertial data to the HoloLens for the placement or control of virtual objects.

A Brief Tangent: Getting More from BLE

BLE is generally focused on packet-wise communication, rather than streaming. However, the communication protocol does not explicitly require the acknowledgement of packets since this action is carried out on the link layer. The consequence of this is that, with the right programmatic patterns, this need for continuous acknowledgement of data can be avoided, allowing considerably higher data throughput. It should be noted that this was made possible due to the reliability of the connection, itself the result of the two devices remaining roughly a metre apart. Packet losses were relatively minor and generally obvious due to the bit sequence becoming shifted by a single increment. This allowed the use of simpler data checks that could be performed on the fly, such as checksums, maintaining the higher rate of transmission. When the session was ended, the connection could be reset, preserving some of the energy-saving advantages of the low-energy protocol.

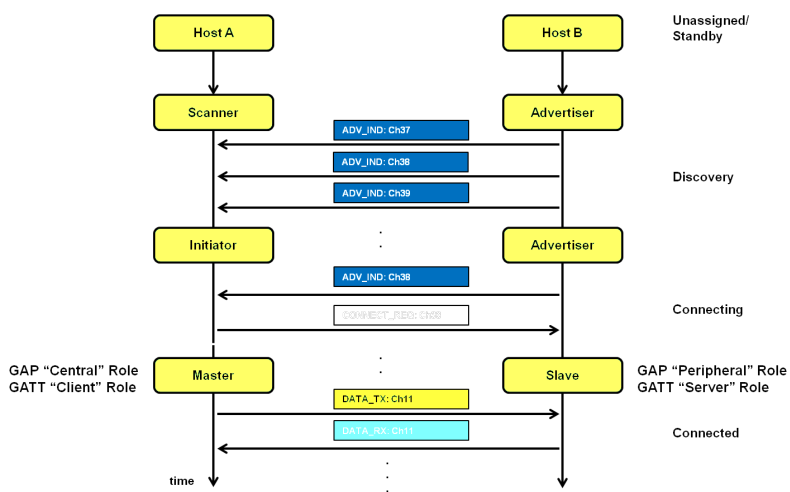


Figure 27: BLE connection flow.

Where Figure 27 ends is where the link layer begins its management of the data flow. In several experiments related to those described in Chapter 5, it was found that removing this requirement could double the amount of data flowing through the channel, though it required a lower level of access to the hardware, together with more specialised knowledge of the various components. The drawback to this modification is that it is no longer possible to anticipate the end of the data stream; the receiving device must explicitly send a command that, when received through an interrupt, would cease the transmission of data. The easiest way to do this was with a system restart, unless there was cross-session data that needed to be preserved.

Returning to the main plot, the BLE adapter functioned by creating several instances of a process that would connect to, and receive data from, an armband. In a standard Universal Windows Platform environment, this parallel operation is easily implemented. Within the Unity environment, where threads are more tightly controlled, the careful execution of parallel tasks played the same role. Figure 28. shows the user interface and an example of the services and characteristics discovered on a device.

Figure 28: Screenshot showing GATT services and characteristics.

The software was used to identify and track the armband's GATT signature and to develop algorithms for automatic connection and disconnection. These included features such as connection auto-restart, haptic readiness confirmation, and the functions connecting to the built-in pairing mechanism on the Myo. Since the original application was made using the XAML framework, it was easily converted to a placeable panel display within the AR training interface, as seen in Figure 29.

Once this software was relatively autonomous, the second piece of software was developed: a tool to provide continuous streaming data used to locate a virtual hand. Since the mode of interaction is AR, this virtual entity is not a replacement but rather an extension of your body, in this case, one capable of visualising muscle activity. The virtual hand could equally be imbued with contextual significance, allowing the information related to position and muscular activity to be correlated with events within the work flow and allowing these data inputs to affect changes to the state of the overall training system.

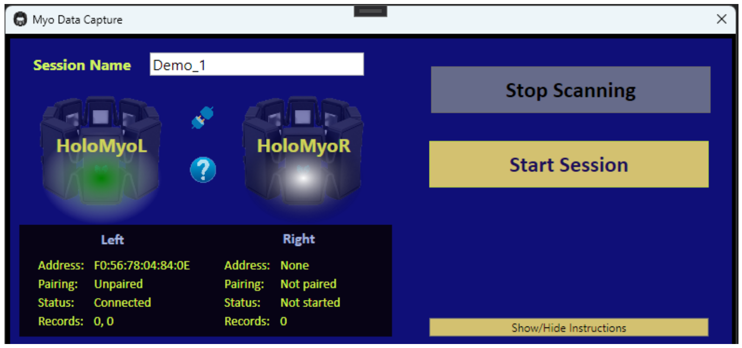


Figure 29: The MyoCapture interface, showing one of two armbands connected.

The data streams produced by this software allowed for an investigation of the use of forearm-based motion capture in concert with the HoloLens. When the devices were recognised and detected, the virtual hand could be created at a given offset from the central, head-worn device. The IMU data stream contained both a gravity-aligned measurement of attitude and heading, and raw values were available for the accelerometer and gyroscope components. There was not, however, a magnetometer output, meaning that it was not complete enough to reconstruct a complete attitude and heading reference system (AHRS) for the device.

EMG data was filtered at 50 Hz to remove electrical interference, after which it was streamed into a data buffer. During this process, there was the opportunity to capture additional information that, with minimal computational overhead, could provide insight into the data. This opportunity arose from the way in which the EMG data was transmitted; four BLE characteristics were used, arriving in a repeating sequence. Each of the eight channels would send two sequential packets, which meant that simple calculations could be done in the interval between the packets arriving, and the overall latency of the process would not be affected.

5.2.3 Building the Hand Track Instruction

The approach used for this implementation relied on a rudimentary inverse kinematic model for the placement of the virtual hand, fused optical tracking data from HoloLens, and orientation data from forearm-worn motion controllers. When the hands were within the field of view of the HoloLens, the location was determined by the camera system, and the orientation was taken from the IMUs. The hand-tracking capabilities of this first-generation headset meant that the orientation of the hands was not estimated by the vision system. When optical tracking was lost, control of position was assigned to the second differential of the accelerometer measurements, which, though less accurate, allowed for the persistence of these virtual appendages during the workflow, providing a continuous data stream that could be connected to muscle activation signals.

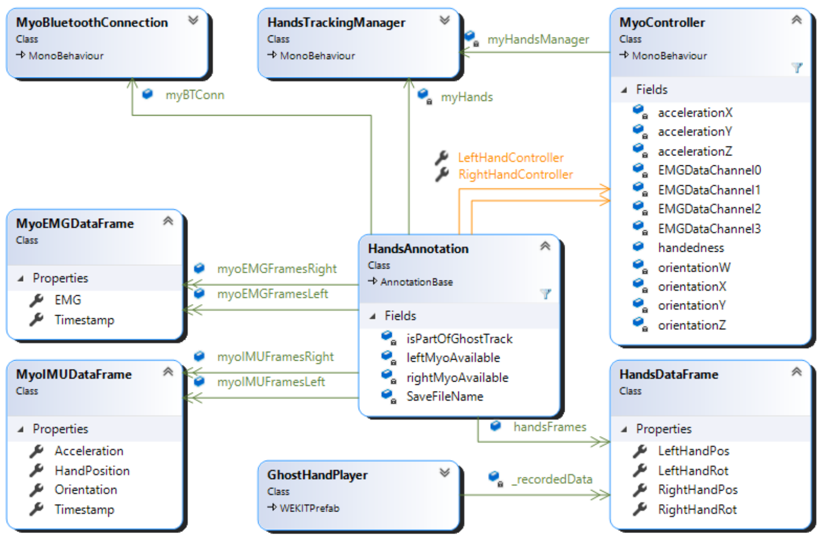


Figure 30: Hand annotation class diagram, including key data elements.

One of the first steps was to detect hand detection events and provide switching between the two modes, which was taken care of by the 'HandsTrackingManager', shown earlier in Figure 22, where it was used to localise the person in the space. This class can be seen in Figure 30 where it provides one form of input to the 'HandsAnnotation' component. When it was time for the armbands to provide data to the training system, frames were enabled and connected, feeding both movement and EMG data to the annotation. Depending on whether the hand-visualisation annotation was part of the "ghost track", described in Section 5.1.1, this data could be alone or as part of a larger data frame describing pose and gesticulation in the virtual space. During the expert capture phase of the procedure, a set of 'HandDataFrame' entries would be kept, down-sampled to 25 frames per second. On playback, when the trainee arrived at the relevant point in the work flow, the 'GhostHandPlayer' would place virtual objects in the correct pose, and they would re-enact the recording.

The 'MyoController' did the bulk of the data processing and transformation. After storing the raw data, shown as the 'MyoEMGDataFrame' and 'MyoIMUDataFrame' classes, for further feature analysis, the controller would arrange the incoming data into suitable arrays. In both cases, the data storage elements were prepared during the initialisation of the device connection to minimise the delay in recording the first packet. Programmatically, this means creating new files and buffers with the correct dimensions. Timestamps were applied at the very last moment, in terms of code, to minimise inaccuracies, although an understanding of frame-rate timing and management in the Unity framework was found to be a more critical factor. Connection as well as reconnection events were fully automated and signified with a double pulse on the built-in vibration motor, and errors in the connection or data flows were passed to the visual interface, the most common of which were BLE-related quirks, especially when working with several devices in a user trial environment.

The placement of the augmented appendage relied on the use of several frames of reference. The first was that of the workplace, which allowed both the HoloLens and the armband to maintain an absolute reference in space (you can think of this as an"world-locked" frame of reference). The second was the moving coordinate system of the HoloLens (a "head-locked" frame), which could be used to calibrate the armbands but not for ongoing placement. Finally, there were intermediate frames of reference introduced in order to fine-tune the virtual hand position, using offsets from the head position, but in the world-locked frame. The final result of this algorithm was a sequence of 'HandDataFrame' entries that describe the positions and rotations of the virtual hands explicitly in the reference frame of the activity, meaning that data recording made had a tangible connection to the workplace and recordings could be played back with reference to a single AR marker: the origin of the activity.

The connection to the overall experience recording framework in which augmented activities were designed was the final piece of the puzzle. The 'HandsAnnotation' class was the interface to the wider ECS, and this could be used in conjunction with the 'GhostTrackAnnotation' described in Section 5.1.1 or as an independent annotation type. This modularity was beneficial for at least two reasons. Since this annotation required additional (wireless) hardware, it could not be guaranteed that this data stream was available, hence the 'Ghost Track' was not dependent on the hand position to work. Conversely, the position of the head can detract from the design of an action step if, for example, the hand position is used to indicate a location or shape, but the appearance of the body would obscure the target. The data field 'isPartOfGhostTrack', denoting the state of inclusion, can be seen in Figure 30.

5.3 Summary

This chapter outlined the development of two features: incorporating body and arm tracking into the ECS. These instructions, which could be utilised by trainers to provide embodied affordances, were designed in a similar fashion to other instructions but were derived from a wider and different set of requirements. Whereas body tracking was facilitated by the VISLAM algorithms on the headset, the tracking of Bluetooth-connected armbands relied on 6-axis accelerometer data, which was insufficient for accurate placement, especially in the presence of sensor drift. Despite this limitation, significant potential was found in the use of both movement and muscle-activation data streams.

Chapter 6: System Evaluation

Adding features that allow for body and arm tracking completed the set of proposed affordances for the ECS. Based on the use of this immersive training system in three knowledge-intensive use cases, described in Section 4.1.1, data could be collected about the users' perspectives of the technology.

Technology acceptance models investigate characteristics of human behaviour that relate to the adoption and use of specific technologies. Their strength lies in their broad application and adaptable structure. The model used here seeks to capture a broad range of perspectives in acknowledgement of the wide range of possible influences that may exert themselves, ultimately leading to the intention to use a technology, in this case, augmented reality and wearable sensor networks.

It is important to recognise that these models aim to reflect the nuance of human intention. This necessarily reflects a balance between model specification and parsimony; having more constructs can capture latent connections, but at the cost of complexity and potential redundancy, allowing the model to reflect unknown attitudes and beliefs.

Structural equation modelling (SEM) is a technique for compiling, arranging, and studying a set of observations and then testing their correlation against a model comprised of theoretical constructs, arranged in a testable way. Probabilistic metrics for the fit between the model and the observations can be taken, along with measures such as model parsimony, factor loading or regression strength. Relations in the model can be modified or removed, and the impact of this change can be measured or abstracted. Modern tools can anticipate misspecification and a growing number of statistical tools support optimisations across a variety of metrics.

This chapter details two iterations, each subject to verification, estimation and optimisation, followed by a third and final model that incorporates the findings from the second stage. The data relate to the use of the technology in general, rather than the specific instance of the ECS described earlier, although all of the procedures carried out were recorded with that system. We will follow the development of the measurement questionnaire, AR-specific constructs and overall regression models. It uses both exploratory and confirmatory factor analysis to inspect the structures within the data, reliability analysis and modification indices for statistical insight into model misspecification and refers to several (robust) test statistics when measuring the predictive strength of the model to describe the data. High-fold Monte Carlo cross-validation is used to prevent gross over-fitting of the data. Other metrics are also used, such as item complexity plots, to help illustrate features in the data landscape.

UTAUT2 Constructs

Performance Expectancy (PE) is the belief or view that the technology will help them to achieve gains in job performance. This kind of overall assessment can take into account several factors, such as the level of motivation or changes to outcome expectancy, after social cognition theory, or relative advantage, following ideas on innovation diffusion. It is commonly found to be a reliable predictor for behavioural intention [11] and is sometimes seen to share a residual correlation with the effort needed to use the technology.

Effort Expectancy (EE), defined by Venkatesh et al. as "the degree of ease associated with the use of the system" [125, p.450], is also regularly viewed as a driver of attitudes and behavioural intention [26]. Views on PE and EE arise from a connection between a task, the technology and the user. Their correlation points to a broader sense of the reward-for-effort conception that people have towards the technology and its use in aid of a particular task.

Facilitating Conditions (FC) are another important concern and represent views around those organisational or technical factors external to the AR system that support its use. This may, for example, be in the form of a wider technical mission or development programme that provides the use of hardware, offers training, or includes technical support. FC has been shown to be a helpful concept when predicting initial uptake of technology as well as its continued use, meaning changes to a person's perspective about FC can directly drive the usage frequency of a technology, hence the factor loading (UF <- FC). Another representation of FC is perceived behavioural control (from TPB), which has also been shown to be a powerful indicator of engagement with a technology [3].

Social Influence (SI), the last of UTAUT's original constructs, is used to capture the perceived impact of technology adoption in terms of organisation influence, assuming that either those people (in the organisation) who were more influential may encourage adoption or that, though their own adoption, their influence within the organisation may increase. In the consumer world, where there are often looser connections between those using a system, this factor may be diminished. Indeed, some studies have shown no impact of this construct on the intention to use the technology [49].

Hedonic motivation (HM) is the fun or pleasure derived from using the technology. Venkatesh et al. [124] point to research on perceived enjoyment as an equivalent, a construct that has been shown elsewhere to be a predictor of PE, EE, and BI [121].

Finally, there is habit (HT), which is the belief or view that the use of the technology happens in a way that is automatic, implying a sense of familiarity, but in a way that is distinguishable from prior experience, which has also been put forward as either a predictor of habit or a mediating factor for it. The interpretation used in this work follows that of Venkatesh et al. [124] and Limayem et al. [65] who, in their popular work on this topic, propose that it moderates the influence of BI, such that the significance of BI decreases as the behaviour becomes more habitual.

6.1 First Technology Acceptance Study

This section begins with a summary of research first published in 2018 [43] by the author. This study analysed a dataset of one hundred and thirty responses, collected across three organisations (and areas of work): LuftTransport (aeronautic engineering), Altec (astronautic engineering), and EBit (sonography). Each of these organisations specified a trial procedure—a linear sequence of actions that, although common enough to be carried out by someone with trainee-level understanding, was also complex enough to be found challenging.

6.1.1 Specification of Augmented Reality Constructs

In the first study conducted, there were eleven questions relating to the core UTAUT2 model, out of a total of nineteen. The other eight were split between four other AR constructs: interoperability, learnability, augmented reality and wearable technology fit, and image. Each of these constructs was framed around a key aspect of the use of the technology, and their connection to the core UTAUT2 model was considered, leading to the first testable layout for an Augmented Reality technology acceptance model, shown in Figure 36.

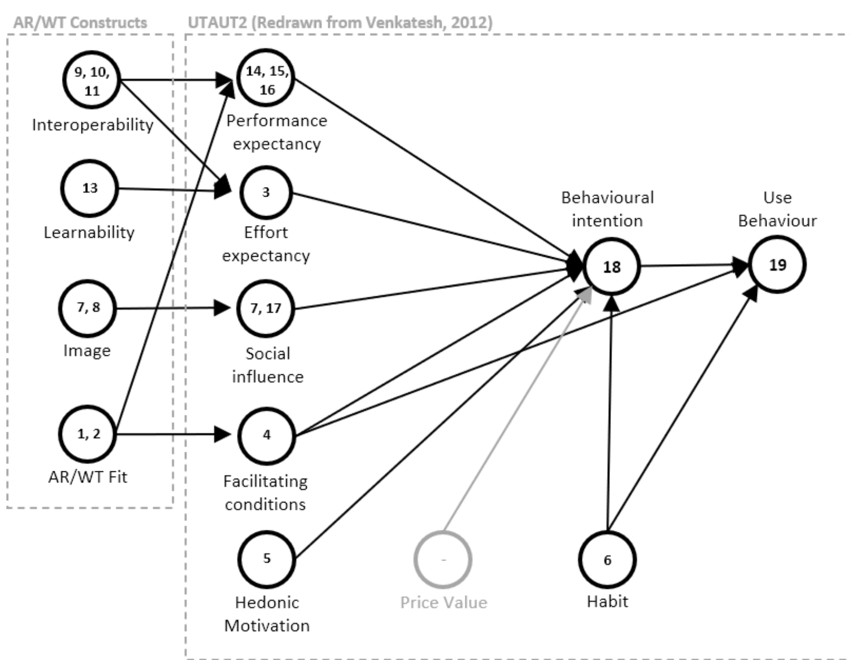


Figure 31: AR Technology Acceptance Model showing initial regressions for proposed AR constructs.

Interoperability (IOP) is the functional cooperation between systems that allows for the fulfilment of a task that can only be carried out with their combined operation. This definition recognises the AR system as being functionally modular, composed of a number of wearable components (headset included) that must operate in tandem for certain functions to be used or benefits to be realised. Using a hand gesture to bring up a visualisation is an example of such coordination. Since interoperability is directly connected to the completion of a task, there is the possibility for overlap with other attitudes that relate to task performance (found in PE) and, due to increasing complexity of use (i.e. more than one device) those attitudes that consider ease of use (found in EE).

Learnability (LRN) conveys the ability of a system to facilitate progress towards specific learning goals. It is tied to both the operational functionality related to the ease of learning the task (does the inclusion of AR make it easier to learn?) as well as its own inherent learning curve, in cases where the user is an AR novice (how much do I need to learn before I can improve my learning with this technology?). In the questionnaire, the construct was investigated by comparing the learning curve when learning to use AR with the value brought by the technology. Although this concept is somewhat abstract, suggesting that it could be applied in many ways, the particular wording as a "learning curve" shows a correlation to a subset of effort expectancy perceptions; hence, the construct is thought to predict some of the variance explained by EE, and a regression is drawn in.

AR/WT Fit (FT) is the view or belief that the experience delivered with AR and WT operates in a manner that is suitable (i.e. fits the requirements) for both the activity and the person using it. In considering the wearable system as an intermediating agent, this is equivalent to the extra affordances provided by the technology, either through the reinforcement of existing, localised information or the inclusion of something new. This construct was motivated by studies on task-technology fit, beginning with the seminal work by Goodhue and Thompson [37], who formalised the distinction between the utilisation of a technology and the ability of the technology to meet certain task requirements and showed that both are involved in impacting performance. These ideas found footing in later work by Parkes, who framed task-technology fit in terms of decision support systems: "the extent to which the complexity of the task being undertaken matches the decisional guidance provided by the technology" [86, p.999]. This definition is useful in this context, not only because we are considering the use of AR and WT in the context of workplace training, where decisional guidance is important, but also because it frames the benefit of the technology in terms of the task itself rather than features solely attributable to the technology. Task-technology fit has the potential to affect factors external to the user since the design of each experience was done in consultation with expert trainers in each workplace. For this reason, a regression is added between fit and facilitating conditions, which represents those beliefs that relate to the level of support provided by the task context.

Image (IMG) is synonymous with 'social approval of AR and WT'. When technology is worn on the body, and especially the head, it affects how we are seen, in both a literal and social sense. This construct investigates a person's perception of the social attitude towards the use of technology worn on the head and body. As with SI, a positive attitude indicates that the use of the technology improves their standing or influence within a social group or organisation, predicting a greater intention to use it in the future. At the other end of the scale, there is a belief that the use of the technology is disapproved of by others. When investigating this latent element, questions focused on the connection between the technology and prestige within the organisation as well as the intention to be an early adopter. Due to this notion being linked to a perception of the beliefs of others within our organisation (who matter to us), the construct was hypothesised to explain some variation within SI, and a regression was added to the model.

6.1.2 Questionnaire and Dataset

Responses were gathered from these workplaces over the course of three months, during which time participants would attend the workplace, receive an introduction to the project and the technology, and then proceed to use the technology to complete the relevant task. The questions were designed to investigate constructs within the UTAUT2 model as well as three others relating specifically to the use of augmented reality and wearable technology. The questions followed the use of the first iteration of the WEKIT.one experience capture system (ECS), described in Chapter 4, which was used to create AR training materials for novices to follow in their respective workplaces.

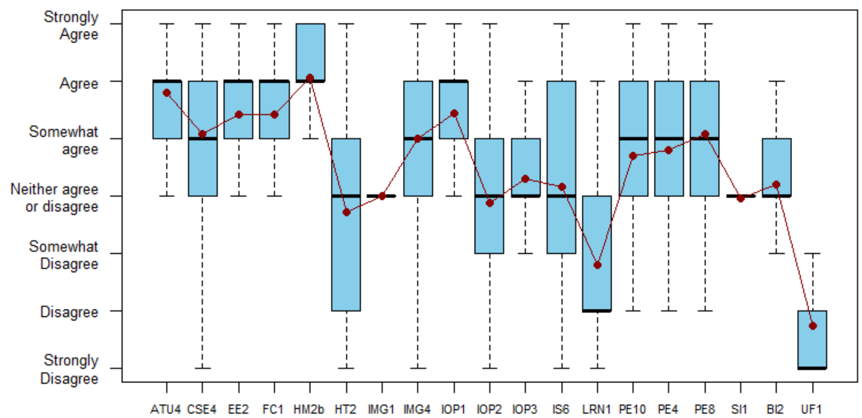
The set of questions was itself the output of a mixed-methods investigation, first described by Wild et al. [135] and shown in Table 12. The first dataset that was collected consisted of 130 responses to these 19 questions, investigating participants' attitudes and beliefs with regard to technology acceptance.

Table 12: Technology acceptance questionnaire, first iteration.

#	Code	Statement
1	ATU4	I look forward to those aspects of my job that require me to use AR & WT.
2	CSE4	I could complete a job if I had used similar technologies before this one to do the same job.
3	EE2	My interaction with AR & WT is clear and understandable.
4	FC1	I have the resources necessary to use AR & WT.
5	HM2b	I like working with AR & WT.
6	HT2	I am addicted to using AR & WT.
7	IMG1	People in my organisation who use AR & WT have more prestige than those who do not.
8	IMG4	I use AR & WT solutions because I want to be a forerunner in technology exploitation.
9	IOP1	AR & WT need to work together with the existing software systems to help me do the task.

10	IOP2	I worry that I could become too dependant on a single AR & WT supplier.
11	IOP3	Integration costs of AR & WT with other software systems in use are high.
12	IS6	I would find it useful if my friends knew where I am and what I am doing.
13	LRN1	The learning curve for AR & WT is too high compared with the value they would offer.
14	PE4	Using AR & WT increases my productivity.
15	PE8	AR & WT increases the precision of tasks.
16	PE10	With AR & WT, I immediately know when a task is finished.
17	SI1	People who are important to me think that I should use AR & WT.
18	BI2	I will always try to use AR & WT in my daily work.
19	UF1	Please choose your usage frequency of AR/WT.

A statistical summary of the responses to these questions is shown in Figure 32, which demonstrates a generally positive response to the technology and, unsurprisingly, a low usage frequency. Note the negative framing of items LRN1, IOP2, and HT2 when interpreting the graph.



	ATU4	BI2	CSE4	EE2	FC1	HM2b	HT2	IMG1	IMG4	IOP1	IOP2	IOP3	IS6	LRN1	PE10	PE4	PE8	SI1	UF1
Mean	5.80	4.19	5.07	5.42	5.41	6.07	3.71	4.01	5.00	5.44	3.87	4.30	4.16	2.80	4.70	4.80	5.09	3.96	1.74
Std.Dev	1.01	1.17	1.35	1.19	1.36	0.94	1.79	1.06	1.17	1.13	1.28	1.23	1.68	1.31	1.33	1.16	1.14	1.13	1.21
Median	6.00	4.00	5.00	6.00	6.00	6.00	4.00	4.00	5.00	6.00	4.00	4.00	4.00	2.00	5.00	5.00	5.00	4.00	1.00
Skewness	-0.93	-0.16	-0.54	-1.45	-1.22	-1.12	0.09	-0.16	-0.32	-0.88	-0.13	-0.30	-0.20	0.60	-0.42	-0.07	-0.49	-0.14	1.64
Kurtosis	0.83	0.24	-0.17	2.50	1.09	1.21	-1.12	2.36	-0.01	1.15	-0.24	0.24	-0.89	-0.46	-0.78	-0.37	-0.11	0.96	1.83

Figure 32: Responses from trainees (N=130)

A note on immersive data and ethics

Datasets such as this hold valuable, multifaceted insights, and the ethical production, use and storage of them is of primary concern. This is not simply because they have the potential to indicate important aspects of the technology, but because they simultaneously highlight pervasive beliefs and attitudes within the population that may also be subject to external influence. It is, therefore, essential that such data be kept anonymously and used only for the purposes for which it was gathered.

AR also enables new types of data to be collected, both at the raw data level (e.g. head and hand movements) and that of inference, such as which features of a space a person is attending to. Given the existing value placed on accumulating user interaction data by technology companies, where it is common to forego payment in exchange for personal usage information, it is likely that movement and attention data from immersive technology will be similarly aligned.

Insights from factor analysis

Values obtained for both Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy suggest that there is sufficient correlation in the data to make factor analysis appropriate. A test of sphericity determines whether the correlation matrix is an identity matrix, which would make the data unsuitable for factor analysis. The null model has 120 degrees of freedom, giving an approximate chi-square value of 536 and a p value of less than 0.001, clearly rejecting the null hypothesis. The KMO is a measure of how well some parts of the data explain others. It is another requisite for factor analysis suitability, and values closer to 1.0 indicate that this property exists. A lower bound for acceptability is always greater than 0.5, though some consider 0.8 a more appropriately stringent value. The data showed good partial correlation, with a value of 0.81.

With this support for further investigation, we turn to a key question in quantitative psychology studies: the determination of an appropriate number of factors in the model. With this aim, several techniques were used to draw out useful features within the data, including the Very Simple Structure criterion (VSS), a measurement of model complexity, the extended Bayesian Information criterion (eBIC), and the Standardised Root Mean Residual (SRMR). A summary of these metrics is shown in Figure 33.

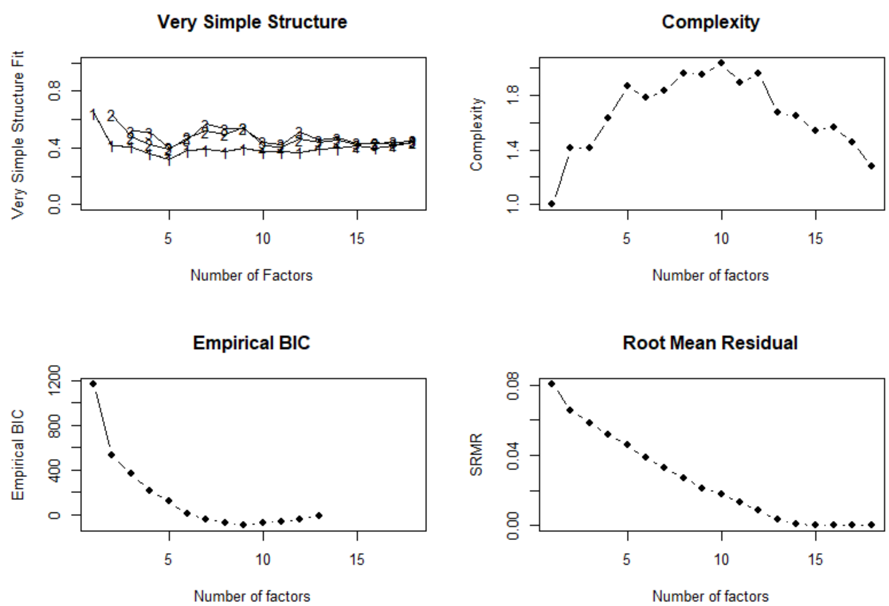


Figure 33: A number of factors analysis of the TAMARA1 dataset.

The VSS criterion compares the correlations in the model to a simple equivalent, where each latent variable is composed of the singular largest factor loading onto it. In this case, the VSS criterion alone has few discernible features and, despite greater confluence above twelve factors, demonstrates poor discrimination when considering the various numbers of factors. The dataset's unidimensionality, which VSS investigates, can also be drawn out through a measurement of item uniqueness (the inverse of communality), which describes the degree to which an item's weight is not accounted for through latent connections within the model.

Complexity, referring to Hoffman's index, represents the number of factors required to explain each item. This index shows bounds at five and twelve factors, outside of which the dataset would be better represented by significantly fewer factors, a sign that the model has diminishing descriptive power. Figure 34 sets complexity against uniqueness, giving a picture of the model's unidimensionality as well as the observed variables' multidimensionality.

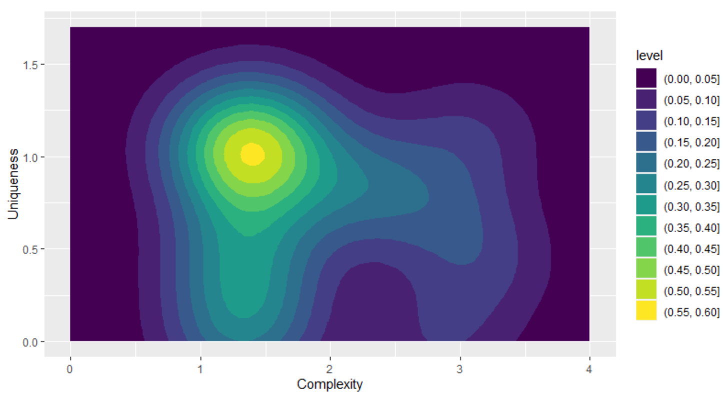


Figure 34: Contour plot showing factor dimensionality.

BIC is a measure of the information lost due to a misfit between the data and the model while varying the number of factors in that model. When computing BIC, Figure 33 uses a hypothetic fit, using a larger sample size. This was done because, for the current size of the dataset, the additional variable penalties applied (above a two-factor model) by this technique dominate the result, preventing any meaningful comparison. With this correction, BIC indicates a stable local minimum centred on nine factors, though there is a relatively wide region (6 to 14 factors) where the criterion falls close to or below zero.

The sample-size-adjusted BIC (SABIC) provides a metric for the theoretical information density. When plotted against the changing degrees of freedom, it is possible to avoid the BIC factor penalty from obscuring the trend and improve the resolution around each factor. The result indicates that models with 75 to 125 degrees of freedom offer minimal losses in information due to structure. The graph shown in Figure 35 has a minimum at 117 degrees of freedom.

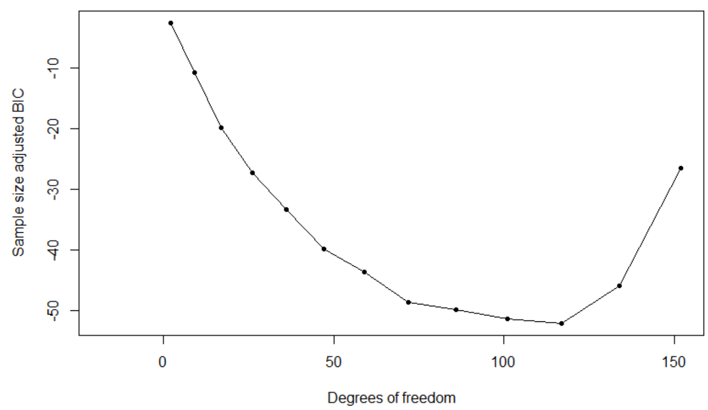


Figure 35: Evolution of the sample size adjusted BIC with degrees of freedom of the model.

SRMR describes (roughly) the average standardised residual covariance in the model. It has been shown to be a safe measurement of the degree of misfit of an ordinal factor analysis model [111], though with somewhat larger datasets than the one here. An additional index worthy of note is uniqueness, which is the proportion of a variable's common variance not associated with the factors. A standardised index, it is the opposite of communality. The SRMR drops below 0.05 with five factors and is negligible when there are fifteen, suggesting overfitting at this level.

6.1.3 Model Fitting and Optimisation

The first dataset, consisting of one hundred and thirty complete responses describing twelve constructs, using the Spearman correlation method, converged in 84 iterations. This method was selected over Pearson as it does not make the assumption that the variables are linearly correlated, something that is yet to be investigated in AR and WT research.

During optimisation, a variety of goodness-of-fit metrics were used to assess the predictive power of the model. These indices, along with covariance matrices, modification indices and residual correlations were used to pinpoint model misspecification. Monte-Carlo cross-validation was used to diminish the risk of overfitting. For more details on the changes made during the first iteration, the reader is referred to the previously published study [43], which contains further qualitative assessment of some of the observed variables wording and interpretation.

In summary, IOP3 was identified as contributing little information to the model; it showed up in several of the initial modification indices and shows minimal correlation with any other variables. The combination of these factors led to the removal of the item. LRN1 was similarly isolated within the correlation matrix and, in addition, was thought to be somewhat incomprehensible and was removed. The IMG construct, while considered valuable, was merged with the social influence construct in order to improve model parsimony. Of the two factors loading onto the image construct, one was changed to load directly to SI (IMG1), and the other was removed (IMG4).

The result of these optimisations was the reduction of the number of AR-specific constructs to two: interoperability and AR/WT fit. The remaining UTAUT2 model was not changed, other than the removal of the price value construct, as previously mentioned. A path diagram showing all regressions within the model is shown in Figure 36, where more significant paths are shown in bold. This model's statistically significant regressions are shown in Table 13.

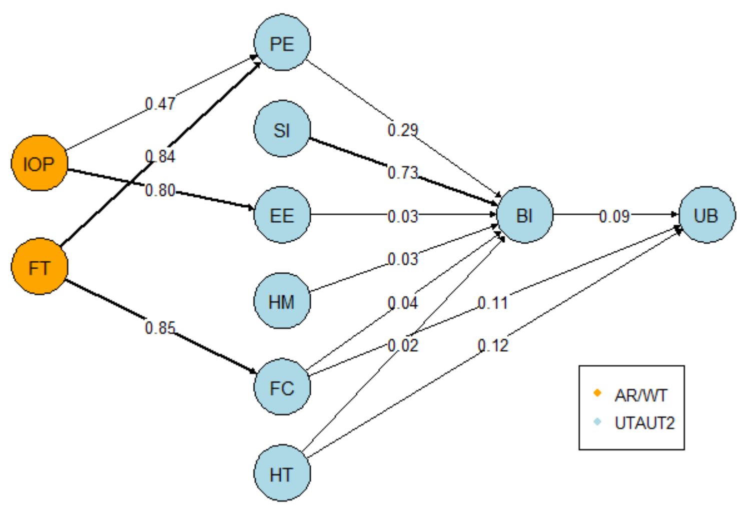


Figure 36: Path diagram of the optimised AR technology acceptance model.

The most significant regressions appear in relation to the AR/WT constructs; around 85% of the variance in both PE and FC can be predicted by variance in the FT construct, and 80% of the variance in EE is correlated with variance in IOP, indicating that these new constructs are both influential and statistically relevant; all three connections within the model have a p value of less than 0.05. The upper confidence interval of these regressions also extends high enough to suggest that measuring these constructs alone can give a good

Table 13: Significant Regression Estimates

lhs	op	rhs	est	se	z	pvalue	ci.lower	ci.upper	stars
FC	~	FT	0.854	0.224	3.820	0.000	0.416	1.293	***
PE	~	FT	0.842	0.223	3.778	0.000	0.405	1.278	***
EE	~	IOP	0.799	0.272	2.943	0.003	0.267	1.332	**
BI	~	SI	0.726	0.292	2.491	0.013	0.155	1.298	*
UB	~	HT	0.121	0.057	2.096	0.036	0.008	0.233	*
BI	~	PE	0.293	0.142	2.056	0.040	0.014	0.572	*

approximation of how important PE and FC are to the person, though the fact that FT is predictive of two separate constructs underscores the relevance to the underlying UTAUT2 model. Other regressions with the core UTAUT2 model are also found to be significant, namely the covariance of behavioural intention with both social influence and performance expectation and the description of usage frequency with the habit construct. PE is shown to explain almost 30% of the variance of BI, a finding routinely encountered in technology acceptance studies. More unexpected is the high level of predictive power of the SI construct, which shows over 72% covariance with BI, making this an essential element when determining the impact of the technology on intention and, consequently, future adoption.

Though significance is shown in the UB HT regression, an inspection of these variables (see Figure 32) offers an alternative interpretation: that the nature of AR as an emergent technology prevents both regular use and the possibility for people to develop habitual influences. Not only does this explain both values receiving lower scores, but also the wider spread of answers to HT2.

As well as the model's paths, it is also useful to look at significant factor loadings, which are shown in Table 14. Two constructs, PE and SI, were both informed by factors with strong statistical influence. Exogenous variables PE8 and PE10, which ask about task precision and their awareness of task completion, are shown to drive change in the latent variable by 0.72 and 0.47, respectively. Two of the three factors determining SI, namely IS6 and IMG1, also produced a measure of statistical influence with a negligible p-value. These items refer to the technology conferring prestige to those who use it (a loading of 0.77) and also to a measure of trust for the organisation in terms of data privacy (a loading of 1.13), indicating that they both point to an underlying attitude or belief, but both are strongly influencing the result. There are two remaining factor loadings that show significant influence in shaping their constructs. One is the expression of AR/WT fit associated with computer self-efficacy; that is here interpreted as individual-technology fit (item CSE4) loading onto the FT construct. This item explains around half (0.57) of the variance in the construct. Finally, interoperability is shown to be influenced to roughly the same degree (0.52) by the item IOP2, which asked about their concern for vendor lock.

The final values for the fit indices were the result of a high-fold Monte-Carlo

Table 14: Significant Factor Loadings

lhs	op	rhs	est	se	z	pvalue	ci.lower	ci.upper	stars
PE	=~	PE8	0.722	0.105	6.866	0.000	0.516	0.928	***
PE	=~	PE10	0.473	0.122	3.884	0.000	0.234	0.712	***
SI	=~	IS6	1.134	0.294	3.852	0.000	0.557	1.712	***
SI	=~	IMG1	0.774	0.212	3.643	0.000	0.358	1.191	***
FT	=~	CSE4	0.569	0.185	3.071	0.002	0.206	0.931	**
IOP	=~	IOP2	0.518	0.251	2.062	0.039	0.026	1.010	*

cross validation, which was preceded by an investigation of the impact of sample size on the fit indices, which was conducted to give further insight into the likely impact of performing this cross-validation. This was done by varying the ratio of training data to that of test data and observing the instability in the fit indices. As seen in Figure 37, a train-test ratio of more than 1.4, corresponding to a 76–54 split in data points, shows greater stability in the fit metrics, suggesting that some additional data would benefit the stability of the model fitting process, but the sample size is not far from being sufficient.

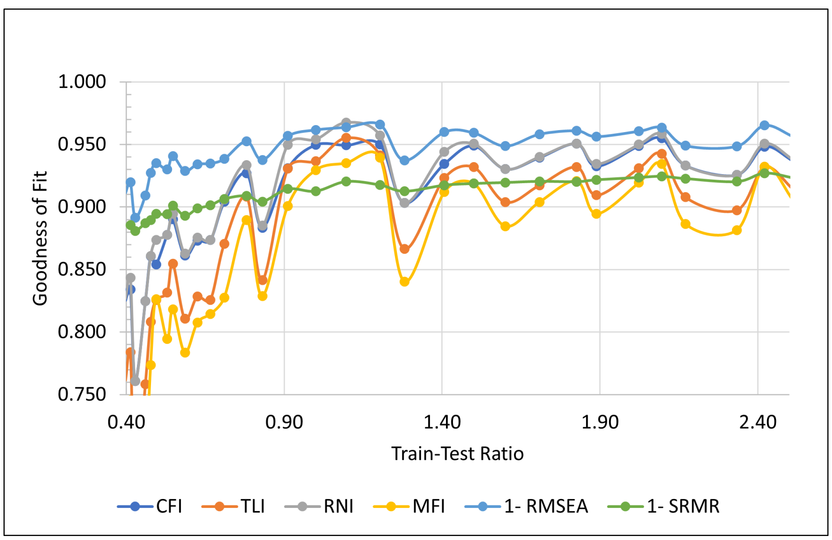
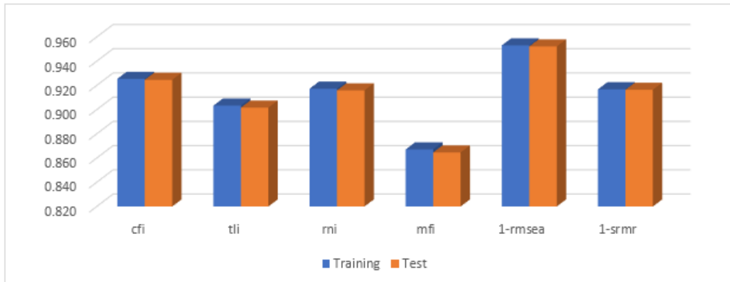


Figure 37: Model fit when varying train-test dataset ratio.

The final fit metrics were obtained after cross-validating the dataset with 1000 folds and are shown in Figure 38. Values for RMSEA and SRMR, instead of their usual scales, are written as '1 - value' so that they can be presented on the same scale as the other metrics.

Both absolute and relative fit indices are included in the summary. Where



Mean Values	cfi	tli	rni	mfi	1-rmsea	1-srmr	chi-sq	p-value
Training	0.925	0.903	0.917	0.867	0.953	0.917	108.684	0.134
Test	0.925	0.902	0.916	0.865	0.952	0.916	106.834	0.151

Figure 38: Model fit when varying train-test dataset ratio.

possible, all metrics use a robust correction, adjusting the chi-square statistic for non-normality in the data. Several common fit indices are reported; the comparative fix index (CFI), Tucker-Lewis index (TLI), and McDonald fit index (MFI). Each of these reports that the model has moderate predictive power, not quite reaching cut-off values for good fit, at 0.95, 0.93, and 0.90, respectively.

Other fit indices suggest a similar valuation of the model. The Chi-squared p-value, relative non-centrality index (RNI), root mean square error of approximation (RMSEA) and the standardised root mean square (SRMR) were used to describe the model's correspondence to features in the data. The RNI's value of > 0.916 indicates moderate-to-good power, and the RMSEA has a value of < 0.048, suggesting that there is a good fit. The SRMR, another absolute fit metric, falls at around 0.08, also suggesting a well-fitted model. Overall, these are encouraging signs that these constructs can meaningfully describe features of technology acceptance within the users' responses.

6.1.4 Conclusions and Recommendations

The results yielded several quantitative insights into the functional coupling between the AR training system and a model for technology acceptance. It validates parts of the UTAUT2 model, suggesting that it represents a robust base from which to perform statistical analysis. All meaningful correlations indicate relationships that are well-defined in UTAUT2 and there was no indication that any part of this model was at odds with the patterns seen in the data and, as such, no modifications were seen as necessary or even helpful. That is not to say, however, that improvements cannot be made, but rather that the underlying principles have descriptive power.

Table 15: Results of first-iteration hypothesis testing.

#	**Hypothesis**	**Outcome**	**Reason**
H1	IOP is significant in predicting acceptance of AR and WT	Rejected	IOP, though showing a strong latent correlation to FT as well as modest predictive strength for EE, cannot be said to influence acceptance.
H2	LRN is significant in predicting acceptance of AR and WT	Null	Construct removed in optimisation.
H3	IMG is significant in predicting acceptance of AR and WT	Null	Construct removed in optimisation.
H4	FT is significant in predicting acceptance of AR and WT	Accepted	FT seems to play a significant part in predicting PE as well as FC, both of which are subsequently seen to predict BI.
H5	IOP is positively related to PE	Rejected	With a p-value of 0.072, the covariance of these is not significant.
H6	IOP is positively related to EE	Accepted	The model claims to predict 81% of the covariance of these items, though more data is needed to confirm this.
H7	LRN is positively related to EE	Null	Construct removed in optimisation.
H8	IMG is positively related to SI	Null	Construct removed in optimisation.
H9	FT is positively related to FC	Accepted	The model claims to predict 85% of the covariance of these items.
H10	FT is positively related to PE	Accepted	The model claims to predict 86% of the covariance of these items.

The inclusion of AR constructs generally improved the model fit, though the connections between these new latent variables and those of UTAUT2 were worthy of further investigation. The AR/WT construct, in particular, showed high covariance with both PE and FC. This is potentially very significant, as these UTAUT2 constructs are often strong drivers of the adoption of technology. To develop this further it would be helpful to refine what 'AR Fit' means in the context of workplace training and how this relates to the more specific details of PE (expectations of increased job performance when using AR) and FC (the sense that AR acts as an external scaffold to support the task).

Interoperability showed a fairly strong correlation with EE—the anticipated ease of use of AR and WT in performing tasks—suggesting, in particular, that integration with existing systems would make these tools more accessible. Since this work includes systems that track body movement using a wearable network of devices, this finding can be explored further, looking at how the perception of interoperability relates to various parts working together to improve task performance. Interoperability was not seen to relate to PE, rejecting the hypothesis that the perception of a more integrated system is indicative of one that offers greater performance gains.

Several of the earlier hypotheses were untestable due to their exclusion from the model during optimisation. This is both routine and inefficient; on the one hand, the presence of a hypothesis should in no way validate its use; however, there is little purpose in generating hypotheses that are later not available for analysis. Good research habits, diligence in constructing new items, and discussion are helpful in minimising lost hypotheses. In this case, the wording of some of the items was overly complicated, as evidenced in part by the greater number of gaps in the data. Simple, balanced question formulations were a clear goal for future iterations.

6.2 Second Technology Acceptance Study

The second iteration of the ECS saw a number of improvements made to the wearable prototype as well as the software being run. The inclusion of sensors on the body and arms, as well as new types of visual augmentation accompanying them, such as the ghost track augmentation described in the previous chapter. For SEM, the second study extended the original dataset and also investigated the previously added constructs in more detail. This allowed the core model to be tested with more samples while deepening the investigation into those constructs directly related to the acceptance of AR technology. An additional item was added to probe interoperability, and four other items were included to examine the internal structure of the AR/WT Fit construct. All items relating to the UTAUT2 core model were identical.

The studies were carried out with the same industrial partners as the first round: LuftTransport in Norway, covering an updated maintenance procedure on an air ambulance; ALTEC, where the procedure involved inspecting and interacting with a model of a Mars rover; and Ebit, where participants performed a carotid artery ultrasound examination.

6.2.1 Further Development of Augmented Reality Constructs

From the first iteration, two constructs were brought forward that specifically investigate AR technology, interoperability and AR fit. The first has been shown to be somewhat predictive of effort expectancy, suggesting that the technology is perceived as easier to use when there is a corresponding view that the parts of the system work well together. The second was shown to have a connection to a person's estimation of the performance of the system and their sense of the external support available to them. Both were subject to further investigation, and a third construct—information security—was included, as it was deemed important and, crucially, not part of the existing model.

Information Security

This construct considers someone's perception of data privacy. In this iteration, it is used in an organisational context, based on item IS7: "In the future I would feel comfortable sharing the personal data captured with my organisation", which was hoped to resonate more strongly with the participants as the demonstration centred on workplace training. This item replaces IS6: "I would find it useful if my friends knew where I am and what I am doing", which focused less on the organisational context of the training, though it could be a useful measure when looking at the use of the technology in a broader social context.

This new conception is aligned with Weinhard et al.'s 'willingness to provide personal information' [132], which is presented as a composite construct, predicted by levels of personal interest, trust and privacy concern. The latter two are also thought to arise from an attitude towards privacy risks. In an organisational setting, it is not clear how these constructs would change. In this study, the item chosen to reflect this construct is more aligned with notions of trust and privacy than personal interest. While the questionnaire only uses a single question to test this construct, its inclusion was thought essential, as privacy concerns are frequently and necessarily addressed with the introduction of new technologies, especially those capable of capturing movement or physiological data about the user, which is doubly true for this prototype, although in a nascent form. Additionally, it should be noted that the participants were not members of the organisations they were visiting; however, since many came from nearby places, it is likely they were aware of the nature of the organisations and could reasonably develop attitudes relating to them.

Interoperability

The questionnaire included a new measure of interoperability, directly asking about the combined use of the hardware platform used by the project. IOP4 stated: "The smart glasses, vest, and armbands worked well together" and participants were asked to express their (dis)agreement with this statement on the usual seven-part Likert scale. A majority of participants had no prior experience with AR, so general remarks about its use with other wearable technology were not requested; instead, the question makes direct reference to the three wearable components, leaving little room for ambiguity.

Both interoperability and AR fit attempt to uncover beliefs that relate not only to the technology at hand but also the use of it to achieve an activity-oriented outcome. This nexus between individual, technology and activity is the central structure that this research investigates, and the further development of this idea is where the rubber meets the road; it is through the construct formulation and the specification of the associated factors that measure it that we may uncover the most direct relations between the three parts.

Augmented Reality Fit

The first round of SEM indicated that this construct is correlated with the level of performance expectancy for the technology. In other words, when someone considered the technology to be a good match for the situation, they were also likely to think of the technology as capable and performant. In addition, perception of AR fit was positively correlated with that of facilitating conditions, meaning that the assessment of a well-suited technology was likely to go hand-in-hand with a sense of being well-equipped for a task. While this may seem self-evident, drawing this conclusion demonstrates the equivalence of these ideas and allows for further examination of the construct, which was done in the second iteration.

In this section the internal structure of the AR Fit construct is investigated, measured, and integrated into the existing model. Rather than considering the fit between activity, individual, and technology as a single concept, there is an additional delineation between individual-technology fit (ITF) and another between activity and technology, named activity-technology fit (ATF). Each was studied with two additional questions, probing the person's view about the respective relationship. The two-part structure and the factors loading onto each sub-construct are shown in Figure 39.

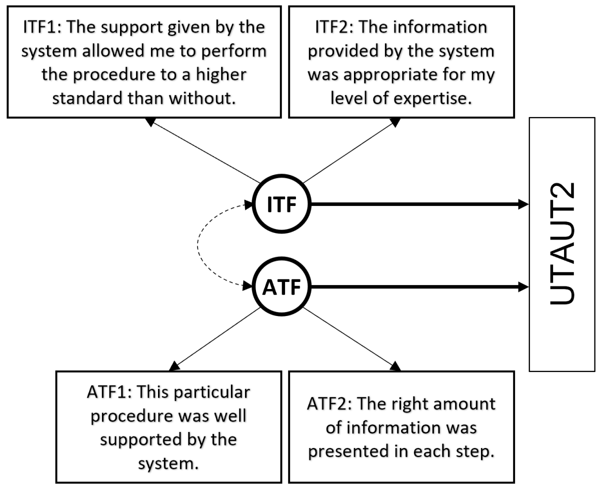


Figure 39: AR Fit construct structure and its observed variables.

The curved, dotted line connecting the ITF and ATF is a hypothetical residual correlation, representing the implication that there is further, underlying information that could tie these constructs together. This is, perhaps, a less radical position than assuming the two constructs are independent and could be thought of as task-individual fit, the remaining link in the triad of activity, trainee, and AR system. Parkes provides us with a definition for such a relation: "the extent to which the complexity of the task being undertaken matches the expertise of the individual" [86, p.999].

Returning to the latent variables, we can provide definitions that relate to the perceived coherence between the technology and either the individual or the activity. ITF is the degree to which the system is perceived to support one's individual competence, whereas ATF is the perceived ability of the technology to support the task. Both of these boundaries are context-dependent and could potentially act as contextual factors for one another. For instance, should the system significantly affect a person's ability to check for and identify defects in a surface, this close linkage provides a strong scaffold for learning, one that people of a wide range of competence can grasp. That is to say that the coupling to the individual competence becomes less significant in reaching certain learning outcomes. In a similar way, where the technology is closely suited to the trainee's expertise, the level of task performance could reasonably be expected to increase more than where there is a poor fit. Since we are dealing with learning affordances, it is important to note that the reverse is also possible. Where there is a lack of coherence, this can provide a more confusing context that might obfuscate otherwise apparent connections and work destructively, ultimately reducing task performance.

6.2.2 Questionnaire and Dataset

The second dataset was larger than the first, ultimately contributing 157 new records spread across 21 items, as shown in Table 16. Many core technological features were already in place and, more importantly, the routine for deploying and testing the technology with each industrial partner was established, allowing for a greater number of trainee simulations. The proportion of complete responses was, however, somewhat lower than in the first study, as this set was taken from an initial list of 300.

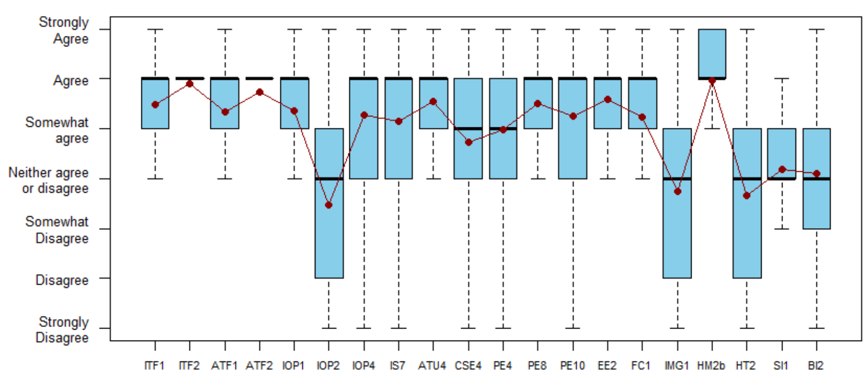
As in the first round, each of the three use cases was represented by a partner in industry who was responsible for identifying a suitable work flow and breaking it down into a series of task steps. Following some consultation with the technical team, an augmented workflow was recorded. Sections of tasks as well as whole task flows were recorded until a final version was approved for use by the trainer. This was done using the Experience Capture System (ECS) described in Chapter 4 and, thanks to the ARLEM format for storing activity and workplace information, was able to be edited, as the content was broken into human-readable steps with clear naming structures.

Table 16: Technology acceptance metrics for trainers using the ECS.

#	**Code**	**Statement**
1	ATU4	I look forward to those aspects of my job that require me to use AR & WT.
2	CSE4	I could complete a job, if I had used similar tech. before this one to do the same job.
3	EE2	My interaction with AR & WT is clear and understandable.
4	FC1	I have the resources necessary to use AR & WT.
5	HM2b	I like working with AR & WT.
6	HT2	I am addicted to using AR & WT.
7	IMG1	People in my organisation who use AR & WT have more prestige than those who do not.
8	IOP1	AR & WT need to work together with the existing software systems to help me do the task.
9	IOP2	I worry that I could become too dependant on a single AR & WT supplier.
10	IOP4	The smart glasses, vest, and armbands worked well together.
11	IS7	In the future I would feel comfortable sharing the personal data captured with my organisation.
12	PE4	Using AR & WT increases my productivity.
13	PE8	AR & WT increases the precision of tasks.
14	PE10	With AR & WT, I immediately know when a task is finished.
15	SI1	People who are important to me think that I should use AR & WT.

16	ATF1	This particular procedure was well supported by the system.
17	ATF2	The right amount of information was presented in each step.
18	ITF1	The support given by the system allowed me to perform the procedure to a higher standard than without.
19	ITF2	The information provided by the system was appropriate for my level of expertise.
20	BI2	I will always try to use AR & WT in my daily work.
21	UF1	Please choose your usage frequency of AR/WT.

Each participant would be given a briefing of the trial they were about to undertake, understand the consequences of giving their consent and complete a number of initial forms, also acknowledging their contribution of anonymised data to the project. Following this, they would be given an introduction to the AR headset and invited to complete the built-in interaction training process, which would teach them how to interact with the device and the holograms shown in the workplace. Since the mode of interaction—air-tapping—and general form factor were almost always entirely novel to the participants, this was a necessary and valuable step. A member of the project team would be available to answer questions or help out if the system was not responding as expected. The relevant activity was then presented to a trainee, giving them guidance at each step of the activity. Each of these so-called "task stations" would contain a set of augmentations, representing spatio-temporally relevant information. A detailed description of the data collection process can be found in Chapter 3.



	ITF1	ITF2	ATF1	ATF2	IOP1	IOP2	IOP4	IS7	ATU4	CSE4	PE4	PE8	PE10	EE2	FC1	IMG1	HM2b	HT2	SI1	BI2
Mean	5.48	5.89	5.34	5.73	5.36	3.48	5.28	5.15	5.54	4.73	4.98	5.51	5.26	5.58	5.23	3.75	5.97	3.66	4.20	4.10
Std.Dev	1.37	0.94	1.22	1.03	1.24	1.55	1.15	1.41	1.22	1.57	1.48	1.20	1.37	1.11	1.36	1.60	1.14	1.94	1.44	1.63
Median	6.00	6.00	6.00	6.00	6.00	4.00	6.00	6.00	6.00	5.00	5.00	6.00	6.00	6.00	6.00	4.00	6.00	4.00	4.00	4.00
Skewness	-1.20	-1.34	-0.97	-1.45	-1.22	0.17	-0.70	-1.06	-1.19	-0.71	-0.76	-1.09	-0.75	-1.36	-1.01	0.00	-1.66	0.12	-0.28	-0.23
Kurtosis	1.06	2.36	0.83	2.36	1.26	-0.89	0.36	0.81	1.65	-0.35	-0.07	1.23	-0.26	2.20	0.52	-0.73	3.59	-1.28	-0.02	-0.71

Figure 40: Responses from participants in the second study (N = 157).

Having completed the activity using the technology, the participant would complete a number of questionnaires, including one on technology acceptance. Participants who were assigned to the control group would perform the activity without AR support and, while this was useful in providing a baseline for other studies, it provided no benefit to an investigation of technology acceptance.

The responses to the questionnaire demonstrated a generally positive sentiment towards the technology. Item UF1, denoting the actual usage frequency of the technology was a clear and easily explainable outlier. AR headsets of the type used in this study, even at the time of writing, are not available to retail customers. For this reason, the focus was instead placed on understanding which elements were correlated with the intention to use the technology (BI2).

A box plot showing the mean, median, quartiles, and range of the data for the second set of trials can be seen in Figure 40. It shows a generally positive response to the technology, in particular in areas related to positive affect (HM2b) and the anticipation of using it in the future (ATU4). Curiously, the two negatively framed questions, HT2: "I am addicted to using AR" and IOP2: "I worry I could become dependent on a single AR/WT supplier", produced similar distributions of responses, though HT2 had an additional grouping at the low end of the scale (ranks 1 and 2). Generally the responses show light-tailed distributions (kurtosis values much less than 3) as compared with normal distribution, with the notable exception of HM2b: "I like working with AR/WT", whose leptokurtic value of 3.59 is most likely an artefact of the item's high mean (5.97) within a 7-point Likert scale.

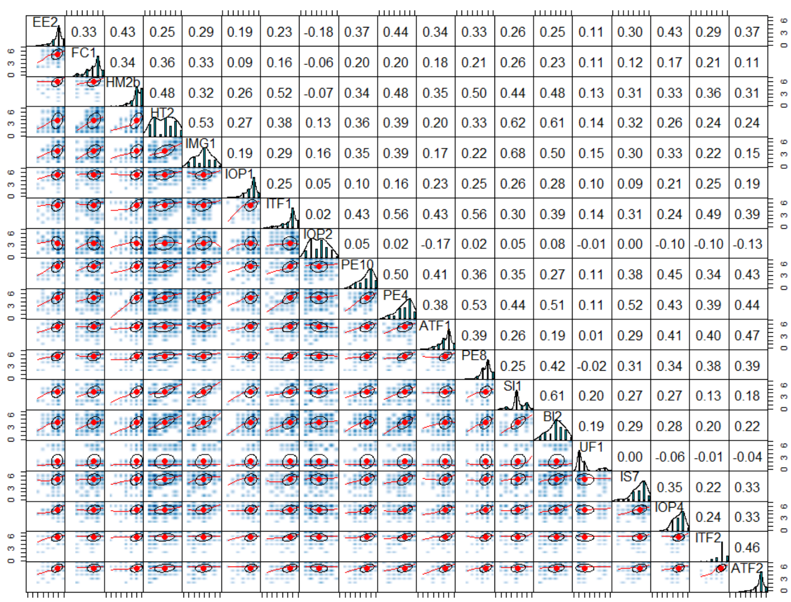


Figure 41: Correlations for the extended AR technology acceptance model.

The inter-item microstructures within the data, also called the degree of lumpiness, can be observed in a correlation plot, shown in Figure 41, which shows Spearman correlation coefficients, correlation ellipses, and trends for each variable pair. Along the diagonal are histogram plots, showing the distribution of responses to each item. These charts, although small, demonstrate clear distributions within most observables, aside from the deviations noted above. While not descriptive in itself, this provides an important sense check at this stage of data aggregation.

6.2.3 Model Fitting and Optimisation

The starting point for modelling these constructs is the final model from the first iteration, seen in Figure 36. The updated questionnaire made two changes to the model and one addition. The inclusion of IS7 re-distinguishes information security (IS) as a measurable variable. Interoperability saw item IOP4 join the list of observables. Both of these changes are easily represented as additions to the model, and a connection is made between the IS construct and behavioural intention (BI), differentiating the notion of trust from the social influence (SI) construct, which serves to examine the social pressure felt to use the technology. The AR Fit construct, however, now has more granularity and, as a result, more conceptual power. The ability of this new distinction—between activity-technology fit (ATF) and individual-technology fit (ITF)—to describe features in the data and the relevance of both to the acceptance model as a whole is the topic of much of this section. A visualisation of the initial, conjectured first model can be seen in Figure 42, which can also be represented as a set of testable hypotheses.

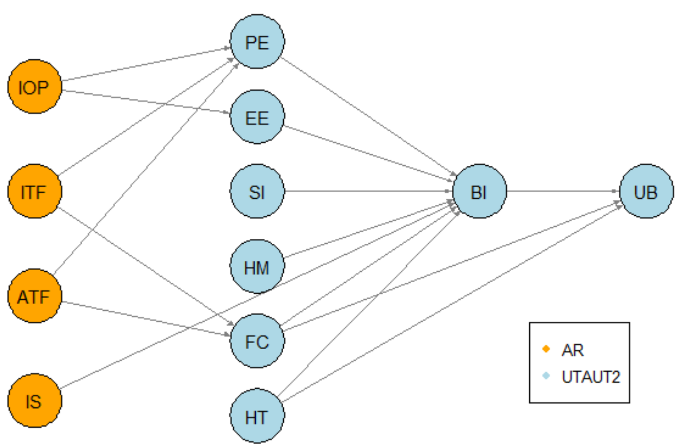


Figure 42: Second iteration acceptance model: initial structure.

In this path diagram, we are especially interested in the connections between the AR construct and those of UTAUT2. Following the outputs of the first iteration, connections between the AR constructs ITF and ATF are drawn to both PE and FC, indicating possible permutations when these new constructs are included. The paths relating to IOP are unchanged, connecting to PE and EE, whereas that from IS is new, and a regression from BI is included. These seven paths represent specific hypotheses and are shown in Table 17, along with more general assertions about the role of the AR constructs in predicting overall acceptance of the technology.

Table 17: Hypotheses for the second acceptance study

#	**Hypothesis**	**Interpretation**
H1	IOP is positively related to PE	Interoperability is a predictor of expected performance gains.
H2	IOP is positively related to EE	Interoperability predicts perceived ease of use.
H3	ITF is positively related to PE	Individual fit predicts performance gains.
H4	ITF is positively related to FC	Individual fit acts as a facilitating condition.
H5	ATF is positively related to PE	Activity fit predicts the degree of expected performance gains.
H6	ATF is positively related to FC	Activity fit acts as a facilitating condition.
H7	IS is positively related to BI	Information security is a predictor of behavioural intention to use AR.
H8	IOP is predictive of acceptance	Interoperability is a significant factor in determining AR acceptance
H9	ITF is predictive of acceptance	Individual fit is a significant factor in determining AR acceptance.
H10	ATF is predictive of acceptance	Activity fit is a significant factor in determining AR acceptance.
H11	IS is predictive of acceptance	Information security is a significant factor in determining AR acceptance.

With this structure in place, confirmatory factor analysis was used to determine the quality with which the model explains variance within the data. The initial run, which converged after 160 iterations, showed a moderate fit with a chi-square p-value of 0.003. The metrics for goodness of fit are the same as those used in the first iteration and include both relative and absolute fit indices. The set of absolute metrics included the (robust) comparative fit index (CFI), the Tucker-Lewis index (TFI), and the McDonald fit index (MFI), while the relative fit indices comprised the non-centrality index (RNI), the root mean square of approximation (RMSEA), and the standardised root mean square (SRMR). A justification and discussion of the various cut-off values considered to be indicative of a well-fitted model can be found in an earlier publication [43]. The initial goodness-of-fit can be seen in Table 18.

Table 18: Initial fit indices for the second iteration's model

Index	CFI robust	TLI robust	RNI	MFI	RMSEA robust	SRMR	DoF	Chi-Sq.
Value	0.941	0.918	0.922	0.777	0.054	0.056	122	201.2

To begin the optimisation of this model, a reliability analysis was carried out using the 'alpha' function, part of the *psych* R package. Guttman's Lambda 6 (G6), a measure of the variance of the errors, improves (0.837 to 0.899), as does Cronbach's alpha (0.815 to 0.874). These values and positive shift indicate there is sufficient reliability in the data and support the changes made in the second iteration. The same package also provides a measure for the fit between the modelled correlations and the off-diagonal elements of the matrix. This value was not largely affected by the change in construct design, indicating that the change is positively affecting one region of the data space, without affecting the overall descriptive power of the data, since values for both alpha and G6 are sensitive to the presence of microstructures within the data [94].

Performing a per-item reliability analysis can also be illuminating in identifying areas of low descriptive power or weakness within the model. In this case, the item IOP2 stood out as problematic. The overall reliability score increased (from 0.87 to 0.88) on the condition of dropping this item. Additionally, where the average reliability for the entire dataset was around 0.55, IOP2 produced the lowest standardised score of 0.21. Besides UF1, a necessary inclusion in the model (discussed earlier), no other item had a value of less than 0.43, suggesting that this item was not good [18]. The score for item overlap and scale reliability was also low, at 0.13. These values are shown in Table 19. The correlation matrix for this item similarly showed very little relevance to other items, leading to the removal of this item from the model.

Table 19: Per-item reliability scores

item	n	raw.r	std.r	r.cor	r.drop	mean	sd
ATF1	289	0.60	0.61	0.59	0.51	5.34	1.22
ATF2	293	0.54	0.56	0.54	0.47	5.73	1.03
BI2	272	0.69	0.67	0.67	0.63	4.10	1.63
EE2	297	0.58	0.61	0.59	0.53	5.58	1.11
FC1	289	0.45	0.43	0.38	0.35	5.23	1.36
HM2b	297	0.64	0.65	0.63	0.59	5.97	1.14
HT2	287	0.62	0.61	0.59	0.56	3.66	1.94
IMG1	230	0.62	0.61	0.60	0.57	3.75	1.60
IOP1	274	0.46	0.46	0.41	0.38	5.36	1.24
IOP2	281	0.25	0.21	0.13	0.13	3.48	1.55
IOP4	240	0.52	0.50	0.46	0.41	5.28	1.15
IS7	271	0.43	0.44	0.39	0.35	5.15	1.41
ITF1	291	0.69	0.69	0.68	0.62	5.48	1.37
ITF2	292	0.51	0.54	0.50	0.44	5.89	0.94
PE10	293	0.61	0.61	0.59	0.54	5.26	1.37

PE4	287	0.71	0.72	0.72	0.66	4.98	1.48
PE8	293	0.64	0.66	0.64	0.58	5.51	1.20
SI1	241	0.71	0.67	0.66	0.64	4.20	1.44
UF1	272	0.27	0.27	0.20	0.19	1.83	1.57

With this change we may re-inspect the fit indices to see what impact it has. The new values are shown in Table 20, demonstrating that the fit has improved according to all but one metric, RMSEA, which changed from 0.054 to 0.055. The number of degrees of freedom in the model is reduced by 17, and the Chi square value falls to around 176, or around 147 for a. The p-value for this model increased to 0.005, still suggesting that there is some misfit. Interestingly, many of the fit indices show better results than their non-robust counterparts, suggesting that data non-centrality may have an impact on the data. Indeed, the assumption of normally distributed data is one that is often relied on to form conclusions and, while steps have been taken to account for this, such as the use of the Santorra-Bentler correction, some residue of this may remain.

Table 20: Fit indices, following the removal of IOP2.

Index	**CFI robust**	**TLI robust**	**RNI**	**MFI**	**RMSEA robust**	**SRMR**	**DoF**	**Chi-Sq.**
value	0.948	0.924	0.930	0.797	0.055	0.052	105	176.1

While the new values indicate a better fit than the initial structure, they suggest that further improvement is possible. A complementary approach of theoretical review of the paths, together with modification indices, is used to make further analysis, specifically the connections from ITF and ATF to the UTAUT2 core structure.

Reviewing AR Fit model paths

Until now we have assumed and hypothesised that both elements of AR Fit are predictors of PE and FC. Before looking at the data related to these connections, it is important to consider the theoretical underpinnings and which changes are reasonable. From there, we may inspect the modification indices and test variations.

First, recall that ITF is the degree to which the system is perceived to support individual competence, so let us consider this in terms of both PE and FC. On the one hand, it is reasonable to think that, where a higher degree of competence support is perceived, greater performance benefits are expected? Similarly, when competence is effectively supported by the technology, could we reasonably anticipate a greater sense of being well equipped to perform the task? In both cases, I would argue the answer is ‘yes’, the better the adaptation to a person’s needs, the more they will get out of it, or at least expect to.

In terms of ATF, the picture is less clear. The perception that the technology is well-suited to the task may not necessarily correlate with a belief that higher task performance can be accomplished. If, for example, you are a novice attempting to perform a complex task, such as the assessment of a Mars rover's solar panels, the belief that the system is capable of transmitting relevant information (itself relying on a number of assumptions) may not impact the degree to which you believe that you will perform the inspection to a high standard. This connection is made more tenuous if the person is inexperienced in the use of the technology, something that was most certainly the case in this study. When we look at the connection between ATF and FC, there is a more natural association, since FC pertains to the degree to which the individual considers themselves to be well-equipped to undertake the task. If there is the perception of a good fit between technology and activity, it follows that this would bolster this notion, as there would be additional tools at their disposal.

Based on this, the connection between ATF and PE was removed and the model was retested, leading to the measures of fit shown in Table 21. A small improvement in fit can be seen, causing the CFI to reach "good fit" territory. Additionally, the Chi-square p-value after this change increased to 0.006, moving marginally outside the realm of strong statistical significance, marking a reduction in the certainty of model misspecification. In this precise, but rather convoluted, phrase, it is better for the p-value not to be statistically significant.

Table 21: Fit indices, following the disconnection of PE to ATF.

Index	**CFI robust**	**TLI robust**	**RNI**	**MFI**	**RMSEA robust**	**SRMR**	**DoF**	**Chi-Sq.**
value	0.950	0.927	0.930	0.800	0.054	0.052	106	176.2

Using Modification Indices

We use modification indices to help us understand where there may be remaining flaws in the model. They work by perturbing the model in a piece-wise fashion, looking at how the model fit would change when paths are added or an existing constraint is freed. This allows for a more fine-grained inspection of the model components, though it is recognised that their use should be accompanied by theoretical insight. As a guide, modification indices (MIs) with a value greater than 3.84 suggest that the model would be improved by the respective change, corresponding to a p-value of 0.05. Values larger than 10.83 are equivalent to a p-value of 0.001. The approach used here is to take into account the expected parameter changes (EPC), the power of each MI, and its statistical significance. It is based on the approach put forward by Saris, Satorra, and van der Veld [103]. In this paper, the authors suggest using an absolute value greater than 0.4 for factor loadings and greater than 0.1 for correlated errors, which are used here.

Beginning with the correlated errors and using statistically significant MIs ($>$ 3.84), we may inspect the standardised results, which are shown in Table 22. The second and third lines, showing correlations from ATF and ITF, respectively, to EE, are marked as they stand out from the others. Power values less than 0.75 are generally 'low'. These relationships represent latent variables in the table (the others are exogenous), have moderate powers, and have significant MI values. Their EPC magnitude, however, is low, meaning that the enforcement of this constraint would not lead to a significant change in the relationship. Their non-centrality parameters (NCP), a measure of the model-consequent chi-square deviation, are also high.

Table 22: Correlation error modication indices.

lhs	op	rhs	MI	epc	sepc.all	ncp	power	decision
ATF1	∼∼	SI1	10.130	0.230	0.400	1.860	0.280	**(m)**
ATF	∼∼	EE	9.760	-0.140	-0.200	4.640	0.580	**(m)**
ITF	∼∼	EE	8.850	0.140	0.160	4.850	0.600	**(m)**
IOP1	∼∼	PE4	7.220	-0.240	-0.260	1.290	0.210	**(m)**
ITF1	∼∼	EE2	7.210	-0.220	NA	1.460	0.230	**(m)**
ATF1	∼∼	PE4	6.980	-0.230	-0.300	1.280	0.200	**(m)**
IOP4	∼∼	ATF1	6.650	0.200	0.230	1.630	0.250	**(m)**
ITF2	∼∼	FC1	6.330	0.220	NA	1.320	0.210	**(m)**
IMG1	∼∼	HM2b	5.480	-0.200	NA	1.330	0.210	**(m)**
ITF1	∼∼	ATF2	4.740	-0.180	-0.270	1.400	0.220	**(m)**
IOP4	∼∼	ITF1	4.500	-0.180	-0.180	1.330	0.210	**(m)**
ITF2	∼∼	SI1	4.490	-0.120	-0.250	2.990	0.410	**(m)**
ITF1	∼∼	PE8	4.400	0.170	0.190	1.500	0.230	**(m)**
ATF2	∼∼	PE4	4.000	-0.150	-0.260	1.880	0.280	**(m)**

When dealing with factor loadings, the value for ‘delta’ was set to 0.4 in lavaan’s ‘modicationindices’ function, meaning that changes with factor loading less than this value are ignored. This number is used in the calculation of the ‘power’ value, shown in the table. It should be noted that this relies on an unstandardised ECP value, while the values shown are standardised, though it was observed that the change through standardisation was small. The use of this particular cut-off for theoretical factor loading follows Saris, Satorra, and van der Veld [103]. Only significant MIs with a power of greater than 0.2 are shown. Table 23 shows the results, ordered by MI value. The first entry has a combination of a high MI value (38.19), a high EPC (0.88), and a high power (0.802), suggesting that this is an item that we need to pay attention to, a fact asserted in the ‘decision’ column.

Table 23: Factor loading modication indices.

lhs	op	rhs	mi	epc	sepc.all	ncp	power	decision
EE	=∼	FC1	38.190	0.880	0.730	7.890	0.800	*epc:m*
PE	=∼	ATF2	19.060	-1.540	-1.820	1.290	0.210	**(m)**
ATF	=∼	ITF2	13.390	0.840	0.730	3.030	0.410	**(m)**
ITF	=∼	PE4	6.690	-0.670	-0.420	2.380	0.340	**(m)**
PE	=∼	IOP4	6.330	0.780	0.850	1.640	0.250	**(m)**
ITF	=∼	PE10	4.540	0.530	0.360	2.560	0.360	**(m)**
PE	=∼	IOP1	4.140	-0.590	-0.610	1.880	0.280	**(m)**

The first modification index, connecting EE to FC1, suggests that were we to load this observed exogenous variable onto effort expectancy, there would be a benefit to model fit. More broadly, this suggests that there is some crossover between the latent variables EE and FC. The two correlation errors mentioned above, both of which relate to EE and factors that load onto FC, also suggest that there is information connecting these elements that is not being captured in the current model. Framing this in terms of the theoretical constructs, the implication is that a person's sense of how well-equipped they are for the task (when using AR) overlaps with an estimation of how much effort is required to complete the task with the support of the system. While a broad and potentially multi-faceted statement, it is certainly plausible. Without performing additional studies, we may represent this missing information by adding a residual correlation between FC and EE. The impact on the quality of model fit after this change is shown in Table 24.

Table 24: Fit indices, after adding a residual correlation between EE and FC.

Index	**CFI robust**	**TLI robust**	**RNI**	**MFI**	**RMSEA robust**	**SRMR**	**DoF**	**Chi-Sq.**
value	0.957	0.937	0.939	0.822	0.050	0.051	105	166.484

This third change leads to an improvement in all measures of goodness-of-fit, bringing both CFI and TLI past their cut-off values and reducing the Chi-square value to around 166. The p-value for this metric also increases to 0.013, clearly leaving significant territory and indicating that the model is less misspecified than before. Further inspection of the MIs after this change showed, for correlated errors, none with a power greater than 0.3. For factor loadings, the modification indices were almost all versions of the regressions between latent variables already in the model (such as BI= PE10) and acted to confirm those connections. The exception was a theoretical loading of ITF2 onto the ATF construct, suggesting that there might be additional information that connects these constructs, a topic that will be mentioned in the next chapter, but did not represent a basis for further changes to the now well-fitted structure.

Final Model

Through the use of goodness-of-fit indices, modification indices, and theoretical review, an optimised structural equation model has been produced. The final step is to interrogate this model to see which factor loadings and paths are significant in their representation of the data and, in doing so, provide results for the hypotheses generated above. Of the 18 observables included, 11 showed significance within the model. For all of these, shown in 25, the p-value was extremely small—less than 0.00001—so this column is not included. The reason for this is most likely because only a small number of exogenous variables inform each latent construct.

Table 25: Significant factor loadings for the final acceptance model.

lhs	op	rhs	est.std	se	z	ci.lower	ci.upper
SI	=~	SI1	0.891	0.032	28.038	0.829	0.953
PE	=~	PE4	0.807	0.045	18.096	0.720	0.894
ATF	=~	ATF2	0.752	0.053	14.143	0.647	0.856
SI	=~	IMG1	0.746	0.045	16.517	0.657	0.834
ATF	=~	ATF1	0.667	0.069	9.621	0.531	0.803
PE	=~	PE10	0.653	0.061	10.759	0.534	0.772
PE	=~	PE8	0.653	0.061	10.730	0.534	0.773
ITF	=~	ITF1	0.653	0.056	11.613	0.543	0.763
ITF	=~	ITF2	0.535	0.063	8.520	0.412	0.658
IOP	=~	IOP4	0.503	0.062	8.173	0.382	0.624
IOP	=~	IOP1	0.407	0.089	4.593	0.233	0.581

This information gives us a set of weights that give the latent variables their predictive power. In Figure 43 they are shown graphically, along with the significant regression in the model. Note that this figure does not give a complete picture of the structural equation model, but rather is an overview of the statistically significant elements of the prior data analysis.

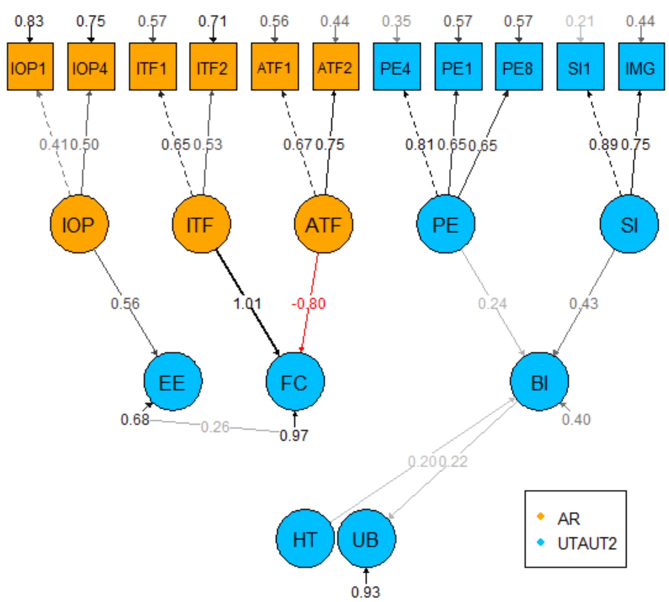


Figure 43: Final technology acceptance model: all significant paths.

It can be seen from Figure 43 that there are two unconnected groups of variables. Three of the four AR constructs connect to EE and FC, while five other UTAUT2 constructs form another arrangement. This will be discussed in more detail in the next chapter, but it is important to note that the absence of connections between these groups is not evidence of a disconnect. The regressions shown between the latent variables can also be seen, alongside non-statistical significant relationships, in Table 26.

Table 26: Final model: regressions

lhs	**op**	**rhs**	**est.std**	**se**	**z**	**p-value**	
EE	~	IOP	0.563	0.074	7.577	3.53E-14	***
BI	~	SI	0.432	0.115	3.766	1.66E-04	***
FC	~	ITF	1.005	0.326	3.087	2.02E-03	**
UB	~	BI	0.221	0.079	2.796	5.17E-03	**
BI	~	PE	0.238	0.101	2.355	1.85E-02	*
FC	~	ATF	-0.802	0.351	-2.284	2.24E-02	*
BI	~	HT	0.197	0.093	2.111	3.48E-02	*
PE	~	IOP	1.879	1.051	1.787	7.39E-02	
BI	~	HM	0.089	0.065	1.365	1.72E-01	
UB	~	FC	0.088	0.066	1.347	1.78E-01	
PE	~	ITF	-0.940	1.051	-0.895	3.71E-01	
BI	~	EE	-0.048	0.062	-0.767	4.43E-01	
BI	~	FC	-0.031	0.061	-0.514	6.07E-01	
BI	~	IS	-0.021	0.054	-0.382	7.03E-01	
UB	~	HT	0.017	0.096	0.176	8.61E-01	

The first seven rows in the table show, to varying degrees, significant predictive ability within the model. IOP is the most certain, explaining around 56% of the variation of the EE construct. From the AR constructs, two other conclusions can be drawn. First, ITF works as a direct predictor of FC, strongly co-varying with it. Second, there is a negative relationship between ATF and FC, where ATF is, inversely, predicting around 80% of the variance there. Both the unity power relationship from ITF and the counter-indication from ATF suggest that both are correlated with FC, but in different ways. Where ITF is reported as stronger, there was a greater sense of being well-equipped for a task. A stronger ATF, however, was more likely to be accompanied by the opposite perspective on 'well-equippedness'. This is an unexpected, but interesting, feature and will be discussed in the next chapter.

The other significant regressions relate to the UTAUT2 core model. Variance in BI is seen to be correlated with both PE, HT, and SI, with the first two explaining around 20% of this construct each. The connection between SI and BI is stronger; it is more significant and explains around 43% of the variance of this fundamental attribute. Returning to the overall model, Figure 44 shows the significant regressions overlaid.

We may now test the hypotheses. A summary of the outcomes is shown in Table 27. In this case, only one hypothesis was rendered null through optimisation; three item-specific hypotheses were accepted, while another three were rejected. Two of the AR constructs are seen as important within the larger

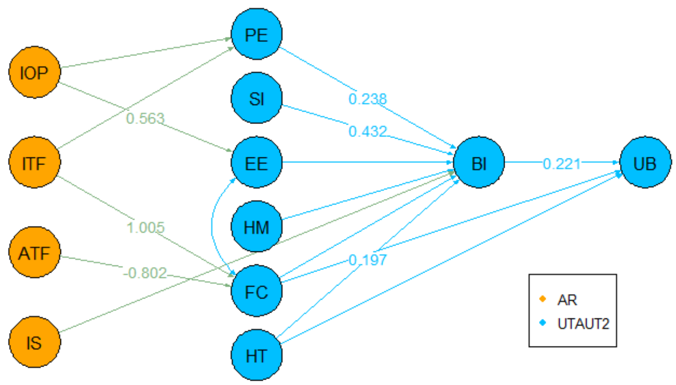


Figure 44: Final technology acceptance model. Significant path estimates are marked with coefficients.

sense of technology acceptance; one (IS) is not, and the remaining one (ATF) is not well understood but provides an intriguing route for further investigation.

Table 27: Results of second-iteration hypothesis testing.

#	**Hypothesis**	**Outcome**	**Reason**
H1	IOP is positively related to PE	Rejected	With these data, IOP cannot be said to predict performance expectancy.
H2	IOP is positively related to EE	Accepted	IOP is, again, found to be a good predictor of effort expectancy.
H3	ITF is positively related to PE	Rejected	There was no significant correlation between these items.
H4	ITF is positively related to FC	Accepted	ITF was a powerful and significant predictor of facilitating conditions.
H5	ATF is positively related to PE	Null	Construct removed in optimisation.
H6	ATF is positively related to FC	Accepted	There was a fairly significant, but negative, predictive power.
H7	IS is positively related to BI	Rejected	The covariance between these constructs was very small and not significant.
H8	IOP is predictive of acceptance	Accepted	The repeated presence of this strong predictor suggests this construct is valuable in understand acceptance.
H9	ITF is predictive of acceptance	Accepted	The strong predictive influence on FC indicated that this construct has value in the wider acceptance model.
H10	ATF is predictive of acceptance	Inconclusive	While showing predictive strength, more information is needed to understand the role of this construct.

H11	IS is predictive of acceptance	Rejected	This construct did not demonstrate predictive power within the model.

6.2.4 Challenges and Constraints

The implementation of the system in three separate domains was a significant technological and organisational challenge. The nascent nature of the hardware, combined with the need for bespoke software (itself built on emerging systems) meant that, in order to solve many of the functional issues, regular communication with (and sometimes the presence of) a developer was needed. While more of a logistical challenge, it was nevertheless an important factor that needed consideration through the implementation process.

Environmental factors, especially temperature and lighting, affected both device- and component-level functionality. It was discovered, for instance, that with a high ambient temperature (> 30 °C), the microphone's functionality was impaired while running the application, but that this could be remedied by connecting the HoloLens to a power bank during operation. Rooms with large fluorescent lights would also cause chromatic aberration in the display of holograms and sometimes cause the device's tracking system to fail, though this could be mitigated to some degree with careful configuration of the application (in particular, settings related to the spatial mapping and stabilisation planes).

Some time constraints were present during the set-up and recording of workplace procedures. While additional peripheral hardware was used (and is described later), the pre-emptive need for recorded activities meant that the software was restricted to controlling only those sensors within the HoloLens. This limitation was also seen in the number of statements that could be considered meaningful when probing the trainers' acceptance of the technology, leading to a reduced data set for evaluation. Again, later trials with learners were able to incorporate additional metrics.

6.3 Summary

Over two iterations, this chapter has demonstrated the construction and optimisation of a technology acceptance model for immersive technology. The technology acceptance model, validated and interrogated by an array of statistical tools, has not only tested hypotheses but also provided insight into the types of hypotheses that would add information to the existing model. The process of model optimisation through the combination of theoretically grounded assertions and data-driven statements offers a methodology for continued, iterative development. Constructs focused on understanding the relationship between individuals, tasks, and technology provided strong support for their utility in this area of research. In the next chapter, these relationships will be explored further.

Chapter 7: Discussion

This thesis has sought to differentiate and understand some of the key considerations when applying augmented reality and wearable technology to workplace training. It first looked at how instructional design methods (IDMs) can aid the training process and at how propositional affordances can be translated into AR features that both engage and inform the user. Body sensor networks and, to some degree, the capture of torso and arm movements were integrated into the framework to support the construction of embodied affordances. Finally, the use of structural equation models allows for an inspection of the meaning-making process itself, looking at various attitudes that drive the adoption of the technology.

This chapter, while still focused on these developments, takes insight from a broader set of sources and seeks to motivate and operationalise the findings so that they can be interpreted and used by the wider community, including both academics and practitioners. In service of this amalgamation, there is first a presentation of Social Systems Theory, a wide-ranging conception of system interaction with parallels in immersive technology. Following this, there is a discussion around the processes of designing and implementing propositional affordances. On the topic of embodied knowledge and the use of physical movement in AR, this chapter takes a look at how connection protocols and sensors can be improved to support future work. Section 7.4 addresses the results of the technology acceptance model and operationalises the statistical findings from the previous chapter. Finally, there are some more speculative notes on how the components could form part of a wider, integrated framework.

7.1 Theoretical Framework

Two underlying theories have supported the work reported in this thesis: - Social Systems Theory and Affordance Theory. Both are relational in nature in that they frame the training process in a way that places importance on the interaction between the learner and the environment. Since it is this immediate experience that the technology seeks to augment and, through a process of intermediation, provide greater opportunity for improving expertise, these theories are extremely relevant to the design of systems used to capture data and display information.

Social Systems Theory

Social systems theory, developed by Niklas Luhmann (1927–1998), is an expansive framework for describing system definition and interaction. It envisions systems—of any kind, but in this case, the person, the activity, and the technology—as operating within an overlapping possibility space, where each is capable of affecting the other in a myriad of ways.

A central idea within this theory is the importance of *functional differentiation*, where it is the operations that each system performs (within itself and towards its environment) that are most significant. In this context, the trainee is an operationally closed system separate from, but contained within, a learning environment. Each person is said to possess three systems: their biological system, a psychic (here, 'cognitive') system, and a communication system. Luhmann points out that the cognitive system, while useful for sense-making, does not communicate with other brains directly but relies on a communication system to do so. For expertise held by one person to be given to another, it must use the communication system via a complex set of interdependencies. In the same way, practical competence is not communicable simply by watching it happen, but by enacting the task itself and sensing the result.

It is important to note that the theory explicitly rejects the idea of a single nexus or mechanistic causation from one system to another. Just as the immune system within the body is comprised of many different functions working in coordination (and yet a disruption in one can affect the entire body), so the cognitive and communication systems are multifaceted while each forms a functionally distinct system. The interfaces between these systems are referred to as their *structural coupling*—a multitude of overlapping effects and interrelations that determine how one affects another given, of course, their shared environment.

Take, as an example, a classroom with a teacher and several students. The teacher will have a lesson plan, and the desired outcome is that the students learn. Since we cannot access the thoughts of the teacher or learners, we must interpret their intentions and knowledge through the media of speech, movement, gesture, writing, and so on. Thus a dialogue occurs between, say, a student and teacher. This communication reveals something about each person's intentions towards the other and (hopefully) establishes some understanding regarding the didactic model. In response to this communication, the teacher, student, classroom, or didactic model may change as a result of new contingent information. It may become apparent that the two speak different languages or that the classroom is on fire. While these are unlikely events (themselves contingent on a very particular set of circumstances), each would necessitate significant changes to the entire situation, perhaps by including an interpreter or relocating the teaching activities altogether. Under more typical circumstances, teaching would progress, though the nature of the instruction would develop to suit the needs of the students and, perhaps, to make better use of the classroom itself.

In order to address the sense-making of constructivist learning, we must consider the cognitive system that constructs this meaning. In Luhmann's words, "cognition is the realization of combinatorial gains on the basis of the differentiation of a system that is closed from its environments but nonetheless 'contained' in that environment" [91]. Unpacking this definition, we may outline three events that occur that can be said, together, to construct meaning through cognition. First, there is a distinction between those things that are of interest and those that are not. He terms this the 'operation of selection', and it relates to the physical or conceptual lines that are drawn to distinguish an object (of observation) from its environment. Second, given that the object is connected to its environment, we recognise these *structural couplings* and develop an understanding of the significance of the thing. Lastly, our 'realisation of gains' is not only the apprehension of a difference but also the application of this understanding through interaction, further observation, or reflection.

Generally, Luhmann sees this experiential arc as moving from a place of necessity—we are driven by our needs—to one of contingency—the interaction from each system gives rise to a specific state that is resultant of, but not driven by, the states of all the systems it contains. Luhmann's theory, whilst impressive in its scope, does not easily describe the role that necessity plays when selecting a particular path for action. The second theory, however, is especially useful in this regard.

Connections with Affordance Theory

Interaction with our environment and reflection on our interactions are, to the operative constructivist, the two most fundamental processes of sense-making or cognition. We learn by inquiring about and interacting with the people, places, and things around us. We do this by taking advantage of the affordances that we find, either as physical attributes or as informational cues. In reflection, or self-observation, we take stock of these inquiries and interactions and develop an understanding that helps us tackle a greater range and complexity of situations. In doing these things, we realise (in both senses of the word) combinatorial gains, improving our understanding and, when there is an element of practice, our embodied competence.

Here there are direct parallels with Luhmann's work, which identifies the communication—in this case, the information specifying the affordance—as the functional unit of the system. An affordance may benefit sense-making, or it might prompt a particular posture or manner of action that aids the learner in a task. In either case, the unit of change relates to the affordance, and both the quality of the interpretation and action determine the outcome.

Gibson does not expand greatly on tools as a particular class of object but does remark that, when used, a tool becomes "a sort of extension of the hand, almost an attachment to it or a part of the user's own body and thus no longer a part of the environment of the user" [36, p.41]. Given that an affordance is also defined as an "invariant combination of variables" [35, p.58], that is, that their affordances do not change when we start using them, we can interpret this as a shift from an object that possesses *opportunities for action* to a tool that possesses *opportunities for use*. When we use it in our environment, we have new capabilities and, by extension, a modified set of competences.

7.2 Designing Propositional Affordances

Chapter 4 demonstrated the use of AR and wave-guide technology, to integrate digital content into our physical environment. These tools allow for the precise localisation of the device and digital content relative to the person or their environment. The WEKIT Experience Capture System (ECS) is an implementation of this technology that allows trainers to generate a standardised description of the augmentation used to enhance both the procedure and the workplace while engaged in the task at hand. This description is captured in the Augmented Reality Learning Experience Model (ARLEM) standard, making the content transferable between both devices and workplaces.

The chapter focused on propositional knowledge, or "knowledge-that", and how trainers can use AR technology to provide affordances to learners that represent this knowledge. It describes how this knowledge, together with ARLEM, can be used to specify the necessary functions of a system that can take advantage of the trainer's expertise and provide a description of their didactic choices in a standardised form. The development of the system took a number of technological constraints and educational affordances into account, using affordances to connect learning objectives to software functionality.

Rather than look again at the specifics of the activities that we carried out with the use of AR, I will instead look at the general considerations when constructing learning affordances. There are, I suggest, two important concepts that can support this work: the relevance of the instruction given by the system (both to the trainee and the task) and the person's competence with regard to both the task at hand and the use of technology. Figure 45 illustrates the interplay between person, instruction, and task, showing how the overall goal of 'augmented performance' is accomplished using, and framed by, these notions. The dotted line represents the AR affordance boundary; person and technology working in tandem towards the task's goal.

The relevance of the instruction to the trainee determines whether they will recognise that the information provided is of use. If, for example, the message provided is in the wrong language or is far too small, it immediately loses relevance. This becomes a more subtle determination if we work with icons or symbols that represent a particular action. Design considerations when creating a library of glyphs or 3D models are clear examples. Since we are working with spatially defined elements, their position also determines their relevance; if an instruction targets a precise location, misalignment may lead to the wrong object being highlighted. Similarly, if small items within the space are to be marked out, there is the risk of obscuring the object of interest with the instruction itself. All of these failure modes ultimately cause the instruction to lose relevance, diminishing the amount of benefit the system can provide.

The second consideration, the trainee's competence in accessing the instruction, refers to the capability of the person to interact with the instruction. In simpler cases, where information must be read or listened to, this amounts to a physical ability to engage with it. It may be important to consider if physical handicaps or disabilities can impact this or if the environment prohibits certain

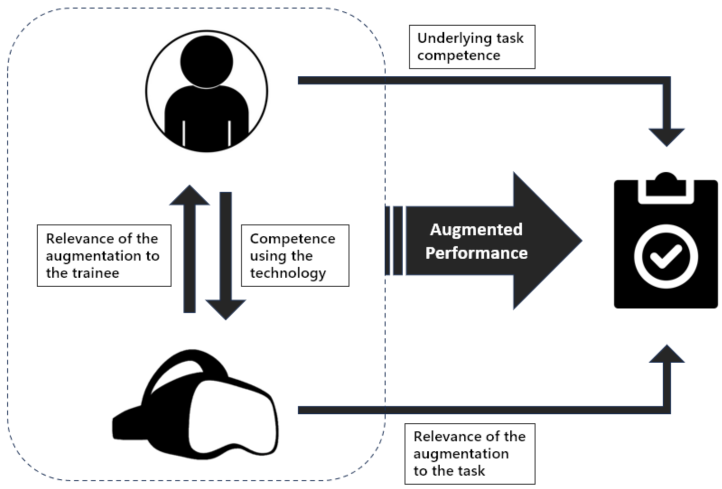


Figure 45: Affordance design considerations.

modes of interaction, such as a noisy factory floor. These kinds of limitations are generally overcome with a better understanding of the environment in which the technology is used and the people who are using it.

The fact that someone is inhibited during their use of the AR system is only half the story. A lack of competence may also arise from a lack of understanding about how interaction happens, such as not knowing how to use the air-tap gesture. It was found, for example, that performing this gesture with an open hand (on the first generation of HoloLens) meant it was frequently not registered. Other interactions, such as moving objects around, also improved with some instructional guidance. As the variety of interaction methods and styles increases, effective communication of how and what can be accomplished with gesture will likely play a more significant role in delivering effective tools. As more complex interactions are developed, it is conceivable that effective use will only be achievable with dedicated training exercises.

The final two considerations can be considered environmental factors, as they are not under the control of those delivering the technology. The first is the pre-existing competence of the person when performing the task. This is typically developed with the pedagogical idea of scaffolding [33], where support for a task is gradually removed as a person's expertise grows. Thus, we should view augmented performance in relation to the amount of scaffolding it provides. In this case, little adaptation was needed since the trainees were assumed to be at the same level. For the aeronautic and astronautic tests, everyone who participated experienced the procedure for the first time, whereas the medical use case focused on providing an experience for undergraduate students. In situations where a wider range of expertise is represented, this consideration becomes important.

Augmented performance is also framed by the relevance of the augmentation to the task. As this work has shown, augmentations can take many forms, so assessing their coherence with any complex task is often not a simple matter. Here, since we are concerned with propositional ('knowledge that') instructions, the types of augmentations are the seven shown in Chapter 4. For each type, we might ask which type of activities benefit most from the information they contain (see Table 28). From there, we can ask how these collectively inform the concept of augmented performance in a meaningful way.

Table 28: The task relevance of different types of augmentation.

Augmentation	**Task Relevance**	**Examples**
Augmented Path	Tasks where the identification of obscured routes or selection between multiple possibilities of where to move (or position oneself) for a given task.	Hazard avoidance, perspective selection.
Directed Focus	Tasks where attention given to a particular object, state, or action improves the outcome or depth of knowledge acquired.	Using assemblies or control systems, differentiating similar objects.
Point-of-View Video	Here a first-person perspective supports understanding through task mimicry, or where the extra viewpoint allows an observer to pick out new, important information.	Patient assessment, remote operation.
Think-aloud Protocol	Also from a first-person perspective, this augmentation supports tasks where time-sensitive information is useful or it is useful to ensure the proper sequence of events.	Maintenance operations or safety protocols.
Annotations	More benefit seen when the context more closely aligns with the content, or the augmentation is placed in a meaningful pose.	Warning signs, example images supporting inspection practices.
Animated 3D Models	Tasks involving precise or complex movement of objects or where there is a strong temporal component whose communication can improve task performance.	Action related to evolving situations, multi-part assembly.
Ghost Track	Visualising even simple representations of body posture supports tasks that have a physical component or are in hard-to-reach places.	Manual handling, inspection routines.

Identifying synergies between tasks and augmentation that exist outside of personal interpretation is not an exact science. Although the aim is an objective measure of the fit between the two, the range of possible augmentations as well as the tasks that they might be applied to is enormous and increasing. I will, however, make two points about the design and use of the augmentations presented here.

First, it is important to have a clear idea of the desired learning outcome when understanding how augmentations fit together, either sequentially or in concert. From here, we can ask which AR affordances—opportunities for action in the augmented space—can be offered to the learner so that they might reach that goal with greater ease. In doing so, we are forced to consider who is using the system and must adapt our mode of support accordingly, since the action, and thus the result, arises from the opportunity, not from the instruction that informs it.

Second, keep in mind the idea of familiarity when designing an augmented work flow. For novices, this is mostly their familiarity with the technology, so interactive elements should be kept simple, and the way in which content is overlaid should be obvious and as simple as possible. When people are more comfortable with the technology, the benefit comes more from their familiarity with the task (or expertise). In this case, the goal is to match the experience to the optimal learning path. For those who are both expert practitioners and familiar with the technology, it is important to emphasise the interrelationships between the physical and digital elements of the activity. Becoming familiar with how these systems overlap will enable them to gain new insight into the best use of the technology, allowing them to conceive of new designs or iterations of augmented experiences, even if they are not software developers themselves.

This more systematic type of familiarity was a major goal of the wider WEKIT project. Not only did we, as a team, facilitate the introduction of various subject matter experts to AR and wearable technology, but we also provided, in the form of the ECS described here, tools that could tap into their expertise and, through patterns of interaction, build an augmented experience that was both nominally useful and also representative of their personal understanding. It should be acknowledged then that even the notion of 'performance augmentation' will mean different things depending on the context. Sometimes it is simply that we experience a set of tasks in a new way, gaining additional insight in the process. In other cases, it will mean heightened capabilities, a greater potential for success, or that it is simply easier to learn new things. In yet other situations, it is that this new medium can yield profound insights or radical shifts in approach or perspective. Context engineering also happens in context.

To arrive at a working definition, we can say that performance augmentation is the aggregated impact on both activity and competence arising from the apprehension and use of a set of AR learning affordances.

7.3 Supporting Embodied Competence

It is clear that our ability to perform tasks well is, at least in part, derived from the quality of the information that is presented to us and the accumulation of knowledge around a particular topic. What is less well-defined, perhaps since it is harder to quantify, is how we gain "hands-on", procedural expertise; how we become more adept with our physical movements so that the accuracy and precision with which we can manipulate tools and objects in the world around improves. Also referred to as implicit memory, know-how, or simple practical experience, embodied competence is generally acquired through repetition and refinement.

In this mode, information is felt rather than interpreted. As competence increases, the task will tend to feel more natural, and while it may still require focused attention, activities tend to have better flow and coherence. Take, for example, the use of an ultrasound probe when performing a scan. To begin with, the actions of preparing the patient, applying gel to the probe, and preparing the examination settings on the machine will be, from a purely kinaesthetic perspective, disjointed, inefficient, and over-corrective. After several sessions, the sonographer will be able to anticipate the next action and have more confidence that the patient and diagnostic environment are suitable to proceed. Their physical movement will become more fluid, attention can be given to more than one action at a time, and their physical gestures are more likely to represent their implicit familiarity with the situation.

Here, the main goal is the measurement of movement, and other biological signals, that represent both expert and learner representations within the augmented space. Since the workplace is already populated with augmentations, these data form part of a larger landscape from which features and patterns can be extracted. The task at hand, then, is one of engineering rather than design. We are concerned with factors such as modularity, comfort, ease of integration, and data fidelity. Chapter 5 demonstrated some of the channels that can be used to gather movement data. It looked at the hardware, software architecture, and low-level functionality, even down to the modification of the BLE protocol to increase data throughput. Going to such lengths is indicative of the challenge of maximising the output from the devices involved, since we are still working at the edge of what is computationally possible (for a reasonable budget).

On a more powerful system, the same adapter was tested with custom-made motion tracking and demonstrated the ability to connect to five devices simultaneously and deliver a stream of position and orientation data from each. These were successfully mapped to an avatar whose corresponding body part (lower arm, upper arm, or chest) would follow the position of the sensor. A demonstration of this in action, together with updates from two features, can be seen in Figure 46. The red line represents the elbow angle, and the blue line is the elevation of the upper arm.

Overall, it was found that inertial data can be delivered at sufficient rates to provide accurate data for positional tracking and that the on-board noise filters were good enough that features were clearly visible in the data and noise was

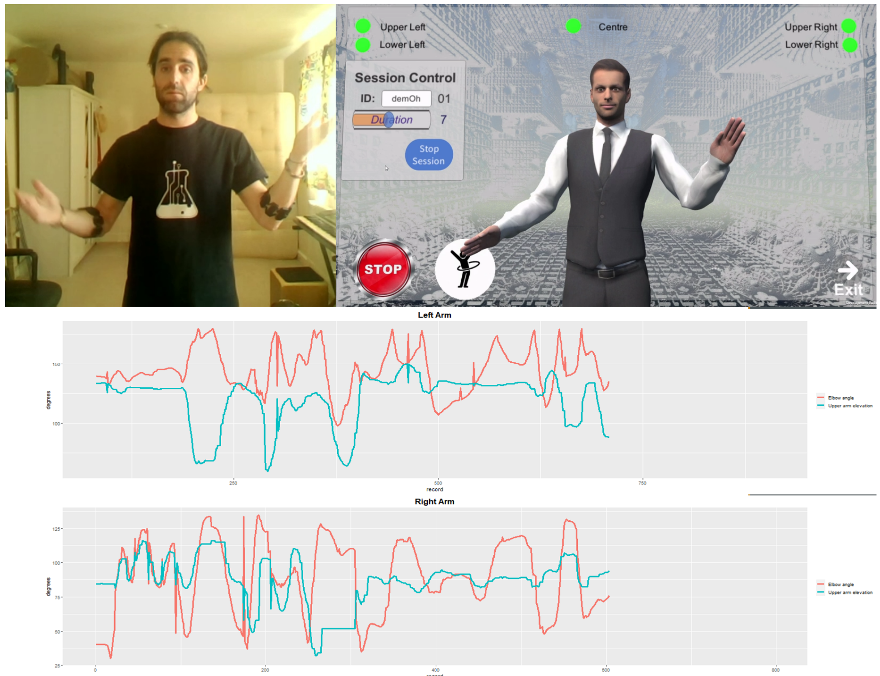


Figure 46: A demonstration of Eric, animated with five sensors.

not a significant factor, although placement of the sensor was.

The most significant technical challenge is related to accurate axis mapping in the augmented space. Distortion in the data arose from at least three sources: device localisation, drift in the sensor readings, and placement on the body. The first version of HoloLens, locating itself with simultaneous localisation and mapping (SLAM) algorithms, is precise to a centimetre scale. Individual measurements of distances with the depth camera system are an order of magnitude better, though errors are necessarily stacked, when considered overlays onto physical objects.

The same cannot be said for EMG data. Although the data could be reliably recorded from a single device over eight channels at over 200 Hz, adding a second device and attempting to clean and store a total of 16 channels, while also updating the positional system so that virtual hands could be displayed at over 25 fps, proved to be too much for the head-mounted system. In addition, even at this measurement rate, features of the EMG signal cannot be measured, so the data is the aggregation of signals from large clusters of muscle fibres. In this case, placement on the forearm is a much more sensitive issue, and while some instructional or mathematical solutions were considered, it suggested significant additional complexity in an already underperforming sensor.

Returning to our augmented learning environment, aided by a positional tracking system, we look at the connections between our newly available data streams and the task at hand. Two types of augmentation used these data to provide: the 'ghost track', a visualisation of the data from all positional sensors, and the 'hand track', which used only that from the hands and forearms. The premise behind their inclusion was that when people pay specific attention to how they position themselves or their hands for a particular task, the more quickly they will develop competence, and the skills that they learn will more closely reflect those of an expert.

The trouble with embodied competence is that you take it with you when you leave. This means that it is difficult, and in some cases impossible, to simply describe what it feels like to perform a certain task. Take, for example, the use of a torque wrench on an aircraft wheel nut. The amount of torque that you apply is, by design, capped for the task, and yet there are still incorrect ways to use it and certain postures and techniques that will make the job easier and reduce the risk of accident. Having an expert demonstration of the tool in use, especially when combined with a verbal commentary on their actions, can provide a lot of procedural know-how.

The procedural knowledge communicated by these types of augmentations, like the propositional knowledge above, will be received differently depending on the level of existing competence and its relevance to the task. Complete novices are more likely to treat it as a demonstration, after which they may attempt the same actions. People with more familiarity might use imitation to align themselves with an expert recording, noting the differences during different stages of the procedure, while those who are experts themselves may use the experience to design new modes of working or put together new forms of augmentation that can better represent the movement and sensation of the task.

7.4 Immersive Technology Acceptance

We generally use acceptance measures that work to detect positive intentions to use the technology again. Intention has long been shown to correlate with the frequency with which people use a technology, notwithstanding various biases and circumstances that can affect the degree to which it has an effect. In the case of newer technologies, especially those that are not available on the consumer market, we therefore rely on this factor in place of actual usage frequency to give a forward-looking estimation of the success of the implementation.

The model used here attempts to capture a set of (reported) beliefs about the systems' utility, ease of use, and impact on the activity. These sentiments are grouped into constructs, such as performance and effort expectancy, which reflect a position towards a particular aspect of the technology's use. Chapter 6 detailed how these constructs were defined, developed, and tested with several hundred people, aggregating a dataset that was subject to detailed statistical analysis. It used SEM to uncover patterns and links within the data that can tell us more about how people responded to using the system.

The strongest indicators of an intention to use the technology came from the technology-specific constructs that were introduced in this work. Interoperability (IOP), individual-technology fit (ITF), and activity-technology fit (ATF) all contributed meaningfully to the attitudinal structure under investigation. In both iterations, IOP was strongly correlated with effort expectancy, indicating that a robust link exists between these notions. This suggests that, where there is a stronger perception that the various elements within the technological framework work well together, the amount of effort needed to use the system decreases. This suggests that users of the technology not only pay attention to the linkages but also ascribe importance to them. This goes against notions that immersive technology users focus solely on the immediate needs of the task and supports a strategy of more in-depth user feedback.

The second observation relates to ITF, or the degree to which the technology matches the user's expertise. This construct was found to be strongly correlated with facilitating conditions; where there was a closer match, there was a sense that the technology acted as a helpful scaffold for learning. This is important because it suggests that one-size-fits-all approaches are likely to be lacking, especially in situations where there is greater variation in a trainee's initial competence. Immersive training tools should, therefore, be at least partly responsive to the progress of each trainee and, if possible, adjust their method of instruction accordingly.

Third, the role of ATF, or the perceived alignment between the technology and the activity it supported, was shown to be a significant element of a trainee's experience. While strongly linked to the same topic of facilitating conditions, an inverse correlation was observed. This unexpected finding shows that when there is a strong perceived technology-activity link, the trainee has a corresponding sense that they do not have the tools needed to complete their task. While this may seem counter-intuitive, there are at least three possible explanations that might explain this link.

The first possibility is that, as the trainee approaches the task, which is understood as both complex and unfamiliar, they are cognisant of both their own abilities (or lack thereof) and the needs of the task. When the technology is seen as more aligned with the latter, it may give the impression that it is an external factor, broadening the gap between their current state and that of task completion. The perception of the technology being designed for someone with greater competence than the one they possess may lead to a more acute feeling of lacking the tools required to complete the task, even if those tools were designed to support them. Put another way, when the task is already complex and the technology is felt to add to this complexity, it can result in a stronger feeling of unpreparedness.

The perceived distinction between task and technology leads to another possibility, which is that the trainees exhibit a kind of authority bias, overestimating the accuracy of the authority figure—here, the augmented training procedure—in relation to their own ability. Since the immersive nature of the training is thought to have a more subtle impact on the actions of the trainee, it is conceivable that a greater range of biases play a role in the interaction. The system does, after all, aim to mimic the active support of an expert trainer, perhaps giving rise to additional sources of bias that normally only appear when engaged with face-to-face training.

Other findings related to technology acceptance reinforce the value of the underlying UTAUT model in specific ways. There was a strong correlation between the social influence construct and behavioural intention, suggesting that there is a general perception that using immersive technology offers a competitive advantage or that there is well-founded pressure to adopt this technology within an organisational setting. The notion of habitual use, perhaps related to a common understanding that technology is prone to such effects, was also shown to have predictive power, although to a lesser degree than other influences.

Finally, a significant relationship was found between the behavioural intention to use the technology in the future and the actual usage frequency reported. This provides important validation of a key pathway within the technology acceptance model, but was taken with the caveat that these data expressed significant non-centrality due to the participants being exposed to this type of immersive technology for the first time. It is likely, therefore, that selection bias can play a significant role, as can other sources of error that tend to diminish as people become familiar with new types of digital tools. It is also likely, given that the mode of instruction is so different from paper-based training, that the adoption curve may differ compared to other technologies, such as hand-held devices.

7.5 Notes on an Integrated Framework

While the various strands of work required different modes of study, they worked together to support immersive learning in the workplace. An established science is still forthcoming; many more large-scale studies are needed to investigate the many quirks and vagaries of using immersive technology in practical situations. It has been the task of this thesis to communicate in-depth knowledge of some of the areas in which progress is needed to realise the combinatorial gains present in this augmented space. Before closing the chapter, there is a brief look at how the findings can be brought together under the theoretical frameworks described at the start.

An operative constructivist's perspective on immersive technology

The operative constructivist, following the principles of Luhmann's Social System Theory, considers two things to be extremely important. The first is the system-environment distinction, and by extension, its boundary, but not following the common anthropocentric approach but instead delineating the learning process according to functional difference. Three systems—cognitive, biological, and social—operate as environments for the others, each connected via a set of interdependencies known as a structural coupling.

The immersive training system is introduced into this arena and, because of its form factor and capabilities, is able to modify these couplings across all three domains. It provides various media channels that comprise the interrelations with the social system via units of communication, such as 3D objects or audio instructions. Experientially, it will have an impact on the user, connecting to the cognitive system in a subtle and intricate way, affecting a variety of opinions and attitudes towards its utility, effectiveness, comfort, and so on. Finally, it is also worn on the body and capable of measuring physiological signals as well as giving haptic feedback, providing a link to the biological system.

In an effort to define the structural couplings of the immersive training system, we first need to identify our unit of measurement. For this, we use affordances, which are opportunities for action by the training system in each of the three domains. For each domain, the training system must operate and be developed in different ways, depending on the system it seeks to change.

Affordances for immersive training

Training with immersive technology requires the deliberate use of each type of affordance. In an effort to operationalise this framework, the chart shown in Figure 47 was constructed, which shows each type of affordance being employed at increasing levels of complexity, marked according to the cognitive process that it relates to. The cognitive processes seen in Bloom's Revised Taxonomy [4] are used here, unchanged.

Social systems, made of communication, transmit propositional affordances, which are made of information. This is expertise that can be transmitted or exchanged through written or other media. Biological systems are accessible through embodied affordances, which lean on factors such as mimicry and dexterity to encourage greater practical skill. The cognitive system changes according to experience and can be influenced by affordances affecting the learning process, such as an improved sense of how to motivate understanding or remove bias or a deeper appreciation of one's own learning style. These are referred to as meta-constructive affordances.

Affordances can be thought of as the building blocks of learning in each system. We may possess expertise in any or all of the three fundamental areas; we may have a more detailed and informative mental model; we may be capable of significant practical skill or workmanship; or we may be communicators that describe the other two, as well as the agency of communication itself.

The hierarchy expressed in Figure 47 is neither strict nor linear, so while it is generally beneficial to have information about a task before practising it, in some cases this is not essential. In the same way, long-standing practitioners of a particular skill may be well placed to say how another should learn a task, compared to one who has only a semantic understanding, though this too will bring up counter-examples where this is not the case. With this in mind, we may think of a progression upwards as indicative of greater relevance of the task to the learner, and a progression to the right as pointing to increased competence in that task.

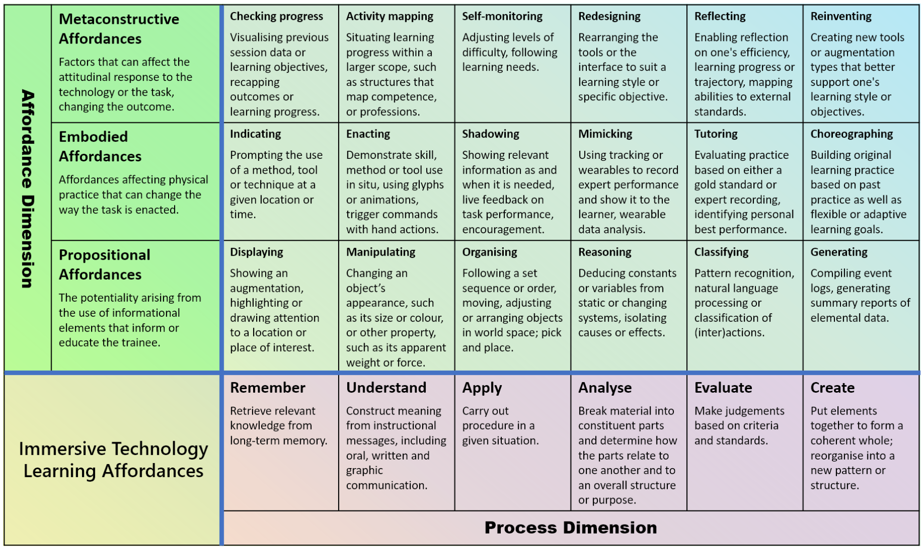


Figure 47: Immersive Technology Learning Affordances.

7.6 Summary

This chapter has brought the practical and statistical conclusions from earlier chapters and consolidated them under a theoretical framework. After attempting a brief summary of Luhmann's Social Systems Theory, it connects it to the practical notion of affordances. In terms of propositional knowledge construction, both the relevance of the augmentation and the competence of the user are seen to contribute to augmented performance. Competence is also seen to arise from embodied knowledge, and through the use of wearable sensors, new opportunities can be made available to the trainee. The grasping of opportunities contained within these augmentations is thought to be mediated by the attitudes that the person holds towards the technology, which was investigated with the use of immersive technology acceptance. Situations in which the technology may be perceived to help or hinder are considered, backed by findings from the structural equation models.

Chapter 8: Conclusions

This final chapter begins by reviewing the initial research objectives and connecting them to the lessons, insights, and discoveries in the earlier chapters. It also mentions some of the limitations that were found, either by necessity or design, in the process of marking out and investigating this tract of academic territory. Additional research that either extends or supports that done here is then shown, and there are a few final words.

The indications and predictions presented in this research have, of course, emerged from particular research and organisational contexts, aligned with the purpose of building the most optimal tools for training. Other interpretations are possible, and constructive challenges to this work are encouraged in the name of building better tools.

8.1 Summary of Contributions

The research objectives laid out in the first chapter pointed to a need for context-agnostic training systems that included tools to support non-propositional learning. As part of the WEKIT project, the ECS of Chapter 4 and body-sensor network of Chapter 5 are a response to this need and provide contributions of both a practical and theoretical nature.

In order to be truly universal, training procedures need to be captured by the experts while they carry out the procedure in question. The ECS described here is an in-situ authoring solution that does just this, producing an ARLEM standard report that describes the augmented workplace and the activity being conducted. This work described the software architecture and information flow that drove this functionality, as well as the visualisation, selection, and capture of the various data streams or augmentations.

This immersive implementation of an authoring system connected a wider variety of types of augmentation than much of the earlier research. From the propositional set of affordances, one stood out as possessing particular educational value: point-of-view video. In this instruction, the visual data it contained was imbued with the position of the wearer, despite this information being implicit in the recording. The ability to easily capture a sequence, event, or action using this method was, however, not always enough to make a good instruction. It was found that, with practice, the experts themselves would improve the quality of the recordings and the quality of the presentations. This highlighted a learning curve present in the use of the hardware and software and speaks to the multi-faceted nature of immersive learning affordances when considering both expert and trainee use.

The second objective—a need for non-propositional capture of training information—was explicitly addressed through the use of motion sensors and EMG-sensitive devices. The system demonstrated the ability to record and play motion capture recordings from both the headset and arm-worn devices. In this sense, the objective was met, although maintaining meaningful, contiguous data streams proved difficult due to sensor drift and bandwidth constraints.

Despite these hurdles, new instructional opportunities were made available with the tools that were built. The 'ghost track' instruction was frequently used by experts and was found, especially when coupled with hand tracking, to be an influential force during training. The sensor fusion algorithm, allowing the hand position to be tracked even when the hands were out of sight, enriches the data stream and opens new avenues for research and development. The ability to serialise action sequences relative to workplace objects, based on the Augmented Reality Learning Experience Model (ARLEM), allows for the creation of contextual data sets, connecting trainee actions to training environments and further extending the potential for technology-enhanced learning.

The third and final research objective motivated the assessment and systematic use of end-user feedback in the development process. To address this, this work presented two iterations of the technology acceptance model, allowing for experimental refinement of the tool as well as the structural equation model. This model produced three important insights:

1) interoperability between devices is essential for immersive technology to be highly effective

2) there should ideally be a close fit between the immersive training and the competence of the trainee

3) for complex tasks, a closer perceived activity-technology fit is not necessarily a good thing.

The final technology acceptance model not only gave broad indications but also specific insights into the use of immersive technology. Individual-technology fit was found to be synonymous with our sense of being equipped for the task, despite these constructs originating from different strands of research. A lack of correlation in the first iteration was explained by the use of activity-technology fit in the second, which was found to be counter-indicative of the same facilitating conditions.

Affordance Theory was found to offer an elegant connection between the surfaces that surround us and both the value and meaning they provide us. The ability of new AR devices to localise themselves and map the geometry of their immediate surroundings provides an avenue of experimentation that directly accesses the theory's core medium. This work goes beyond inferring information about this so-called spatial mapping surface in two ways. We first consider the inclusion of digital visualisation overlaid on this surface, providing new affordances built upon existing ones. Second, we may consider non-visual affordances. These are other environmental factors, such as temperature and humidity, that also furnish the user with other information that directs their actions.

In their design, affordances are seen as having dimensions of relevance that relate to tasks and individuals. Simultaneously, trainees are influenced by their degree of competence in both technology use and the activity. Together, these create the conditions for augmented performance. Competence is accessed through both propositional and embodied means, which are viewed as the structural coupling linking the technological system with the communication and biological systems, respectively, of the trainee. Their cognitive system, in part described by the attitudes and beliefs of the user, can be illuminated with carefully crafted questions. In this case, both correlated and opposing relations were uncovered, indicating that the use and acceptance of immersive technology is a complex and dynamic affair.

8.2 Limitations

Much more research awaits the developers and academics working in the field of immersive training. While this thesis has highlighted several productive avenues, it has also met with challenges and constraints. The overarching strategy was to take a didactic approach to system design and a pedagogical approach to system validation, one using instructional design principles and the other notions of affordances and structured feedback. Although generally suitable, this research design necessarily leads to blind spots, where learner participation could inform the design process and expert input could be allowed to shape the acceptance profile of the technology.

The inclusion of learners in the system design process is made easier by the growing prevalence of immersive technology. This needs to be built into the project from the outset and is a good candidate for further work. The latter category is more easily tackled, as it can be done after the fact. To address this limitation, trainers in their respective workplaces were encouraged to participate in the technology acceptance study. Appendix C is the beginning of a study to meaningfully address the lack of expert feedback, using the final model from Chapter 6 and 40 responses from trainers, but it is incomplete. The dataset is reduced due to some questions lacking relevance, and considerably more analysis can be performed to uncover the trends and features there.

The experience capture system, although capable of a wide range of tasks, represents a tiny part of the developmental possibility space made available with AR and wearable technology. Despite this, the use of ARLEM as a standardised language for describing activity and workplace has some significance beyond the specific implementation. In this regard, there are, however, at least two important limitations. One is the complexity of the activities used to test the system. These were non-branching sets of task steps that were relatively straightforward when determining their completion. In more realistic scenarios, it is highly likely that an activity would express some non-linear or conditional flow and that tentative or partial completion of a step is needed. Another notable limitation comes from the interpretation of the ARLEM standard itself, which is open-ended with regards to the kind of data it can contain. Along with more intricate and branching procedures, a more nuanced form of ARLEM expression would be required than that used here.

In the case of the wearable sensor framework, the initial intention was to provide a continuous data feed, containing both movement and muscle data, to the headset, where augmented appendages would provide a new layer of immersive interactivity. This, however, was not fully achievable. A lack of stability, combined with narrow communication lanes, meant that a dataset of actions and operations could not be collected. Since that work was done, new versions of both components and head-mounted displays have become available, potentially making this realisable.

The design of the technology acceptance study, in its use of convenience sampling for this dataset, ignores trends present due to mediating factors such as age, gender, and other demographic metrics. Additionally, while it was acknowledged that all of the trial participants were using the technology for the first time, steps were not taken to remove the bias that is expected to be associated with such a scenario. Indeed, the enthusiasm of these trainees was often palpable, especially in the case of the Mars rover simulation, which was an example of rare access to an impressive facility. Such effects will surely influence the sentiment of the trainee, and, as a result, a greater emphasis was placed on understanding correlated and relative effects as opposed to normative assessments of the technology. A longitudinal study could escape and/or measure this bias and potentially offer stronger conclusions about factors affecting immersive technology adoption.

8.3 Future Research

This work has also shown that the presence of immersive technology is, by itself, no guarantee of improved tuition or performance. The complex relationship between learner and learning environment can be influenced in a variety of ways, and each will have an impact on learning and task performance. While this research has made progress in investigating and understanding the role of immersive technology, there is much that remains to be done.

A clear extension to this research is a more in-depth study of experts' perspectives on immersive training systems. As mentioned above, Appendix C is the start of such a study, which could compare the style of authoring according to industry or activity. It may be possible to create general guidelines or templates that suit specific needs, or a more intelligent system may be able to interpret an expert's instruction in more natural ways, using a vocabulary tied to natural language or gesture-driven control.

The methodology of model optimisation and refinement used here is a suggestion for a standardised approach. If such a standardisation were achieved, even partly, structural equation models could begin to function as inference engines, powered by behavioural data about each iteration of development. Inferences can be tested, first through the manipulation of model parameters and the introduction of new constructs, and again experimentally, using precise experiential phenomena and semantic structure in the form of attitudinal constructs.

Ideally, such work would be conducted in the open, using published behaviour models and transparent data collection methods. A parallel effort toward the development of good ethical practice is, for many reasons, essential. The potential for data-based exploitation abounds, and more efforts are needed to prevent data misuse. Certainly, the aggregation of large datasets around movement or other physiological data has important security and privacy implications that cross disciplines, such as metrics related to health or attention.

Despite some tentative formulation and classification, the connection between affordances and the specific visualisation of immersive instruction remains unclear. More design, development, and testing are needed to understand what form and grouping of form many of the instructions will take. With higher-complexity processes, this process is more difficult, as the affordances become nested and blend with other elements of expert practice and interdisciplinary knowledge. The construction of a complete set of immersive instructions, together with modulating factors that allow for degrees of perceptibility, is a natural extension of this work.

The work on wearable motion capture presented here is a prelude to a much more in-depth investigation. In such a study, features based on physical motion as well as motion coupled to activity-specific triggers could provide rich insights into how we adapt our movement depending on our attitudinal and physiological state, such as confidence, cognitive load, or the sense of feeling supported by the technology. Connecting the realms of biological, cognitive, and communicative competence could transform not only a trainee's progress towards specific learning objectives but also their embodied competence and ability to articulate and share their knowledge with others.

As discussed, the 'activity-technology fit' construct is in need of further investigation. Understanding how a trainee will develop a sense of collaboration with the technology remains unknown and could interact with a variety of personal and organisational factors. Additional studies in new industries, while testing the hypotheses given here, could also benefit from the core, validated model as well as the optimisation and analysis strategy used here.

In the areas of AR system design, wearable integration, and technology acceptance, this research provides a basis for future work and, in the process of answering its central research question, has uncovered several others. Hopefully, the practical as well as theoretical insights from this thesis will motivate, expedite, and support other work aimed at improving workplace training with immersive technology.

8.4 Closing Remarks

Performance augmentation, the concept that entitles and underscores this work, has been found to be something abstract as well as tangible. Cognitively, the tendency is towards abstraction; as our knowledge grows, we connect, rearrange, and generate ideas that help us respond to situations and anticipate solutions. Kinaesthetically, we move more deftly and adapt our pose to both our current action and that which is anticipated. When it comes to communication, our efforts as experts tend towards a more sophisticated understanding of how to author learning experiences, using both symbols and actions, and, as trainees, how to interpret the instructions coming from the technology framework into knowledge of the task, both in principle and in practice.

Appendix A

ARLEM: Augmented Reality Learning Experience Model

The AR Learning Experience Model (ARLEM) [134] is a standardised format for exchanging information about activities and workplaces. The activity model is a description of the structure and content of the procedure. It specifies how the action steps connect to one another and which information each instruction contains. The workplace model is a description of the (relatively) slow-changing environment, identifying key objects or places with which the system interacts, as well as other devices or sensors that are in communication with the system.

For example, aircraft maintenance procedures can be taken from an on-demand library of previously recorded procedures. By specifying the workplace model around the aircraft rather than the hangar, we can also lock the reference frame, making the maintenance procedures usable from anywhere.

The ARLEM standard also provides an essential piece of the technical puzzle as we move from functionally-driven events in the software's code to a set of outputs that describe the trainer's decisions when using the ECS. As a conceptual model, it shapes the way the data is collected, and as a data model, it connects to the established experience API (xAPI) specification, which functions as an endpoint for experience logs.

The standard is written in extensible markup language (XML), snippets of which will accompany the following description. For a complete description, the reader should look to the standard itself [134].

The Activity Model

The whole procedure sits within a single top-level 'activity' element, which has both machine- and human-readable names and attributes of description and language. At this level, it is connected to the workplace model via a uniform resource locator (URL), and there is a pointer to the first action step.

```
<activity
    id="A unique identifier."
    name="A human-readable name of the activity."
    description="A human-readable description of the activity."
    language="The language of the activity's human-readable prompts."
    workplace="URL of the workplace model."
    start="The <id> of the first action."
>
...
</activity>
```

Listing 3: ARLEM activity element

An activity is broken into action nodes, each of which is a place-holder for one or more instructions. Each indexable action node has attributes that categorise the action in terms of visual location and appearance. Other optional attributes may be read from the workplace model, including the device used, the location and an xAPI predicate, which can connect to one of the domain-specific verbs defined there [56].

```
<action
    id="A unique identifier."
    viewport="The display area where the content is visualised."
    type="The category of the action, used for visual styling."
    [device]="The device on which the actions are executed."
    [location]="The location of the activity."
    [predicate]="An xAPI ‘verb’ that describes the interaction."
>
...
</action>
```

Listing 4: Action element attributes.

An action has four configurable parts: an 'instruction' that provides human-readable information; ‘enter’ and ‘exit’ elements that contain machine-readable instructions on which augmentations to control; and, finally, a ‘triggers’ section that describes how a transition from ‘enter’ to ‘exit’ occurs. Listing 5 briefly describes the attributes and illustrates the sub-structure of each of these parts.

When an action step is first initialised, it performs any functions contained in the entry phase, after which the trigger(s) specified become active. If the trigger conditions are met, or if the ‘removeSelf’ attribute of the entry phase is ‘true’, the action will proceed to the exit phase. If the ‘removeSelf’ attribute of the exit phase is 'true', the action will be removed *before* the functions of that phase are executed, removing the instructions and triggers that are associated with it. Otherwise, the functions will be executed prior to the action's removal.

```
<action
...
        <instruction
                title="The instruction’s headline"
                description="More information about the action"
        >
        <\instruction>
        <enter
                [removeSelf]="A boolean denoting whether to automatically
                     proceed
                                    to the exit phase,once the enter phase
                                         have completed."
                <activate> ... <\activate>
                <deactivate> ... <\deactivate>
                <message...>
                <if...>
        >
        <\enter>
        <exit
                [removeSelf]="A boolean denoting whether to remove the
                                    action step prior to the exit phase,
```

```
                                         removing the instruction and any
                                              triggers."
                <activate> ... <\activate>
                <deactivate> ... <\deactivate>
                <message...>
                <if...>
        >
        <\exit>
        <triggers
                <trigger> ... <\trigger>
        >
        <\triggers>
</action>
```

Listing 5: Action element structure.

Triggers allow an action to progress from enter to exit phase and are identified in reference to an action, tangible (in the workplace model), or sensor that the trigger is observing. They identify when actions have taken place and can signify direct intervention by the operator, the detection of a visual marker, or the acquisition of a sensor value or pattern, represented as the 'mode' attribute.

If detecting a physical object, some form of delay is required to indicate selection, the duration of which is an optional attribute, along with the observation type, display area, and sensor threshold values.

```
<trigger
        id="A reference to an action, tangible or sensor."
        mode="One of three possible modes: UI, detect and sensor."
        [duration]="Duration of interaction or detection, as an integer."
        [type]="The type of entity that this trigger observes."
        [viewport]="The display area where the content is visualised."
        [key,value,operator]="'Sensor' mode only. Specify a variable key,
                                                  set a threshold at a certain
                                                        value and
                                                  optionally compare it with
                                                       another operator."
<\trigger>
```

Listing 6: Trigger element

Both enter and exit phases may be used to activate and deactivate content. This means that the appearance and flow of content need not be tied to the sequence of action steps, but rather to the activity that they represent. This decoupling also eases development requirements, allowing for a separate treatment of action step sequencing and the (de)materialisation of procedural content, as the 'activate' and 'deactivate' sections can be used in either phase.

It is in using the (de)activation elements that the rubber meets the road, so to speak, for it is here that the objects represented in the workplace model are referenced, first through the 'target' property, which references a physical object or location, and second through the 'augmentation' property, which points to the media content used in this instruction. In addition, their type, point-of-interest (poi), and media-specific properties are configurable here. A full description can be found in the IEEE standard [134].

```
<activate
       target="A reference to a workplace tangible."
       type="Type of augmentation: primitive, predicate, warning or action."
       augmentation="The id of the augmentation in the workplace model."
       [poi]="The location where the augmentation will be displayed."
       [option]="A configuration option for the augmentation."
       [viewport]="The display area of the augmentation."
       [url, state, text,
       sensor, key, option]="Media-specific properties; reference a URL or
                                  sensor id, or configure streaming data
                                      flow."
<\activate>
-----------------------------------------------------------------------
<deactivate
       target="The id of workplace tangible."
       type="The type of workplace element (augmentation)."
       augmentation="The id of the element to be removed."
       [poi]="The location from which to remove the augmentation."
       [viewport]="The display area from which to remove augmentations."
<\deactivate>
```

Listing 7: Activate and deactivate elements

When deactivating elements, a wildcard ('*') can also be used to remove multiple augmentations by action, type of augmentation, location, or display area. This is particularly useful when implementing a branched activity structure where the user has a number of choices, each of which can lead to different outcomes. Any number of activate or deactivate statements can be made as part of an action step, and each can be linked to another action ID, providing the functionality to generate any form of action network, though further definition is required to allow for conditionality. The 'if' element does this and uses a typical if-then-else structure, also incorporating a query URL and minimum and maximum bounds for received data.

```
<if
       url="A query URL, with parameters."
       then="The action to trigger if the condition is met."
       else="The action to trigger if the condition is not met."
       [min]="Check the number of results is more than this integer."
       [max]="Check the number of results is less than this integer."
<\if>
```

Listing 8: If element

The final part of the activity model is the 'message' element, which allows communication between people, devices, and sensors ('type'). The message is compatible with the Message Queuing Telemetry Transport (MQTT) protocol and uses a 'key' variable to denote the MQTT topic, which is handled by a message broker. In MQTT, a wildcard ('+') may be used to refer to all topics of the same type. The message can also reference an action that can be launched on the target device. A text option gives human-readable messages.

```
<message
       target="A link a person.id, device.id or sensor.id"
       type="This can be either to a person, device or sensor."
       [viewport]="The display area where the message is shown."
       [key]="The name of the MQTT topic or similar descriptor."
       [launch]="Which action should be launched on the device?"
       [text]="A message to display to the user."
<\message>
```

Listing 9: Message element

The Workplace Model

The workplace model is a document describing useful or interesting parts of the work environment. It contains four types of object: tangibles, configurables, detectables, and augmentations. Table 29 shows the substructure of these resources and a short description of each.

Table 29: Workplace model structure.

Resource	Sub-category	Description
Tangibles	Persons	Particular individuals.
	Places	Locations in the workplace.
	Things	Physical workplace objects.
Configurables	Sensors	Devices that measure physical properties in or of the workplace.
	Devices	Any device connected to the training system.
	Apps	Details of any local or cloud applications that have an interface with the training system.
Detectables	Markers	An identifiable feature in the environment. Markers, although typically flat images, can also be 3D or be features of the surrounding spatial geometry.
	Anchors	These are reference locations, set by the software, relative to the origin of the workplace coordinate system.

Augmentations	Primitives	A set of standard visualisations of common media types, including 3D models, animations, image, video, audio and text.
	Predicates	Visualisations linked to a basic vocabulary of motion and handling verbs linked to manufacturing and maintenance.
	Warnings	Graphical overlays representing the ISO 7010 warning signs for workplaces.

Tangibles are physical entities that interact with the training system. Every tangible has an indexable identifier, a description, a link to a detectable, to make it discoverable, and a point-of-interest (poi). One type, the 'person' element, has additional attributes that facilitate communication with them.

```
<thing
       id="A unique identifier."
       name="A human-readable short description of the thing."
       detectable="The id of a reference with real-world coordinates."
       <poi> ... <\poi>
<\thing>
--------------------------------------------------------------------
<place
       id="A unique identifier."
       name="A human-readable short description of the place."
       detectable="The id of a reference with real-world coordinates."
       <poi> ... <\poi>
<\place>
--------------------------------------------------------------------
<person
       id="A unique identifier."
       name="A person's name."
       detectable="The id of a reference with real-world coordinates."
       [twitter]="The person's twitter handle."
       [mbox]="The person's email address."
       [persona]="Their role within an organisation or group."
       <poi> ... <\poi>
<\person>
```

Listing 10: Workplace tangibles

The poi is assumed to be at the geometric centre of the tangible, unless specified otherwise. It is referenced by the activity model's 'activate' and/or 'deactivate' elements. The numerical entries (i.e. all but 'id') are floating-point values.

```
<poi
    id="A unique identifier."
    z-offset="Offset along x (cm)."
    y-offset="Offset along y (cm)."
    z-offset="Offset along z (cm)."
    z-rotation="Pitch (euler angles)."
    y-rotation="Yaw (euler angles)."
    z-rotation="Roll (euler angles)."
<\poi>
```

Listing 11: A tangible's point-of-interest.

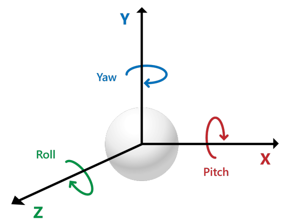


Figure 48: poi pose

Configurables are those entities that connect to the training system, allowing it to reach beyond its own wiring. Low-level sensors as well as equivalent devices can be referenced here, and external applications can be run with the 'app/url' property. Using 'application' often forms a part of more technical procedures where application identifiers can be used in place of whole instruction elements in the activity model.

```
<sensor
       id="A unique identifier."
       url="A human-readable short description of the thing."
       username, password ="Authentication information."
       <data> ... <\data>
<\sensor>
--------------------------------------------------------------------
<device
       id="A unique identifier."
       type="The type of the device being used."
       name="A human-readable name for the device."
       [owner]="A person.id of the owner."
       [url]="The device's message protocol connection string."
       [topic]="The listening channel on the messaging protocol."
       [username], [password]="API authentication information."
<\device>
--------------------------------------------------------------------
<app
       id="A unique identifier."
       type="Either an HTML widget, prototype element or launch command."
       name="A human-readable description of the target application."
       url="A method to invoke functions in other applications."
<\detectable>
```

Listing 12: Workplace configurables.

Detectables are description of the visual elements that have meaning to the visual (and depth) camera system. Markers may be 2D image targets or more complex 3D models or features derived from the spatial mapping system. Point-clouds models may be referenced by the 'url' property.

```
<detectable
        id="A unique identifier."
        type="May be a ‘marker’ or an ‘action’."
        [sensor]="‘Tracking’ objects or ‘mapping’ the environment."
        [url]="The addressable location of trackable content."
<\detectable>
```

Listing 13: Workplace detectables.

Augmentations, in the language of the workplace model, refer to reusable instructional materials. The details in Listing 8.4 centre on the predicate, which is a typical verb (in an xAPI understanding) involving movement or handling. In this subject-verb-object syntax, the ‘url’ property provides the object (or target) of the interaction. Warnings have a dedicated classification.

```
<primitive
        id="A unique identifier."
        [x-size], [y-size], [z-size]="Scale factor for images and objects."
        [volume]="An option to set audio volume."
<\primitive>
-------------------------------------------------------------------
<predicate
        id="A unique identifier."
        type="An identifier for a primitive."
        scale="The scale factor applied to the primitive."
        url="A link to the file containing the relevant augmentation."
<\predicate>
-------------------------------------------------------------------
<warning
        id="A unique identifier."
        type="The primitive id."
        scale="Primitive scale factor."
        symbol="A game engine reference to the prototypical visual element."
<\warning>
```

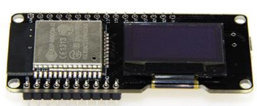
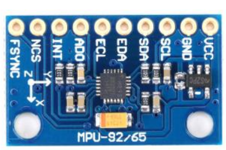
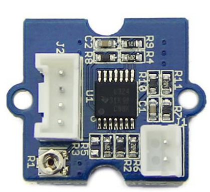
Listing 14: Workplace augmentations.

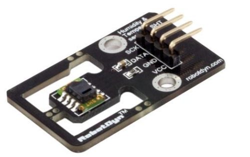
Appendix B

A Body Sensor Network

These biosignals, including heart rate variability (HRV) and galvanic skin response (GSR), are thought to support an understanding of the level of exertion and attention exhibited by the wearer, and, while their use is not the focus of this work, they are included in this description of the system for two reasons. First, the significant efforts of my WEKIT collaborators in developing these tools should not be overlooked, and second, it is anticipated (though yet to be proven) that such signals can enhance information about body movement. Work by Sharma et al. [109] and Ransley et al. [90] on the selection of the hardware components led to a final set of hardware, shown in Table 30, as well as prototypes.

Table 30: Body sensor network hardware selection.

Component	**Sensor Description**
ESP32 Microcontroller 	32-bit LX6 microprocessor. Bluetooth, BLE and WiFi integration. Low operating voltage (3.3V), current (80 mA). Small form factor On board screen display.
MPU-9265 IMU 	Inertial Measurement Unit (2) Gyroscope, accelerometer and magnetometer. Dedicated I^2C sensor bus. High data acquisition rate (up to 200 Hz).
Grove heart rate monitor 	Easy setup and use. 3-5V operating voltage. Low cost Low power consumption (6.5mA) Sampling rate of around 0.5 Hz
Grove galvanic skin response sensor 	Uses ear clip to monitor heart rate. 3.3 or 5V operating voltage. Adjustable sensitivity (with potentiometer).

Temperature and humidity sensor 	Measure environmental conditions. Supply Voltage: 3.5 - 5V Low power consumption: 0.55 mA Absolute humidity accuracy: ± 2% Determined by temperature: -40+128ºS ±0.3% Temperature accuracy: ± 0.3°C at 25°C Sampling rate: 1 Hz
Vibration motor 	Voltage range : 2.5 V - 4.0 V Starting voltage: 2.3 V DC Rated speed : 2500 - 12000 RPM Rated current : 70 mA Starting current : 90 mA

Some details of this framework have appeared in earlier work [43], demonstrating the components as part of a wider architecture. While it briefly described the choice of the MQTT protocol and the communication flows, this section expands on this, showing the wiring and software patterns that were used to handle the data. A redrawn, simplified chart showing the full pipeline, from sensor to ECS, is shown in Figure 49.

Two components handle data flowing from the e-textile. The first, containing the input/output (IO) modules, runs once during system initialisation, using both direct wiring to components and those running through an inter-integrated circuit (I^2C) serial communication bus. The data handler, running in a loop, reads from the various channels, aggregating the data into a single frame, which it passes to the MQTT broker for transmission within a pre-assigned topic.

After transmission, the data must be decoded, integrated with the HMD's localisation system, and used to drive the visualisation of the wearer's movement. Figure 49 shows the sequential set of components that perform these tasks within the ECS. The 'sensorBase' class contains data structures that anticipate the structure of the incoming data, not only from the e-textile's peripherals but also any sensors on board the HMD. The Ghost Track Instruction is the central ECS component that connects the data streams to user interface commands, allowing the wearer to record and visualise the information.

Component Input/Output Modules

The IMUs were connected via the I^2C bus, a 2-wire interface on the ESP32 microcontroller board. The GSR and heart rate sensors required a single pin, and the temperature/humidity sensor and vibration motors each used two pins. The final pin usage is shown in Figure 50.

The OLED display showed the name and version of the software and was used to track error messages and connection events, as well as the packet length and uptime. Understanding the update frequencies of each sensor was critical in

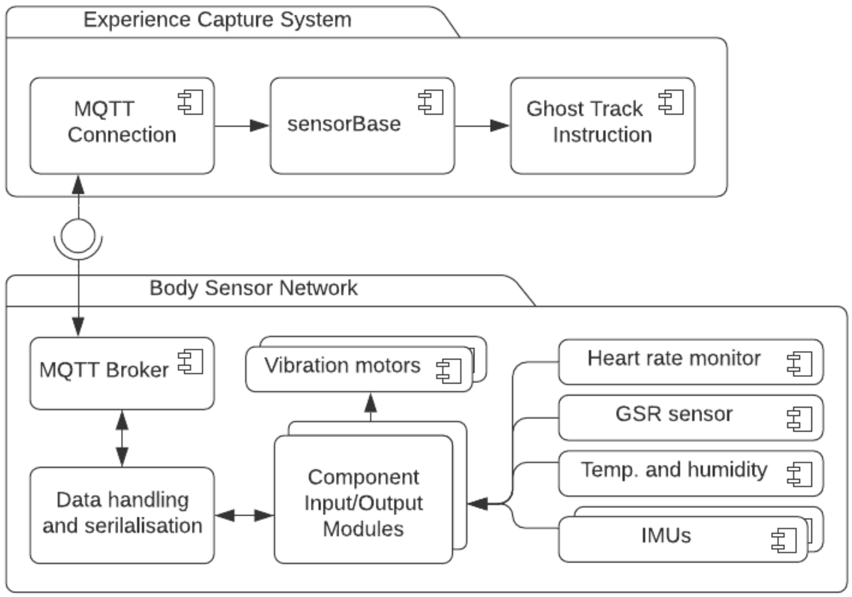


Figure 49: Body sensor network component diagram

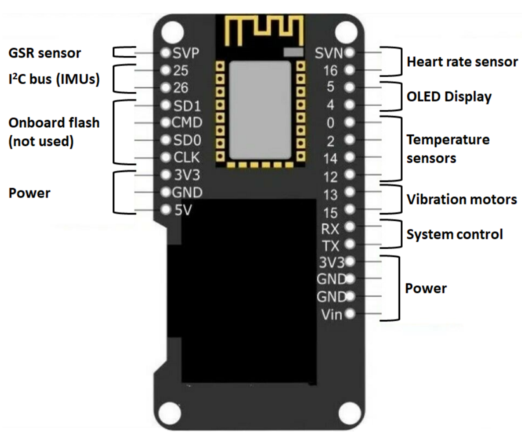


Figure 50: ESP32 pin usage.

determining the method used to collect data. The IMUs and temperature/humidity sensors each provided new data at a rate equal to or higher than the speed at which the processor would loop through the code, so acquisition was possible on each pass. The GSR sensor provides continuous data, so it could be read at any time, though the data was averaged to remove unwanted noise. The estimation of heart rate was performed using a sliding window of two seconds. During this window, each heartbeat was signified by an interrupt. The moving average of these pulses was then reported. The update frequencies, along with the data storage requirements, are shown in Table 31.

Sensor	Signal	Range	Data frame	Frequency
GSR	Skin resistance	0-60 kOhm	integer (1)	10 Hz
HRM	Pulses/sec	$>$ 30 bpm	float (1)	0.5 Hz
SHT11	Temperature	0-70 Celsius	float (1)	100 Hz
	Humidity	0-2000 bar	float (1)	100 Hz
MPU 92/65	Linear acceleration	3-axis (0-360)	float (3)	200 Hz
	Angular acceleration	3-axis (0-360)	float (3)	200 Hz
	Magnetic orientation	3-axis (0-360)	float (3)	200 Hz
	Attitude and heading	4-component	float (4)	200 Hz

Table 31: Body sensor data acquisition.

The vibration motors were switched on and off based on commands received over the MQTT messaging service. The component IO module was triggered by these commands at the start of the processing loop.

Data Handling and Serialisation

Leveraging each of the connection patterns used in the IO component, the data handling component builds, sequentially, a data string in a JavaScript Object Notation (JSON) format. The attribute-value pairs, delivered as a single statement, were constructed with code (in the Objective-C language) shown in Listing 15.

```
String to_json(String attr, String value, bool add_comma=true)
{
  String tstr = "\"" + attr + "\":" + value;
  if (add_comma) { tstr += ","; }
  return tstr;
}
```

Listing 15: JSON string builder

A device identifier and time stamp for each measurement were collected, after which the heart rate, GSR, temperature, and humidity variables were read and converted into message-friendly characters. The data coming from the IMUs—a total of nine values covering linear and gyroscopic acceleration and magnetic orientation—was put through a variation of the Madgwick filter [74], using input from all three sensors to produce a smoothed, normalised attitude. This is essential, as it removes much of the drift error inherent to the gyroscope by maintaining a frame of reference along the magnetic field. The quaternion produced in this process is a complete description of the orientation state (i.e. free of gimbal lock) due to the double-cover property of this type of representation. The quaternion has an "x,y,z,w" structure, where the final part is a scalar, imaginary value.

Putting these values together and including the axis values for all parts of the IMU data produces the final data string, shown in Listing 16. Overall, this message uses 289 bytes, or, including the header used by MQTT, 303 bytes.


```
{
  "client":"WEKIT-VEST-0000F4A2",
  "time":68096,
  "heartrate":65,
  "imus":[{
      "ax":-0.04, "ay":-1.75, "az":-0.06,
      "gx":-7.33, "gy":-46.49, "gz":-7.33,
      "mx":-356.02, "my":28.77, "mz":380.96,
      "q0":-0.58, "q1":0.56, "q2":-0.45, "q3":0.39
    }],
  "gsr":0,
  "temp":20.3,
  "hum":952
}
```


Listing 16: Final MQTT message example

Communicating via MQTT

Connections using this protocol are made after authenticating each device on a local WiFi network. Once this is done, the e-textile software creates an MQTT server and sets up messaging channels for protocol configuration (the control channel) and data transfer (the data channel). Clients on the same WiFi network can subscribe to either channel and receive notifications when data is published to it. The handling of incoming messages is necessarily asynchronous and fulfilled by the event delegation pattern. The ECS, or, more specifically, its connection manager component, operates as the client and uses a C# class containing events that handle connection, the publication of and subscription to topics, and the receipt of data.

An overview of the functions that achieve this is shown in Figure 51, which also illustrates the discrete event-driven method for providing haptic feedback and the looping (though still event-driven) method of sensor data acquisition.

Onboard the wearable component, the MQTT manager class directly handled events related to (dis)connection and the receiving of messages. The structure of these events in this implementation is shown by the call graph in Figure 52.

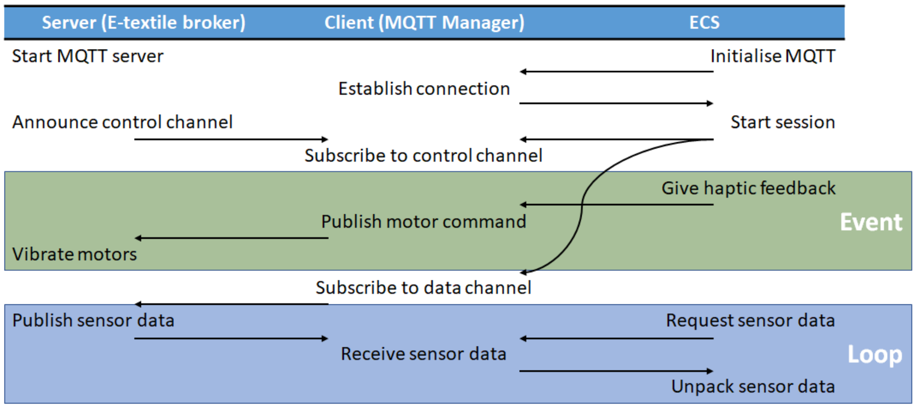


Figure 51: Data flow diagram for MQTT connection

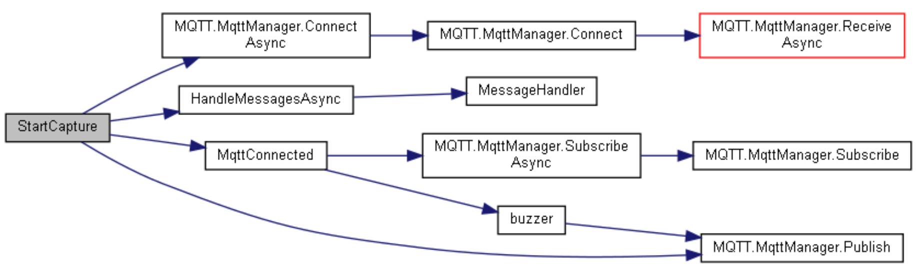


Figure 52: Vest Sensor call graph.

Appendix C

Use Case: Trainer Technology Acceptance

The success of the training system was ultimately determined through end-user (trainee) evaluation. However, the trainers' acceptance of the ECS was also considered invaluable in understanding how the authoring system could be improved to better suit the needs of the trainers when authoring procedures.

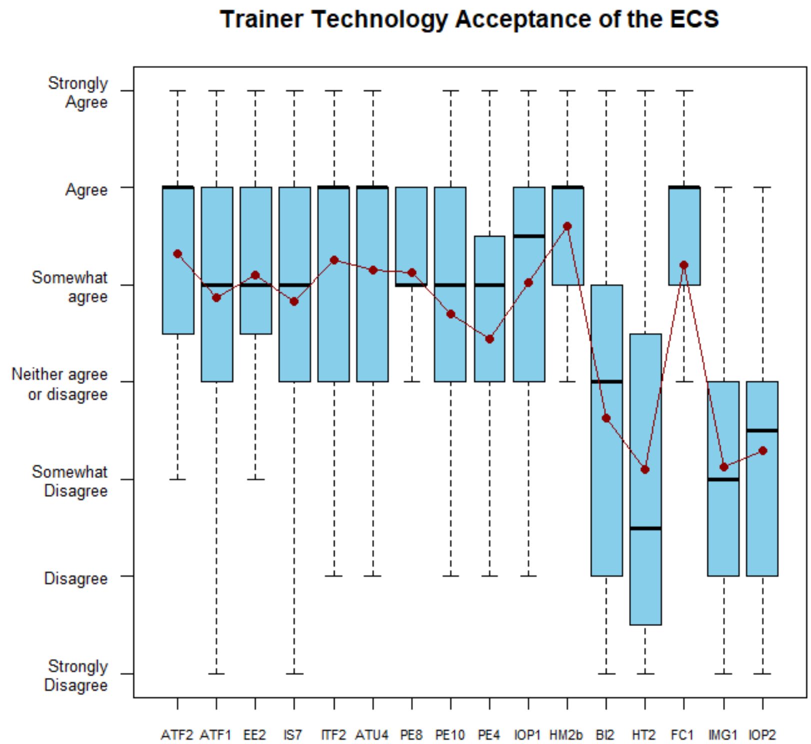
Following the formulation of the technology acceptance model above, it was immediately put to use to investigate trainers' impressions of the AR training tool. We were motivated to investigate their perspective for several reasons. First, workplace trainers are far less numerous than trainees, making it difficult to gather sufficient sample sizes for model validation. Second, they are a valuable trove of practical know-how and instruction, two elements the system aims to specifically emphasise. Third, they are often more central (than trainees) within organisational decision-making structures, meaning that their attitudes are more likely to be translated into organisational adoption or change.

In the medical context, eight interviews were conducted with radiologists (2), students (2), and medical device experts (4). All of them reported enjoying the experience; half described the experience as "very nice", and three said that the experience was "interesting". Comments were made about the effort required to use the tool, saying it was "easy to use" and "intuitive", as well as the performance of the system, reported as "smart and with high potential in training" and having "good overlap of the physical world and AR". The ECS was mentioned as being helpful and a suitable technology: "AR is perfect for this kind of task where reference images are clearly available".

Where the use of the system drew criticism, it was generally related to the human-computer interface—"difficulty with voice interaction and somewhat spatial interaction"—and the stability of the holograms—the visualisations were "sometimes distorted" and one person was "disappointed because the spatial reference was lost". With regards to the device itself (the HoloLens, first edition), three of the interviewees described it as "too heavy" and two mentioned a lack of comfort when using it with eyeglasses.

Other observations of the use of the system highlighted the difficulty with gestural interaction. Often this was due to a larger distance to the holograms, making them appear smaller and more difficult to focus on using the gaze cursor, or their occlusion by other visual elements. Despite these difficulties, it was noticed that the users of the technology would remain still, rather than approach the more distant holograms or step to the side to gain access to occluded parts.

Besides the more informal observations and inquiries about the trainers' experience, a technology acceptance questionnaire was also completed by each person using the ECS. Based on previous work, the questions extended the UTAUT2 technology acceptance model and asked participants to score a set of questions on a 7-point Likert scale, ranging from "strongly disagree" to "strongly agree". The statements are listed in Table 16. Across all workplaces, 40 responses from trainers were collected. The results are shown in Figure 53.



	ATF2	ATF1	EE2	IS7	ITF2	ATU4	PE8	PE10	PE4	IOP1	HM2b	BI2	HT2	FC1	IMG1	IOP2
Mean	5.32	4.88	5.10	4.82	5.25	5.15	5.12	4.70	4.45	5.03	5.60	3.62	3.10	5.20	3.12	3.30
Std.Dev	1.10	1.38	1.26	1.39	1.08	1.21	1.02	1.38	1.50	1.25	1.08	1.61	1.92	1.20	1.44	1.56
Median	6.00	5.00	5.00	5.00	6.00	6.00	5.00	5.00	5.00	5.50	6.00	4.00	2.50	6.00	3.00	3.50
Skewness	-0.65	-0.98	-0.79	-0.75	-0.97	-0.79	-1.38	-0.44	-0.55	-0.81	-0.61	-0.04	0.56	-1.07	0.24	0.38
Kurtosis	-0.62	0.17	0.00	0.62	0.22	-0.34	1.91	-0.73	-0.34	-0.35	-0.69	-0.92	-1.03	0.54	-0.89	-1.05

Figure 53: Responses from trainers across three workplaces. (n = 40)

Those who used the system were positive about its enjoyability and accessibility. Questions investigating hedonic motivation (HM2b: "I like working with AR&WT") and facilitating conditions (FC1: "I have the resources necessary to use AR&WT") were among the highest scored, and there was universal agreement that these were not negative factors.

Other responses that showed positive attitudes and beliefs towards the technology were in the area of technology fit. The fit between the activity and technology (ATF2: "The right amount of information was presented in each step") was scored, on average, at 5.32 (SD = 1.10), suggesting a largely affirmative attitude, although a negative skew (-0.65) and kurtosis (-0.62) indicate that more data or insight is needed. The individual-technology fit, investigating whether the "information provided by the system was appropriate for my level of expertise" (ITF2), also scored well with a mean of 5.25, and, in this case, the data has a spread narrower than that of a normal distribution, with an excess kurtosis of 0.22.

Statements about the two most widely studied predictors of technology acceptance, performance expectancy and effort expectancy, also elicited mean response scores above 5. Of particular note is item PE8, which measured whether "AR&WT increases the precision of tasks". This metric (mean = 5.12) showed the lowest standard deviation (1.02) and a large positive excess kurtosis (1.92), indicating not only positive agreement but consensus amongst the respondents. A histogram of the responses to this metric can be seen in Figure 54, where 80% of the respondents answered with either "somewhat agree" or "agree".

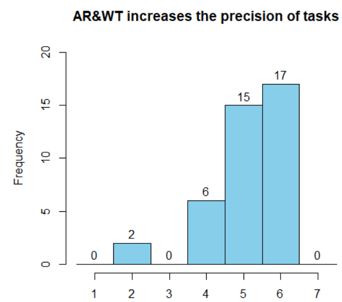


Figure 54: Responses to metric PE8.

Four of the results indicate more negative positions. In two cases, habit (HT2: "I am addicted to using AR&WT") and vendor-lock (IOP2: "I worry that I could become too dependent on a single AR & WT supplier"), a negative response may in fact indicate positive beliefs towards the technology, as either habitual or vendor-locked use could be considered unwanted outcomes. The remaining two, relating to whether the technology may confer prestige within an organisation (IMG1) or whether the experience encourages people to use AR&WT in their daily work practice (BI2), may indicate that a state of technology readiness has not yet been attained in this case. It is worth noting, however, that all four of these results express highly platykurtic distributions (all excess kurtosis values are close to -1), suggesting that several meaningful outliers are present and that there may be additional factors or sources of randomness driving these results.

To understand how the various metrics are connected, their statistical correlations are also reviewed. Figure 55 displays these, highlighting significant relations with various cutoff values.

Groups of items may show where trainers' attitudes overlap. Reading from top-left to bottom-right, the shaded boxes indicate areas of strongest similarity. The metrics are sorted by the magnitude of their correlation eigenvectors, which has the effect of clustering related items. Those items already grouped within constructs that investigated performance expectancy (PE) and activity-task fit (ATF) were, unsurprisingly, found to strongly correlate. However, no such pattern appears for items within the interoperability (IOP) construct.

Behavioural intention (BI) and habit (HT) showed a significant correlation. This may be due to semantic similarity in the framing of the questions, but non-

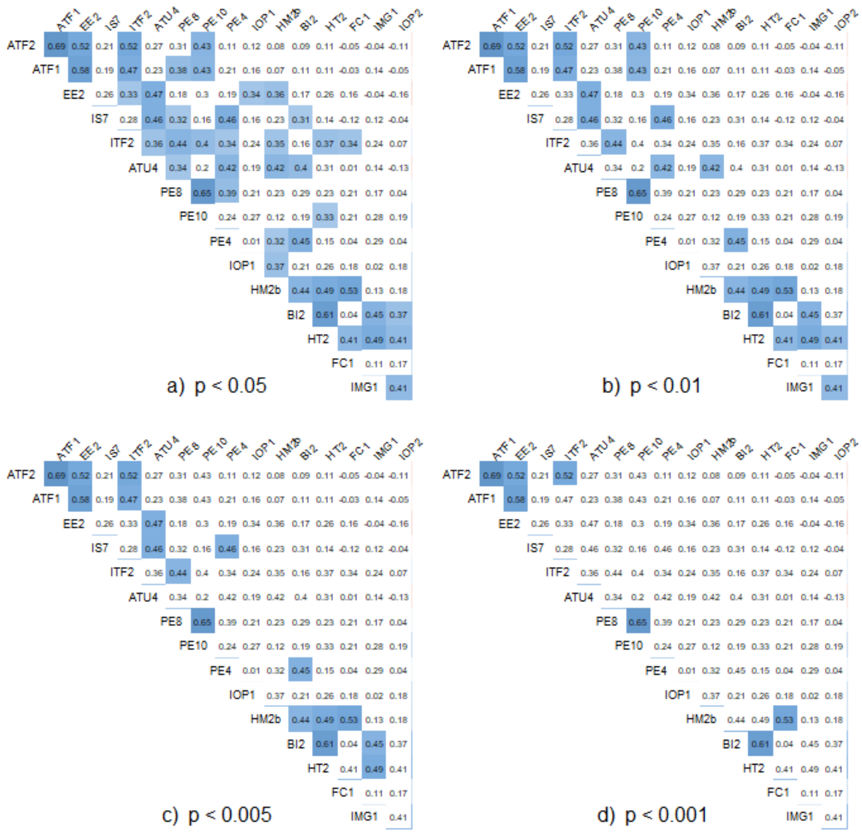


Figure 55: Correlation plots

etheless, it is in agreement with the UTAUT2 model of technology acceptance, in which attitudes towards habit are thought to be a direct predictor of those guiding an intention to use the technology.

The responses to questions about effort expectancy (EE) and individual technology fit (ITF) were also shown to co-vary with each other and those related to ATF. EE is understood as either the perceived or actual ease of use of the technology, also drawing on the notion of complexity from the diffusion of innovation theory [96], which considers the trainers' perception of the technology's difficulty to understand or use.

All of the metrics investigating the EE-ITF-ATF group scored well, with only ATF1 ("This particular procedure was well supported by the system") falling below an average of 5. Thus, the view that the technology is easy to use is suggested to be supported by the view that the technology allows both the procedure and their own expertise to be well expressed.

Overall, trainers reported that the technology was enjoyable to use, that they already have the resources they need to begin using the technology, and that they consider the technology to increase the precision with which tasks can be accomplished. The trainers were not, however, intending to use it in their everyday work and sometimes struggled with interaction techniques.

An area of particular interest that requires further study is that of the relation between effort expectancy and technology fit, surrounding both task and individual. While this is looked at in this work (see Chapter 7), it is done with reference to the learner's experience; the key construct investigated is task-individual fit, in this case relating to the level of competence that the trainer possesses and, by inference, is able to impart to the ECS.

In general, much more can be done to explore the relationship between the trainer's experience of a task and the one they are able to impart using this technology.

References

[1] Daniel F Abawi, Ralf Dorner and Paul Grimm. "A Component-based Authoring Environment for Creating Multimedia-Rich Mixed Reality". In: *Proceedings of the Seventh Eurographics conference on Multimedia*. 2004, pp. 31–40. DOI: 10.2312/EGMM/MM04/031-040.

[2] Icek Ajzen. "From intentions to actions: A theory of planned behavior". In: *Action control*. Springer, 1985, pp. 11–39. DOI: 10.1007/978-3-642-69746-3_2.

[3] Icek Ajzen and Thomas J Madden. "Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control". In: *Journal of experimental social psychology* 22.5 (1986), pp. 453–474. DOI: 10.1016/0022-1031(86)90045-4.

[4] Lorin W Anderson and David R Krathwohl. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman, 2001. ISBN: 0321084055.

[5] Ronald T Azuma. "A survey of augmented reality". In: *Presence: teleoperators & virtual environments* 6.4 (1997), pp. 355–385. DOI: 10.1162/pres.1997.6.4.355. URL: https://doi.org/10.1162/pres.1997.6.4.355.

[6] Juhi Baskar. "HoloPHAM: An Augmented Reality Training System For Upper Limb Myoelectric Prosthesis Users". PhD thesis. Johns Hopkins University, 2017. URL: https://api.semanticscholar.org/CorpusID:64967273.

[7] C. K. Battye, A. Nightingale and James Whillis. "The use of myo-electric currents in the operation of prostheses." In: *The Journal of bone and joint surgery. British volume* 37-B 3 (1955), pp. 506–10. DOI: 10.1302/0301-620X.37B3.506. URL: https://api.semanticscholar.org/CorpusID:29926268.

[8] Steve Benford et al. "Shared spaces: Transportation, artificiality, and spatiality". In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. 1996, pp. 77–86. DOI: 10.1145/240080.240196.

[9] Mark Billinghurst, Adrian Clark, Gun Lee et al. "A survey of augmented reality". In: *Foundations and Trends in Human–Computer Interaction* 8.2-3 (2015), pp. 73–272. DOI: 10.1561/1100000049.

[10] Jonas Blattgerste, Patrick Renner and Thies Pfeiffer. "Authorable augmented reality instructions for assistance and training in work environments". In: *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. Pisa Italy: ACM, Nov. 2019, pp. 1–11. ISBN: 978-1-4503-7624-2. DOI: 10.1145/3365610.3365646. URL: https://dl.acm.org/doi/10.1145/3365610.3365646.

[11] Markus Blut et al. "Meta-Analysis of the Unified Theory of Acceptance and Use of Technology (UTAUT): Challenging Its Validity and Charting a Research Agenda in the Red Ocean". In: *Journal of the Association for Information Systems* 22.9 (2021), pp. 824–867. DOI: 10.17705/1jais.00654.

[12] Tony Bonen and Matthias Oschinksi. *Mapping Canada's Training Ecosystem: Much Needed and Long Overdue*. Institute for Research on Public Policy, 2021. DOI: 10.13140/RG.2.2.13331.76327.

[13] Eleonora Bottani and Giuseppe Vignali. "Augmented reality technology in the manufacturing industry: A review of the last decade". In: *IISE Transactions* 51.3 (2019), pp. 284–310. DOI: 10.1080/24725854.2018.1493244.

[14] Ilona Buchem, Ralf Klamma and Fridolin Wild. "Introduction to wearable enhanced learning (WELL): trends, opportunities, and challenges". In: *Perspectives on Wearable Enhanced Learning (WELL) Current Trends, Research, and Practice* (2019), pp. 3–32. DOI: 10.1007/978-3-319-64301-4_1.

[15] Frank Buschmann, Kevlin Henney and Douglas C. Schmidt. "Pattern-Oriented Software Architecture, a Pattern Language for Distributed Computing". In: 2007. URL: http://www.researchandmarkets.com/reports/2251154/patternoriented_software_architecture_a_pattern.pdf.

[16] Jack CP Cheng, Keyu Chen and Weiwei Chen. "State-of-the-art review on mixed reality applications in the AECO industry". In: *Journal of Construction Engineering and Management* 146.2 (2020), p. 03119009. DOI: 10.1061/(asce)co.1943-7862.0001749.

[17] Subramanian Chidambaram et al. "ProcessAR: An augmented reality-based tool to create in-situ procedural 2D/3D AR Instructions". In: *Designing Interactive Systems Conference 2021*. Virtual Event USA: ACM, June 2021, pp. 234–249. ISBN: 978-1-4503-8476-6. DOI: 10.1145/3461778.3462126. URL: https://dl.acm.org/doi/10.1145/3461778.3462126.

[18] Michael Clark. *Factor Analysis with the psych Package*. https://m-clark.github.io/posts/2020-04-10-psych-explained/. 2020.

[19] European Commission, Social Affairs Directorate-General for Employment and Inclusion. *ESCO, European classification of skills/competences, qualifications and occupations – The first public release – A Europe 2020 initiative*. Publications Office, 2013. DOI: 10.2767/76494.

[20] Dan Curtis et al. "Several devils in the details: making an AR application work in the airplane factory". In: *Proc. Int'l Workshop Augmented Reality*. 1999, pp. 47–60. DOI: 10.1201/9781439863992-11.

[21] F. Cuzzolin et al. "Metric learning for Parkinsonian identification from IMU gait measurements". In: *Gait Posture* 54 (May 2017), pp. 127–132. DOI: 10.1016/j.gaitpost.2017.02.012.

[22] Aparecido Fabiano Pinatti De Carvalho et al. "Of embodied action and sensors: Knowledge and expertise sharing in industrial set-up". In: *Computer Supported Cooperative Work (CSCW)* 27 (2018), pp. 875–916. DOI: 10.1007/s10606-018-9320-6.

[23] Mandy R Drew, Brooke Falcone and Wendy L Baccus. "What does the system usability scale (SUS) measure?" In: *International Conference of Design, User Experience, and Usability*. Springer. 2018, pp. 356–366. DOI: 10.1007/978-3-319-91797-9_25.

[24] Andreas Dunser et al. "Creating interactive physics education books with augmented reality". In: *Proceedings of the 24th Australian Computer-Human Interaction Conference on - OzCHI 12*. Melbourne, Australia: ACM Press, 2012, pp. 107–114. ISBN: 978-1-4503-1438-1. DOI: 10.1145/2414536.2414554. URL: http://dl.acm.org/citation.cfm?doid=2414536.2414554.

[25] Yogesh K Dwivedi et al. "A meta-analysis based modified unified theory of acceptance and use of technology (meta-UTAUT): a review of emerging literature". In: *Current opinion in psychology* 36 (2020), pp. 13–18. DOI: 10.1016/j.copsyc.2020.03.008.

[26] Yogesh K Dwivedi et al. "Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model". In: *Information Systems Frontiers* 21.3 (2019), pp. 719–734. DOI: 10.1007/s10796-017-9774-y.

[27] Johannes Egger and Tariq Masood. "Augmented reality in support of intelligent manufacturing–a systematic literature review". In: *Computers & Industrial Engineering* 140 (2020), p. 106195.

[28] Guangzheng Fei, Xin Li and Rui Fei. "AVATAR: autonomous visual authoring of tangible augmented reality". In: *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry - VRCAI 2012*. Singapore, Singapore: ACM Press, 2012, p. 21. ISBN: 978-1-4503-1825-9. DOI: 10.1145/2407516.2407521. URL: http://dl.acm.org/citation.cfm?doid=2407516.2407521.

[29] Alessandro Filippeschi et al. "Survey of Motion Tracking Methods Based on Inertial Sensors: A Focus on Upper Limb Human Motion". In: *Sensors (Basel, Switzerland)* 17 (2017). DOI: 10.3390/s17061257.

[30] Martin Fishbein and Icek Ajzen. "Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research". In: *Philosophy and Rhetoric* 10.2 (1977), pp. 130–132. DOI: 10.2307/2065853.

[31] Paula Fraga-Lamas et al. "A review on industrial augmented reality systems for the industry 4.0 shipyard". In: *IEEE Access* 6 (2018), pp. 13358–13375. DOI: 10.1109/ACCESS.2018.2808326.

[32] Juan Garzón, Juan Pavón and Silvia Baldiris. "Systematic review and meta-analysis of augmented reality in educational settings". In: *Virtual Reality* 23.4 (2019), pp. 447–459.

[33] Paul van Geert and Henderien Steenbeek. "The dynamics of scaffolding". In: *New ideas in Psychology* 23.3 (2005), pp. 115–128. DOI: 10.1016/J.NEWIDEAPSYCH.2006.05.003.

[34] Yiannis Georgiou and Eleni A Kyza. "Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings". In: *Computers in Human Behavior* 89 (2018), pp. 173–181. DOI: 10.1016/j.chb.2018.08.011.

[35] James J Gibson. "The theory of affordances". In: *Psychological Review* 1.2 (1977), pp. 67–82. DOI: 10.4324/9781315816852-21.

[36] James Jerome Gibson. *The Ecological Approach to Visual Perception.* Lawrence Erlbaum Associates, 1979. ISBN: 9780898599596.

[37] Dale L Goodhue and Ronald L Thompson. "Task-Technology Fit and Individual Performance". In: *MIS quarterly* (1995), pp. 213–236. DOI: 10.2307/249689.

[38] Henry Gray. "Antomy of the human body". In: *Annals of Surgery* 68.5 (1918), pp. 564–566. DOI: 10.1097/00000658-191811000-00012.

[39] Jens Grubert et al. "Annotate that: Preparing event posters for augmentation". In: *ISMAR 2011 Workshop on Authoring Solutions for Augmented Reality.* 2011.

[40] Will Guest et al. "Affordances for Capturing and Re-enacting Expert Performance with Wearables". In: *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12.* Springer. 2017, pp. 403–409. DOI: 10.1007/978-3-319-66610-5_34.

[41] Will Guest et al. "Architecture and Design Patterns for Distributed, Scalable Augmented Reality and Wearable Technology Systems". In: *2019 IEEE International Conference on Engineering, Technology and Education (TALE).* IEEE. 2019, pp. 1–8. DOI: 10.1109/TALE48000.2019.9225855.

[42] Will Guest et al. "Learning in the Real World using Augmented Reality and Wearable Technology". In: *The Envisioning Report for Empowering Universities.* Ed. by G. Ubachs. Maastricht, NL: EADTU, 2021, pp. 35–37.

[43] William Guest et al. "A Technology Acceptance Model for Augmented Reality and Wearable Technologies". In: *Journal of Universal Computer Science* 24.2 (2018), pp. 192–219. URL: http://oro.open.ac.uk/71033/1/71033.pdf.

[44] Michael Haller, Erwin Stauder and Juergen Zauner. "AMIRE: Authoring Mixed Reality once, run it anywhere". In: *Proceedings of the 11th International Conference on Human-Computer Interaction (HCII)*. Vol. 2005. 2005. URL: https://api.semanticscholar.org/CorpusID:15520298.

[45] Han Kyu Yoo and Jong Weon Lee. "Mobile augmented reality system for in-situ 3D modeling and authoring". en. In: *2014 International Conference on Big Data and Smart Computing (BIGCOMP)*. Bangkok, Thailand: IEEE, Jan. 2014, pp. 282–285. ISBN: 978-1-4799-3919-0. DOI: 10.1109/BIGCOMP.2014.6741453. URL: http://ieeexplore.ieee.org/document/6741453/.

[46] Mort Heilig. "Beginnings: Sensorama and the Telesphere Mask". In: *Digital Illusion: Entertaining the Future with High Technology*. 1998, pp. 343–351.

[47] Kaj Helin et al. "User experience of augmented reality system for astronaut's manual work support". In: *Frontiers in Robotics and AI* 5 (2018), p. 106. DOI: 10.3389/frobt.2018.00106.

[48] Kaj Helin et al. "Virtual/mixed/augmented reality laboratory research for the study of augmented human and human-machine systems". In: *2016 12th International Conference on Intelligent Environments (IE)*. IEEE. 2016, pp. 163–166. DOI: 10.1109/IE.2016.35.

[49] Ángel Herrero, Héctor San Martín et al. "Explaining the adoption of social networks sites for sharing user-generated content: A revision of the UTAUT2". In: *Computers in Human Behavior* 71 (2017), pp. 209–217. DOI: 10.1016/j.chb.2017.02.007.

[50] María-Blanca Ibáñez and Carlos Delgado-Kloos. "Augmented reality for STEM learning: A systematic review". In: *Computers & Education* 123 (2018), pp. 109–123. DOI: 10.1016/j.compedu.2018.05.002.

[51] Thalmic Labs Inc. *Myo Armband: A Wearable Gesture Control Device*. https://web.archive.org/web/20140312142915/https://www.thalmic.com/en/myo/. 2016.

[52] Jordi Janer, Gerard Roma and Stefan Kersten. "Authoring augmented soundscapes with user-contributed content". In: *ISMAR workshop on authoring solutions for augmented reality*. 2011. URL: http://mtg.upf.edu/system/files/publications/ISMAR_auth_mtg.pdf.

[53] Hyung-Keun Jee et al. "An immersive authoring tool for augmented reality-based e-learning applications". In: *2011 International Conference on Information Science and Applications*. IEEE. 2011, pp. 1–5. DOI: 10.1109/ICISA.2011.5772399.

[54] Qiao Jin, Danli Wang and Fang Sun. "TanCreator: A Tangible Tool for Children to Create Augmented Reality Games". In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. Singapore Singapore: ACM, Oct. 2018, pp. 82–85. ISBN: 978-1-4503-5966-5. DOI: 10.1145/3267305.3267603. URL: https://dl.acm.org/doi/10.1145/3267305.3267603.

[55] Li Jinyu et al. "Survey and evaluation of monocular visual-inertial SLAM algorithms for augmented reality". In: *Virtual Reality & Intelligent Hardware* 1.4 (2019), pp. 386–410. DOI: 10.1016/j.vrih.2019.07.002.

[56] Jonathan M Kevan and Paul R Ryan. "Experience API: Flexible, decentralized and activity-centric data collection". In: *Technology, knowledge and learning* 21.1 (2016), pp. 143–149. DOI: 10.1007/s10758-015-9260-x.

[57] Byeong Jeong Kim and Seop Hyeong Park. "CARDA: Content Management Systems for Augmented Reality with Dynamic Annotation". In: 2015, pp. 62–67. DOI: 10.14257/astl.2015.90.14. URL: http://onlinepresent.org/proceedings/vol90_2015/14.pdf.

[58] Sonny E. Kirkley and Jamie R. Kirkley. "Creating next generation blended learning environments using mixed reality, Video Games and Simulations". In: *TechTrends* 49.3 (May 2005), pp. 42–53. ISSN: 8756-3894, 1559-7075. DOI: 10.1007/BF02763646. URL: http://link.springer.com/10.1007/BF02763646.

[59] Eric Klopfer and Josh Sheldon. "Augmenting your own reality: Student authoring of science-based augmented reality games". In: *New Directions for Youth Development* 2010 (Dec. 2010), pp. 85–94. DOI: 10.1002/yd.378. URL: https://onlinelibrary.wiley.com/doi/10.1002/yd.378.

[60] Milos Kravcík et al. "Requirements for Scenarios and Prototypes". In: 2016. URL: https://api.semanticscholar.org/CorpusID:115580440.

[61] Bernard C Kress and Ishan Chatterjee. "Waveguide combiners for mixed reality headsets: a nanophotonics design perspective". In: *Nanophotonics* 1.ahead-of-print (2020). DOI: 10.1515/nanoph-2020-0410.

[62] Kai Kunze and Paul Lukowicz. "Sensor placement variations in wearable activity recognition". In: *IEEE Pervasive Computing* 13.4 (2014), pp. 32–41. DOI: 10.1109/MPRV.2014.73.

[63] Kari Kuutti and Liam J Bannon. "The turn to practice in HCI: towards a research agenda". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. Ed. by Jonathan Grudin and Fabio Paterno. Toronto, Ontario, Canada, 2014, pp. 3543–3552. DOI: 10.1145/2556288.2557111.

[64] Tobias Langlotz et al. "Sketching up the world: in situ authoring for mobile Augmented Reality". In: *Personal and Ubiquitous Computing* 16.6 (Aug. 2012), pp. 623–630. ISSN: 1617-4909, 1617-4917. DOI: 10.1007/s00779-011-0430-0. URL: http://link.springer.com/10.1007/s00779-011-0430-0.

[65] Moez Limayem, Sabine Gabriele Hirt and Christy MK Cheung. "How habit limits the predictive power of intention: The case of information systems continuance". In: *Management Information Systems Quarterly* (2007), pp. 705–737. URL: https://api.semanticscholar.org/CorpusID:215932432.

[66] Bibeg Limbu et al. "A Conceptual Framework for Supporting Expertise Development with Augmented Reality and Wearable Sensors". In: *Perspectives on wearable enhanced learning (WELL) current trends, research, and practice* (2019), pp. 213–228. DOI: 10.1007/978-3-319-64301-4_10.

[67] Bibeg Limbu et al. "From AR to Expertise: A User Study of an Augmented Reality Training to Support Expertise Development." In: *Journal of Universal Computer Science* 24.2 (2018), pp. 108–128. URL: http://www.jucs.org/jucs_24_2/from_ar_to_expertise.

[68] Bibeg Limbu et al. "WEKIT. one: a sensor-based augmented reality system for experience capture and re-enactment". In: *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14*. Springer. 2019, pp. 158–171. DOI: 10.1007/978-3-030-29736-7_12.

[69] Bibeg Hang Limbu et al. "Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review". In: *Educational Research Review* 25 (2018), pp. 1–22. DOI: 10.1016/J.EDUREV.2018.07.001.

[70] Thomas Ludwig, Peter Tolmie and Volkmar Pipek. "From the Internet of Things to an Internet of Practices". In: *Social Internet of Things* (2019), pp. 33–47. DOI: 10.18420/ecscw2017-10.

[71] Niklas Luhmann. *Social Systems*. Stanford, California: Stanford University Press, 1995. ISBN: 9780804719933.

[72] Chris Lytridis, Avgoustos Tsinakos and Ioannis Kazanidis. "ARTutor-an augmented reality platform for interactive distance learning". In: *Education Sciences* 8.1 (2018), p. 6. DOI: 10.3390/EDUCSCI8010006.

[73] Steffen Mader and Bodo Urban. "Creating Instructional Content for Augmented Reality Based on Controlled Natural Language Concepts". In: *Proceedings of 20th International Conference on Artificial Reality and Telexistence (ICAT 2010)*. 2010, pp. 313–321. URL: http://icat.vrsj.org/ICAT2010_Proceedings/paper/paper-002.pdf.

[74] Sebastian Madgwick. "An efficient orientation filter for inertial and inertial/magnetic sensor arrays". In: *Report x-io and University of Bristol (UK)* 25 (2010), pp. 113–118. URL: http://x-io.co.uk/res/doc/madgwick_internal_report.pdf.

[75] Andrew Maimone and Henry Fuchs. "Computational augmented reality eyeglasses". In: *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2013, pp. 29–38. DOI: 10.1109/ISMAR.2013.6671761. URL: https://api.semanticscholar.org/CorpusID:7637250.

[76] G. Gales E. McClean and John B. McDonald. "A facade based approach to non-expert authoring in urban augmented reality". In: 2012. URL: https://api.semanticscholar.org/CorpusID:61704365.

[77] David McNeill. *Hand and mind: What gestures reveal about thought*. University of Chicago press, 1992. DOI: 10.2307/415833.

[78] Paul Milgram et al. "Augmented reality: A class of displays on the reality-virtuality continuum". In: *Telemanipulator and telepresence technologies*. Vol. 2351. International Society for Optics and Photonics. 1995, pp. 282–292.

[79] Mohsen Moghaddam et al. "Exploring augmented reality for worker assistance versus training". In: *Advanced Engineering Informatics* 50 (2021), p. 101410. DOI: 10.1016/j.aei.2021.101410.

[80] David Moher et al. "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Chinese edition)". In: *Journal of Chinese Integrative Medicine* 7.9 (2009), pp. 889–896. DOI: 10.1136/bmj.b2535.

[81] Vinh T. Nguyen, Kwanghee Jung and Tommy Dang. "BlocklyAR: A Visual Programming Interface for Creating Augmented Reality Experiences". In: *Electronics* 9.8 (July 2020), p. 1205. ISSN: 2079-9292. DOI: 10.3390/electronics9081205. URL: https://www.mdpi.com/2079-9292/9/8/1205.

[82] Carleandro Noleto et al. "An Authoring Tool for Location-Based Mobile Games with Augmented Reality Features". In: *2015 14th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. Piauí, Brazil: IEEE, Nov. 2015, pp. 99–108. ISBN: 978-1-4673-8843-6. DOI: 10.1109/SBGames.2015.12. URL: http://ieeexplore.ieee.org/document/7785846/.

[83] Don Norman. "Affordances and design". In: (2004), p. 14. URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=538961f9812c70d20c1f48042bab1e8060366ca2.

[84] Donald A Norman. *The psychology of everyday things*. Basic books, 1988. ISBN: 1116535246.

[85] Riccardo Palmarini et al. "A systematic review of augmented reality applications in maintenance". In: *Robotics and Computer-Integrated Manufacturing* 49 (2018), pp. 215–228.

[86] Alison Parkes. "The effect of task–individual–technology fit on user attitude and performance: An experimental investigation". In: *Decision support systems* 54.2 (2013), pp. 997–1009. DOI: 10.1016/j.dss.2012.10.025.

[87] Antigoni Parmaxi and Alan A Demetriou. "Augmented reality in language learning: A state-of-the-art review of 2014–2019". In: *Interactive Learning Environments* 31.1 (2020), pp. 861–875. DOI: 10.1080/10494820.2020.1765392. URL: https://www.tandfonline.com/doi/full/10.1080/10494820.2020.1765392.

[88] Michael Polanyi. *The Tacit Dimension*. Chicago: University of Chicago Press, 1966. ISBN: 9780226672984. URL: https://press.uchicago.edu/ucp/books/book/chicago/T/bo6035368.html.

[89] Ivan Poupyrev et al. "Tiles: A mixed reality authoring interface". In: *(2001) INTERACT 2001 Conference on Human Computer Interaction*. 2001. URL: http://www.ivanpoupyrev.com/e-library/2001/Interact2001_tiles.pdf.

[90] Mark Ransley et al. "WEKIT Wearable Design Solutions". In: *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (2016).

[91] Jens Rasmussen. "Constructivism and Phenomenology What Do They Have in Common, and How Can They Be Told Apart?" In: *Cybernetics & Systems* 29.6 (1998), pp. 553–576. DOI: 10.1080/019697298125515.

[92] Rocco Raso, Dirk Werth and Peter Loos. "Enriching Augmented Reality with Text Data Mining: An Automated Content Management System to Develop Hybrid Media Applications". In: 2017. DOI: 10.24251/HICSS.2017.746. URL: http://hdl.handle.net/10125/41910.

[93] Dominik Renzel et al. "Requirements bazaar: Social requirements engineering for community-driven innovation". In: *2013 21st IEEE International Requirements Engineering Conference (RE)*. IEEE. 2013, pp. 326–327. DOI: 10.1109/RE.2013.6636738.

[94] William Revelle. "Hierarchical cluster analysis and the internal structure of tests". In: *Multivariate Behavioral Research* 14.1 (1979), pp. 57–74. DOI: 10.1207/s15327906mbr1401_4.

[95] Daniel Roetenberg, Henk Luinge, Per Slycke et al. "Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors". In: *Xsens Motion Technologies BV, Tech. Rep* 1.2009 (2009), pp. 1–7. URL: http://edge.rit.edu/content/P10010/public/PDF/MVN_white_paper.pdf.

[96] Everett M Rogers. *Diffusion of innovations*. Simon and Schuster, 2010. ISBN: 9781451602470.

[97] Dariusz Rumiński and Krzysztof Walczak. "Creation of interactive AR content on mobile devices". In: *Business Information Systems Workshops: BIS 2013 International Workshops, Poznan, Poland, June 19-20, 2013, Revised Papers 16*. Springer. 2013, pp. 258–269. DOI: 10.1007/978-3-642-41687-3_24.

[98] Gilbert Ryle. "Knowing how and knowing that: The presidential address". In: *Proceedings of the Aristotelian society*. Vol. 46. JSTOR. 1945, pp. 1–16. DOI: 10.1093/ARISTOTELIAN/46.1.1.

[99] Maryam Safi, Joon Chung and Pratik Pradhan. "Review of augmented reality in aerospace industry". In: *Aircraft Engineering and Aerospace Technology* (2019).

[100] Christian Sandor et al. "Authoring augmented reality as situated multimedia". In: *22nd International Conference on Computers in Education (ICCE 2014)*. Asia-Pacific Society for Computers in Education. 2014, pp. 554–556. URL: https://api.semanticscholar.org/CorpusID:62956772.

[101] Akira Sano. "A Web Application for Creating Real-Size Augmented Reality Content without 3D Modeling Skills". In: *The Journal of Information and Systems in Education* 10.1 (2012), pp. 32–38. ISSN: 1348-236X, 2186-3679. DOI: 10.12937/ejsise.10.32. URL: https://www.jstage.jst.go.jp/article/ejsise/10/1/10_32/_article.

[102] Akira Sano. "An application for creating full-scale augmented reality content without 3d modeling skills". In: *Proceedings of the 2011 ACM symposium on The role of design in UbiComp research & practice - RDURP '11*. Beijing, China: ACM Press, 2011, p. 19. ISBN: 978-1-4503-0931-8. DOI: 10.1145/2030031.2030039. URL: http://dl.acm.org/citation.cfm?doid=2030031.2030039.

[103] Willem E Saris, Albert Satorra and William M Van der Veld. "Testing structural equation models or detection of misspecifications?" In: *Structural equation modeling* 16.4 (2009), pp. 561–582. DOI: 10.1080/10705510903203433.

[104] Margaret Schenkman and Victoria Rugo de Cartaya. "Kinesiology of the shoulder complex". In: *Journal of Orthopaedic & Sports Physical Therapy* 8.9 (1987), pp. 438–450. DOI: 10.2519/jospt.1987.8.9.438.

[105] Jonathan A Schlueter. "Remote maintenance assistance using real-time augmented reality authoring". Pages: 12331520. Master of Science. Ames: Iowa State University, Digital Repository, 2018. DOI: 10.31274/etd-180810-6087. URL: https://lib.dr.iastate.edu/etd/16457/.

[106] Kjeld Schmidt. "Computer-Supported Cooperative Work (CSCW)". In: *The International Encyclopedia of Communication Theory and Philosophy* 12 (2016), pp. 1–4. DOI: 10.1002/9781118766804.wbiect144.

[107] Kjeld Schmidt. “The trouble with ‘tacit knowledge’”. In: *Computer supported cooperative work (CSCW)* 21 (2012), pp. 163–225. DOI: 10.1007/s10606-012-9160-8.

[108] Waseem Shahzad et al. “Enhanced Performance for Multi-Forearm Movement Decoding Using Hybrid IMU–sEMG Interface”. In: *Frontiers in neurorobotics* 13 (2019). DOI: 10.3389/fnbot.2019.00043.

[109] Puneet Sharma, Roland Klemke and Fridolin Wild. “Experience Capturing with Wearable Technology in the WEKIT Project”. In: *Perspectives on Wearable Enhanced Learning (WELL)*. Springer, 2019, pp. 297–311. DOI: 10.1007/978-3-319-64301-4_14.

[110] Sumit Shekhar et al. “ARComposer: Authoring Augmented Reality Experiences through Text”. en. In: *The Adjunct Publication of the 32nd Annual ACM Symposium on User Interface Software and Technology*. New Orleans LA USA: ACM, Oct. 2019, pp. 63–65. ISBN: 978-1-4503-6817-9. DOI: 10.1145/3332167.3357116. URL: https://dl.acm.org/doi/10.1145/3332167.3357116.

[111] Dexin Shi, Alberto Maydeu-Olivares and Yves Rosseel. “Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 27.1 (2020), pp. 1–15. URL: https://api.semanticscholar.org/CorpusID:198455373.

[112] Jinwook Shim et al. “Gesture-based interactive augmented reality content authoring system using HMD”. en. In: *Virtual Reality* 20.1 (Mar. 2016), pp. 57–69. ISSN: 1359-4338, 1434-9957. DOI: 10.1007/s10055-016-0282-z. URL: http://link.springer.com/10.1007/s10055-016-0282-z.

[113] Jens de Smit. “Layar Creator: Self-service AR Authoring for Print”. en. In: *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2nd IEEE ISMAR workshop on authoring solutions for Augmented Reality. IEEE, 2012, p. 3. URL: https://ieeexplore.ieee.org/document/6483978.

[114] Luis Fernando de Souza Cardoso, Flávia Cristina Martins Queiroz Mariano and Ezequiel Roberto Zorzal. “A survey of industrial augmented reality”. In: *Computers & Industrial Engineering* 139 (2020), p. 106159. DOI: 10.1016/j.cie.2019.106159.

[115] Gheric Speiginer et al. “The Evolution of the Argon Web Framework Through Its Use Creating Cultural Heritage and Community-Based Augmented Reality Applications”. en. In: *Human-Computer Interaction: Users and Contexts*. Ed. by Masaaki Kurosu. Vol. 9171. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 112–124. ISBN: 978-3-319-21005-6 978-3-319-21006-3. DOI: 10.1007/978-3-319-21006-3_12. URL: http://link.springer.com/10.1007/978-3-319-21006-3_12.

[116] Jacob H Steffen et al. "Framework of affordances for virtual reality and augmented reality". In: *Journal of Management Information Systems* 36.3 (2019), pp. 683–729. DOI: 10.1080/07421222.2019.1628877.

[117] Lucy A Suchman. *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge university press, 1987. DOI: 10.2307/1423221.

[118] Anna Syberfeldt, Oscar Danielsson and Patrik Gustavsson. "Augmented reality smart glasses in the smart factory: Product evaluation guidelines and review of available products". In: *IEEE Access* 5 (2017), pp. 9118–9130. URL: https://ieeexplore.ieee.org/iel7/6287639/6514899/07927376.pdf.

[119] Aaron Tabor et al. "Designing Game-Based Myoelectric Prosthesis Training". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 1352–1363. ISBN: 9781450346559. DOI: 10.1145/3025453.3025676. URL: https://doi.org/10.1145/3025453.3025676.

[120] Kuttimani Tamilmani et al. "Exploring the Role of 'Price Value' for Understanding Consumer Adoption of Technology: A Review and Meta-analysis of UTAUT2 based Empirical Studies". In: *PACIS 2018 Proceedings*. 2018. URL: https://aisel.aisnet.org/pacis2018/64.

[121] Timothy Teo and Jan Noyes. "An assessment of the influence of perceived enjoyment and attitude on the intention to use technology among pre-service teachers: A structural equation modeling approach". In: *Computers & Education* 57.2 (2011), pp. 1645–1653. URL: https://api.semanticscholar.org/CorpusID:29052549.

[122] Thalmic Labs Inc. *Myo Armband SDK Documentation*. Accessed: 2023-02-18. Thalmic Labs Inc., 2013–2018. URL: https://developer.thalmic.com/docs/.

[123] Jayzon F Ty, Ma Mercedes T Rodrigo and Marc Ericson C Santos. "A Mobile Authoring Tool for Augmented Reality Content Generation Using Images as Annotations". In: *Philippine Information Technology Journal* (2014). URL: https://archium.ateneo.edu/discs-faculty-pubs/8/.

[124] Viswanath Venkatesh, James Y.L. Thong and Xin Xu. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology". In: *MIS Quarterly* 36.1 (2012), pp. 157–178. DOI: 10.2307/41410412. URL: https://www.jstor.org/stable/41410412.

[125] Viswanath Venkatesh et al. "User Acceptance of Information Technology: Toward a Unified View". In: *MIS Quarterly* 27.3 (2003), pp. 425–478. DOI: 10.2307/30036540. URL: https://www.jstor.org/stable/30036540.

[126] Daniel Vlasic et al. “Practical Motion Capture in Everyday Surroundings”. In: *ACM Trans. Graph.* 26.3 (July 2007), 35–es. ISSN: 0730-0301. DOI: 10.1145/1276377.1276421. URL: https://doi.org/10.1145/1276377.1276421.

[127] Olga Volkoff and Diane M Strong. “Affordance theory and how to use it in IS research”. In: *The Routledge companion to management information systems* (2017). Ed. by Munir Khan, Paul Waterson and Ajmal Chaudhry, pp. 232–245. DOI: 10.4324/9781315619361.CH16.

[128] Alla Vovk et al. “SIG: spatiality of augmented reality user interfaces”. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–4. DOI: 10.1145/3290607.3311756.

[129] Alla Vovk et al. “Simulator sickness in augmented reality training using the Microsoft HoloLens”. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–9. DOI: 10.1145/3173574.3173783.

[130] Ming-Jen Wang, Chien-Hao Tseng and Cherng-Yeu Shen. “An Easy to Use Augmented Reality Authoring Tool for Use in Examination Purpose”. In: *Human-Computer Interaction*. Ed. by Peter Forbrig, Fabio Paterno and Annelise Mark Pejtersen. Vol. 332. Series Title: IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 285–288. ISBN: 978-3-642-15230-6 978-3-642-15231-3. DOI: 10.1007/978-3-642-15231-3_31. URL: http://link.springer.com/10.1007/978-3-642-15231-3_31.

[131] Tianyi Wang et al. “CAPturAR: An Augmented Reality Tool for Authoring Human-Involved Context-Aware Applications”. In: *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. Virtual Event USA: ACM, Oct. 2020, pp. 328–341. ISBN: 978-1-4503-7514-6. DOI: 10.1145/3379337.3415815. URL: https://dl.acm.org/doi/10.1145/3379337.3415815.

[132] Alexander Weinhard, Matthias Hauser and Frédéric Thiesse. “Explaining Adoption of Pervasive Retail Systems with a Model Based on UTAUT2 and the Extended Privacy Calculus”. In: *PACIS 2017 Proceedings*. 2017, p. 217. URL: https://aisel.aisnet.org/pacis2017/217.

[133] Fridolin Wild et al. “An augmented reality job performance aid for kinaesthetic learning in manufacturing work places”. In: *Open Learning and Teaching in Educational Communities: 9th European Conference on Technology Enhanced Learning, EC-TEL 2014, Graz, Austria, September 16-19, 2014, Proceedings 9*. Springer. 2014, pp. 470–475. DOI: 10.1007/978-3-319-11200-8_44.

[134] Fridolin Wild et al. “IEEE standard for augmented reality learning experience models”. In: *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE. 2020, pp. 1–3. URL: https://www.researchgate.net/profile/Fridolin-Wild/

publication/349905576_IEEE_Standard_for_Augmented_Reality_Learning_Experience_Models/links/607442aaa6fdcc5f779d0dc2/IEEE-Standard-for-Augmented-Reality-Learning-Experience-Models.pdf.

[135] Fridolin Wild et al. "Technology acceptance of augmented reality and wearable technologies". In: *International Conference on Immersive Learning*. Springer. 2017, pp. 129–141. DOI: 10.1007/978-3-319-60633-0_11.

[136] Bob G Witmer and Michael J Singer. "Measuring presence in virtual environments: A presence questionnaire". In: *Presence* 7.3 (1998), pp. 225–240. DOI: 10.1162/105474698565686.

[137] Mike Wozniewski and Paul Warne. "Towards in situ authoring of augmented reality content". In: *Proceedings of the 2011 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2011. URL: http://mikewoz.com/publications/Farrago-ISMAR-2011.pdf.

[138] Dong Woo Yoo et al. "Augmenting Learning with Augmented Reality: Exploring the Affordances of AR in Supporting Mastery of Complex Psychomotor Tasks". In: *arXiv preprint arXiv:2305.09875* (2023). DOI: 10.48550/arXiv.2305.09875.

[139] Jürgen Zauner and Michael Haller. "Authoring of Mixed Reality Applications Including Multi-Marker Calibration for Mobile Devices". In: *EGVE*. 2004, pp. 87–90. DOI: 10.2312/EGVE/EGVE04/087-090.

[140] Pamela Zave. "Classification of research efforts in requirements engineering". In: *ACM Computing Surveys (CSUR)* 29.4 (1997), pp. 315–321. URL: https://ieeexplore.ieee.org/document/512563/.