Journal of Universal Computer Science, vol. 24, no. 2 (2018), 192-219
submitted: 1/12/17, accepted: 19/1/18, appeared: 28/2/18 © J.UCS

A Technology Acceptance Model for Augmented Reality and Wearable Technologies

William Guest, Fridolin Wild and Alla Vovk
(Oxford Brookes University, Oxford, United Kingdom
16102434@brookes.ac.uk, wild@brookes.ac.uk, 16022839@brookes.ac.uk)

Paul Lefrere
(CCA Ltd, Milton Keynes, United Kingdom
paul.lefrere@cca-research.co.uk)

Roland Klemke
(Open University of the Netherlands, Heerlen, Netherlands
Roland.Klemke@ou.nl)

Mikhail Fominykh
(Europlan UK Ltd, London, United Kingdom
mikhail.fominykh@europlan-uk.eu)

Timo Kuula
(VTT Technical Research Centre of Finland Ltd**,** Espoo, Finland
Timo.Kuula@vtt.fi)

Abstract: Leveraging Augmented Reality and wearable technology for knowledge-intensive training is thought to offer huge potential for improving human performance. The recent introduction of the technology means that much of this potential is untapped, though efforts are needed to understand what makes it useful, entertaining, and easy-to-use. The research presented in this article investigates the implementation of a combined hardware and software application in three use-cases: aviation, medical and space. Following the validation of metrics for a questionnaire, data was collected from 142 participants, and a structural equation model, based on UTAUT2, was proposed in order to interpret the data. Following model improvement, two constructs show significant factor loading and latent variable correlation, Interoperability and Augmented Reality / Wearable Technology Fit. Model optimisation was conducted, and a variety of goodness-of-fit indices are reported. The two additional constructs are found to be covariant and impact the UTAUT2 variables performance expectancy, effort expectancy and facilitating conditions, in some cases explaining more than 85% of the variance in those constructs ($p < 0.001$). A root mean square error of approximation of 0.047 after a 1000-fold Monte Carlo cross-validation indicates a good fit between the model and the data. In all other fit indices, a moderate power has been observed.

Keywords: Augmented Reality, Wearable Technology, Structural Equation Model, Technology Acceptance, UTAUT2, Microsoft Hololens
Categories: H.1.2, L.0, L.3.0, L.3.1, L.3.6

1 Introduction

Our focus is technology-acceptance models, TAM, for Augmented Reality (AR) and wearable technologies (WT) in training. Within that, we pay attention to TAM for knowledge-intensive training, KIT. We define KIT as training whose efficiency and effectiveness is affected and improved through a combination of theoretical and practical know-how. KIT typically involves an evidence-based deployment of various types of formalised knowledge. The training under consideration consists of tasks that require access to and manipulation of large quantities of such knowledge, and that make significant demands on our cognitive capabilities (e.g., short-term memory), so benefitting from repeated exposure where some of these demands are thought of as expertise.

Each wave of technologies used in KIT is built on its predecessors, which will modify expectations for the impact and suitability of the technology and depends, in part, on the level of adoption and familiarity with any such technology. This generational nature is not well reflected in models of technology acceptance, requiring them to be updated as major shifts (such as that recognised with AR) come about. Estimations of performance expectancy and effort expectancy (when using the technology's functions) are likely to be influenced by such factors, as is what is thought to be a facilitating condition in an organisational context and the degree of social influence that accompanies the use of such a technology. Metrics offered for a TAM are therefore expected to evolve with the spread and application of new technologies, such as holographic displays. Further, the degree of acceptance of each such technological wave is increased or decreased by the power and resonance of the stories we tell, incorporating past experience, current needs and prospective benefits. These elements offer both help and hindrance when using current-technology acceptance levels to plan next-generation technologies.

AR allows context-sensitive information and experiences to be offered to users, moment-by-moment, in ways that can direct their attention to pertinent facets of their environment by replacing or augmenting aspects of their sensory field (e.g., what they can see or hear or feel). Used in that way, AR is taken to be an enabling technology for Attention Management, which can help to reduce information overload and associated risk of errors. TAM can aid the design choices made by designers of AR spaces, for example, to model the fit between envisaged knowledge-intensive tasks facing a user, and the affordances of a given design of AR system, which may use combinations of stimuli, instantiated using a mix of technologies. Further, we anticipate that TAM can be extended to investigate the acceptance of attention-management tools and functions for collaborative work (human-human or AI-human), or the acceptance of personalisation tools to enable the capture and enhancement of annotatable and shareable recordings of each AR user's experiences. AR has the potential to leverage other enabling technologies such as holographic displays and wearable experience-capture and -replay, and to enhance the skills and understanding of people working in fast-changing and societally-important domains such as knowledge-intensive business services [Schnabl and Zenker 2013].

Currently (early 2018), Microsoft's HoloLens is a frontrunner in the development of hardware, offering an untethered system with depth scanning sensors, optical waveguides for visualisation of holographic objects and a free, accessible platform for app development. The capabilities of this class of devices, including gestural interaction, hand tracking, accurate image tracking and visual overlay, as well as interoperability with other WT has stimulated a high level of investment in the development of prototypes that can be used to capture, edit, replay and share experiences.

Against this background, the goal of today's KIT is to facilitate the transition from the specific knowledge and skills that are needed today, to the next generation of in-demand knowledge and skills. Ideally, AR-enhanced KIT would be somewhat future-proof, in the sense that self-directed learners can anticipate 'the next' state of knowledge that they should be aiming for, and devise informal training that gives them ways to acquire tomorrow's technologies and experiences early. Associated learning outcomes can include becoming a competent self-regulated learner and at recognising and remedying any gaps in their knowledge or shortcomings in their performance of tasks. In achieving this, trainers have a body of knowledge on how to direct and sustain the attention of learners to key aspects of the domain they are learning about and the skills, performance levels and insights that employers feel they need first to acquire and then to demonstrate, if they are to become acknowledged as competent in that domain.

When combining the above learner/trainer goals with a system intended to support each goal in real-time, especially systems consisting of wearable devices and holographic augmentation, we should consider that this kind of "attention management and support" is a qualitatively different process to that of training on a screen. The design elements under consideration can include reference to physical objects, allowing the location, orientation, state, appearance and usefulness of these objects to be modified to assist learning in the workplace. AR-enhanced training in the workplace changes the variety of affordances that are available to the user compared to a classroom environment, together with the memorability and retrievability of the experiences offered by the digital environment [Guest et al. 2017]. This has direct implications for the perceived utility whilst also impacting the learnability of the task at hand.

When assessing acceptance levels of AR and WT in KIT, the standard models (such as UTAUT2 – see section 3.2) might need extension. Additional constructs, such as those proposed by Wild, et al. [Wild, et al. 2017], are shown to be descriptive and are further developed here.

The rest of this article is structured as follows. A description of the software prototype and its application in three use cases forms section 2. Section 3, the research design, describes the demographic spread and presents the UTAUT2 model as a foundation of this work, together with the selection of metrics and the validation of their selection. Following in section 4 is a description of the structural equation model, and the formulation of both hypotheses and constructs describing the use of AR/WT in this context. Section 5 presents the findings from the model fitting and some practical considerations of applying these metrics to the use case and section 6 demonstrates the cross-validation processes used to test an improved model. Section 7 is a summary, with attention given to existing limitations and future work.

2 Context

The present work on TAM for AR/WT has been conducted in the frame of the project Wearable Experience for Knowledge Intensive Training (WEKIT, wekit.eu). In the project, a new training platform WEKIT.one is being developed and evaluated in three use cases.

2.1 Knowledge Intensive Training Use Cases

The three use cases of applying AR/WT for KIT included Aviation, Medical and Space. For each case, a real-life procedural scenario has been selected.

In the aviation case, a 10-step part of a pre-flight inspection has been used. The inspection procedure included such steps as such as checking baggage compartment, emergency exit handles and control locks. Pre-flight inspection is used to determine if the aircraft is in airworthy condition. In order to conduct a pre-flight inspection, a large amount of paperwork needs to be done, reference information gathered and studied before proceeding to the aircraft to conducting the inspection. The inspection itself takes a considerable amount of time, and the cost of errors can be very high.

In the medical case, a 13-step procedure for teaching how to perform an ultrasound carotid artery examination and take a simple measure using the MyLab8 ultrasound machine has been set up. The procedure included such steps as positioning the probe, selecting different modes on the machine, and taking measurements. The procedure has been taken from the educational libraries on operating the machine which are used in training of medical students and radiologist apprentices to perform an echographic examination.

In the space case, a 10-step procedure that foresees the installation of the Temporary Stowage Rack in the Automated Transfer Vehicle Part Task Trainer of the International Space Station has been used. The installation procedure included such steps as installing studs, connecting and fixing straps. All the components and mock-ups for the procedure are available and the training is conducted on Earth, but the astronauts perform it on the International Space Station.

2.2 The WEKIT.one Application

WEKIT.one is a software platform that allows capturing and delivering human performance using AR and WT. In KIT, the platform is used to capture expert performance by unifying incoming sensor data, and providing synchronisation, processing, and storage. The captured performance is then available for re-enactment by trainees.

The learning *Activities* or experiences are saved and re-enacted. The learning tasks are created and organised into scenarios that better define and delimit a learning experience and correlate it with related learning tasks that can be fetched and downloaded for later purposes. Interactive objects that are placed to describe a particular learning task to be executed by the learner to accomplish part of the scenario. An *Action* serves as a hub of annotations, and for this reason is in *one-to-n* relation with Annotation. WEKIT.one allows two versions of Actions: physical position based Actions and marker based Actions.

The recorder defines a JSON-based format for storing scenarios, task stations, and annotations. All annotations are stored in additional files using binary formats for audio data and images as well as XML formats for recorded sensor data. The Recorder allows for the export of data into the AR Learning Experience Model (ARLEM, IEEE p1589; http://www.techstreet.com/ieee/products/vendor_id/6073) format for exchange.

Figure 1 shows how to create initial Actions, to which annotations can be added. When a sensor annotation is added sensor recordings can be activated for this task station. The recorded data can be visualized for rehearsal including physical position, gaze direction and hand positions (Figure 2).

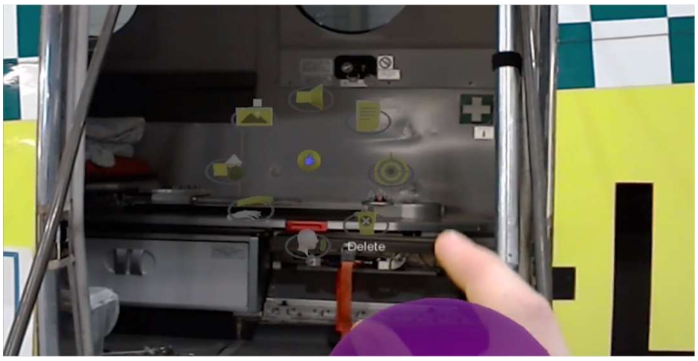

Figure 1: Starting to record a step, defining an action in space

The WEKIT.one Re-enactments system has also been developed in Unity3D and is configured around the Activity JSON and the Workplace JSON files of ARLEM. The workplace JSON describes workplace-related information such as point of interest, sensors, etc. It is parsed with the Workplace manager and information is transferred to the data layer. Activity JSON describes all action steps and what content should be active in each of these steps. It is parsed with the Activity Manager and information is transferred to the AR layer via local storage. The user can act with WEKIT.one Re-enactment system by exploiting a multi-modal User Interface. Three modalities can be used simultaneously: gesture (e.g., doing the 'click' gesture to go next work step), voice commands (e.g., saying "next" to go next work step or "show status" / "hide status"), and physical HoloLens click button (e.g., "click" to go next work step).

3 Research Design

The research design of the work presented in this article consists of two studies. The objective of the first study "Development and validation of TAM for AR/WT" was to develop a model for TAM of AR/WT and validate its metrics. The major outcome of this study was a questionnaire consisting of 19 questions - the validation of these metrics presented in Wild et al., 2017; result that are summarised in this section. In the second study, “Testing a TAM for AR/WT”, we used this questionnaire to derive a model and test several hypotheses relating to the technology acceptance of AR/WT.

Figure 2: Visualisation of a ghost track recording using physical position, orientation, and gaze direction

3.1 Participant Demographics

In the ‘Development and Validation of TAM for AR/WT’ study, we tested a group of 33 professionals in three knowledge-intensive areas [Wild, et al. 2017]. In the ‘Testing a TAM for AR/WT’ study, we put the same set of question to 142 participants in the same use case areas (Table 1).

	Aviation	Medical	Space	Total
Expert	21	9	17	47
Learner	34	39	22	95
Females	8	14	9	31
Males	47	35	30	112
Total	55	48	39	142

Table 1: Use case demographic spread

Age distribution across all test beds: 69 participants were between 18-24 years old, 48 were 25-32 years old, 11 were 35-44 years old, 9 were 45-54 years old, 5 were 55-64 years old, and none of the participants fell into over 65 years category. Many of the participants of the first study had no or little technology exposure, and all of the participants of the second study were experiencing an AR/WT training system for the first time. In terms of applications, participants were identified as either experts or learners. Experts were using a “recorder” application – a tool for experience capturing, learners were using a “player” application – a tool for experience re-enactment.

3.2 UTAUT2 Model of Technology Acceptance

UTAUT2 includes Performance Expectancy, Effort Expectancy, Social Influence, Facilitating Conditions, Hedonic Motivation, Price Value and Habit [Venkatesh, et al. 2012]. Chang [Chang 2012] offers a useful review of the constructs in this model. Using this model, Venkatesh, et al. [Venkatesh, et al. 2012] describe 74% of the variance in behavioural intention and 52% of the variance in technology use.

Performance expectancy has been found to be the most significant predictor of behavioural intention to use a technology, reportedly accounting for over 80% of variance in behavioural intention [Venkatesh, et al. 2003, Williams, et al. 2015]. Effort expectancy relates to the perception of the system's ease-of-use and complexity and is most significant in the first exposure to a system, becoming less predictive with sustained usage [Venkatesh, et al. 2003]. Social influence has also shown to be a reliable predictor of behavioural intention, as have Facilitating Conditions [Chang 2012]. Later inclusions to the UTAUT framework are hedonic motivation and habit, which both show indications of affecting both behavioural intention and use behaviour directly [Venkatesh, et al. 2012].

3.3 Validation of metrics

The validation of metrics used in this study was a multistep process aimed at building a model of what drives acceptance and use of technology in the industrial KIT.

On step 1, the initial collection of items resulted in a pool of 91 statements, many of which belonging to groups of items, investigating the same construct, but asking for different aspects or using different phrases to express the same statement using a 7-point Likert agreement scale [Wild, et al. 2017].

On step 2, we tested reliability and measured internal validity of the constructs and items in the model by asking 15 subject matter experts from a project consortium partnership to provide ratings for all items of the pool.

With that, we measured the correlation (Pearson's r) across the responses with the sum scores of all items to assess that each item is actually measuring what we were interested in, testing for discriminatory power of the item [Diekmann 2007]. The general assumption of this is that in total the chance for error is less likely than with a single item [Diekmann 2007].

Results show that there are several items that do not correlate (directly or inversely) with the sum score (Fig. 2). There are 12 items in total correlating with the sum scores on a level higher than 0.7, a total of 36 items on a level higher than 0.6, and 45 items with a correlation value higher than 0.5. If responses to an item did not correlate with the sum scores of all items' responses, then it was very likely to not measure aspects of acceptance and use of technology, but rather something else. This analysis step allowed sorting out those items not highly correlated with the sum score.

Code	Name	Short explanation	References
PE	Performance Expectancy	The degree to which an individual believes that using the system will help him or her to attain gains in job performance.	UTAUT2: [Venkatesh, et al. 2012]; [Chang 2012]
EE	Effort Expectancy	Synonymous with perceived ease of use; the degree to which a person believes that using a particular system would be free of effort.	
SI	Social Influence	The extent to which consumers perceive that important others (e.g., family and friends) believe they should use a particular technology.	
FC	Facilitating Conditions	The degree to which an individual believes that organizational and technical infrastructure exists to support use of the system.	
HM	Hedonic Motivation	The fun or pleasure derived from using a technology.	
HT	Habit	The extent to which people tend to perform behaviours automatically because of learning.	
BI	Behavioural Intention	Referred to below as 'intention'; a person's perceived likelihood or subjective probability that he or she will engage in a given behaviour.	[Fishbein and Ajzen 1975]
UB	Use Behaviour	The manifest, observable response in a given situation with respect to the use of the system.	[Fishbein and Ajzen 1975]

Table 2: Constructs for the UTAUT TAM

On step 3, we calculated the item-to-item correlations to further identify those items loading onto the same construct. If the correlation between two or more items was high, one or a subset was selected. On step 4, Cronbach's α was measured to estimate interrater reliability, comparing the reliability for the full pool as well as the final subset selection. Analysis of the 36 included items and their item groups finds that several items correlate highly within their group and choices was made to re-phrasing with more clarity or for a more aesthetically pleasing formulation.

It is worth noting that items relating to Price Value were not included in this study. The only item in this group demonstrating high enough correlation in this group was PV5: "There are no standard, off-the-shelf AR/WT solutions", which we dropped as it was considered to be unclear without the inclusion of the other items, which related to the pricing, value for money, and cost of customisation. The reason for the lack of correlation in this group may be that AR/WT is an emerging area, so few or no references exist for what constitutes value for money.

3.4 Predictors

The above process led to the selection of 19 items for inclusion in a questionnaire (Table 3).

#	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 organization 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	Interoperability is important for AR & WT.
10	IOP2	I am worried about vendor lock in with AR & WT.
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	Learning curve for AR & WT is too high compared with the value they would offer.
14	PE10	With AR & WT, I immediately know when a task is finished.
15	PE4	Using AR & WT increases my productivity.
16	PE8	AR & WT increase precision of tasks.
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 life.
19	UF1	Please choose your usage frequency of AR/WT

Table 3: Final metric scale, from Wild, et al. (2017)

The resultant questionnaire was used in our Development and validation of TAM for AR/WT study with 33 participants from three industrial partner companies, described in section 2.1, in the areas of aviation, medicine, and space in order to assess the current level of user acceptance of AR and WT in those industries [Wild, et al. 2017]. A split half reliability test and the predictive quality of the items against measured behavioural intent over this pool of responses from end-user participants assess reliability and internal validity of the model. The findings of technology acceptance of AR and WT of the participants measured with this instrument are summarised below. A known limitation of this study is that the group of participants may not be representative for the whole target group in the three industries, which was addressed in the second study "TAM for AR/WT" (see Section 5).

3.5 Discussion on Metrics in 'Development and validation of TAM for AR/WT'

Generally speaking, the participants were acceptant of the tested technology (mean of means: 4.51 with a standard deviation of the means of 1.22). The majority of participants expressed a positive attitude towards the technology (ATU4) and showed signs of hedonic motivation to work with the technology (HM2b), expected low effort in working with the technology (EE2), and were positive about the availability of facilitating conditions and resources required (FC1). They disagreed with AR/WT technology requiring a too-steep learning curve for low value (LRN1).

In general, participants expressed a tendency to agree that the technology demonstrated a high degree of computer self-efficacy (CSE4), came with an associated positive image (IMG4), required interoperability (IOP1), and promises performance gains to be expected of it (PE10, PE8). They were unsure about whether using the technology would provide personal prestige (IMG1) and whether there was any social influence to use this technology (SI1). They had similarly mixed feelings about improvements in their productivity through the use of AR/TW (PE4) and the risks and costs of interoperability (IOP2, IOP3), albeit some expressing agreement. There seemed to be very mixed attitudes with respect to information security and privacy (expressed in IS6). They currently do not possess addictive habits towards AR/WT technology (HT2).

4 Model Specification

In order to operationalise the above metrics, the questions were fit, where applicable, to the UTAUT2 model [Venkatesh, et al. 2012], informing predictors of behavioural intention (BI) and, consequently, use behaviour (UB). Predictors from UTAUT2, namely Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC), Hedonic Motivation (HM), Price Value (PV) and Habit (HT), are complemented with additional constructs of Interoperability (IOP), Learnability (LRN), Image (IMG) and Technology Fit (TF) (Table 4), which have been introduced as measurable elements that characterise technology acceptance, fine-tuned to AR/WT. When allocating the construct items for the AR/WT TAM, consideration was given to the organisational context in which these studies took place and to the nascency of AR-enhanced KIT.

4.1 Additional Constructs for AR/WT

Four additional constructs for AR/WT are presented here (Table 4).

Code	Name	Short explanation	References
IOP	Interoperability	The necessary minimum cooperation between systems allowing for fulfilment of a task which can only be carried out by combining these systems	[Sobernig, et al. 2006]
LRN	Learnability	The degree to which a product or system can be used by specified users to achieve specified goals of learning to use the product or system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use, , or by product properties corresponding to suitability for learning as defined in ISO 9241-110.	ISO/IEC 25010:2011
IMG	Image	Synonymous with 'social approval'; the degree to which use of an innovation is perceived to enhance one's status in one's social system.	[Moore and Benbasat 1991, Rogers 1983]
FT	AR/WT Fit	The extent to which an activity, thought of as a sequence of action steps, is enhanced through the use of an AR/WT system.	Adapted, in this work, from [Parkes 2013]

Table 4: Constructs used in the AR/WT technology acceptance model.

The *Interoperability* of the system is a description of the cooperation between subsystems through application functionality (WEKIT.one) and application semantics, in the form of the ARLEM standard. Any system said to be interoperable contains at least an 'operationalisable' minimum level of cooperation between subsystems, in the sense that operational tasks which require the use of some or all subsystems complete successfully [Sobernig, et al. 2006], thus it is closely tied to performance expectancy or the C-TAM-TPB concept of perceived usefulness.

So, while IOP is necessarily specified at a data level, *Learnability* requires an assessment on a pedagogic level, investigating those properties that make the system conducive to learning. This means that the affordances available to the user (for those applicable here, see Guest et al. 2017) must be made available and accessible within the system's functionality.

The *Image* construct is taken as that part of SI that relates to the organisational perspective exerting influence on an individual to use the technology was shown to be significant, though largely in situations where the use of the technology was mandatory [Venkatesh, et al. 2003], which is not the case here.

Motivated by the successes of Task-Technology Fit (TTF) models [Goodhue and Thompson 1995, Parkes 2013], a final construct - entitled *AR/WT Fit* - is proposed to describe how well suited an AR/WT system is to a contextual activity, that is, to a sequence of action steps anchored to elements in the workplace. Rooted in TTF, which posits a technology is definable by task and technology characteristics, this construct supposes the *AR/WT Fit* (TF) is composed of an individual-technology fit (ITeF) and an activity-technology fit (ATeF) as in Figure 3.

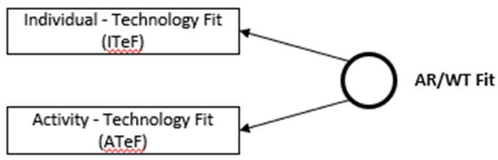

Figure 3: AR/WT Fit construct, adapted from [Parkes 2013]

ITeF is the extent to which the system offering decision support enhances the individual's competence when completing an activity (sequence of action steps). It is investigated with item ATU4: "I look forward to those aspects of my job that require me to use AR/WT". The attitudinal nature of this statement reinforces Parkes' statement that "novices are likely to perceive a technology providing additional information cues as useful, despite the fact that they are unlikely to be able to use that additional information to improve their performance, consequently it is argued that ITeF will affect attitude rather than performance" [p. 998, Parkes 2013]. While is it thought that this stage of learning is both short-lived and unfavourable to the learner [Cooper, et al. 2007], it is itself a fitting statement, since all of the participants involved in the studies in question were experiencing this AR/WT-enhanced training for the first time.

ATeF is defined as the extent to which the level of guidance provided by the system matches the complexity of completing a set of action steps in a given workplace. It is coded as item CSE4: "I could complete a job, if I had used similar technologies before this one to do the same job". A sense of parsimonious suitability, or apt level of guidance, is implied and refers to a specific, though non-specified activity.

While grounded in Parkers' definition of a task-technology fit as "the extent to which the complexity of the task being undertaken matches the decisional guidance provided by the technology" [Parkes 2013], this paper moves from thinking of a 'task' to an 'activity', the latter being more generalised and allows for a degree of cross-context functionality in the technology (across similar, and definable action steps).

4.2 Hypothesis Formulation

This research proposes that UTAUT2 provides a useful, though limited model for assessing the technology acceptance of AR and WT and that additional constructs can be added that are related to these fields in order to better predict acceptance. Based on this, the following subsequent hypotheses are put forward as causal relations (Table 5), testable through an expanded structural equation model (Figure 2).

H1	Interoperability is significant in predicting acceptance of AR and WT
H2	Learnability is significant in predicting acceptance of AR and WT
H3	Image is significant in predicting acceptance of AR and WT
H4	AR/WT Fit is significant in predicting acceptance of AR and WT
H5	Interoperability is positively related to Performance Expectancy
H6	Interoperability is positively related to Effort Expectancy
H7	Learnability is positively related to Effort Expectancy
H8	Image is positively related to Social Influence
H9	AR/WT Fit is positively related to Facilitating Conditions

Table 5: Hypotheses to be tested

These hypotheses can also be represented as a structural equation model (shown in Figure 4):

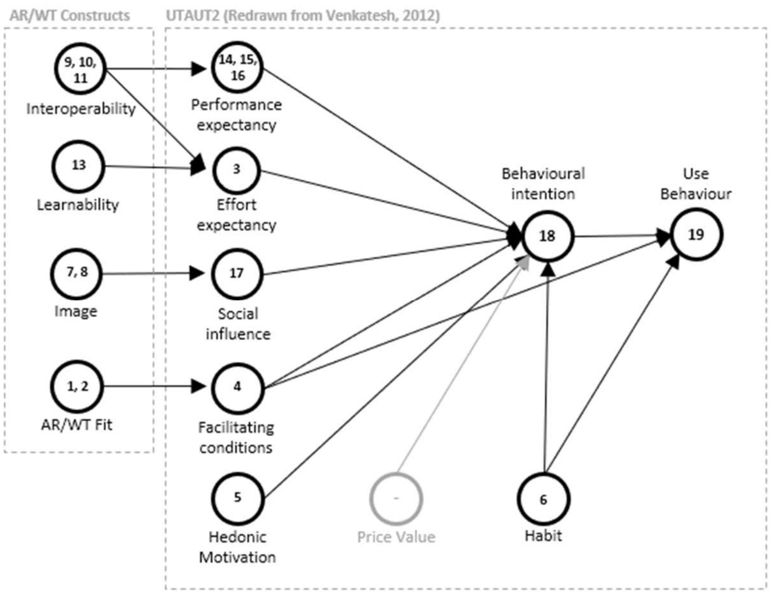

Figure 4: Research Model (showing item allocation)

5 Model Improvement

Model fitting was done using R Studio (version 3.4.2). Correlations and confirmatory factor analysis used the most up-to-date (December, 2017) 'psych' and 'lavaan' libraries and path diagrams were made using the 'semPlot' package. A maximum likelihood estimator was used for fitting, with standard errors and test statistics scaled according to the sample size (as per the Satorra-Bentler correction).

5.1 Preliminary Model Fitting

The model in Figure 2 was coded and included residual correlations between each and all of the AR/WT construct (not shown). This is based on the assumption that there are contextual factors influencing these measures not captured during this study. The model converged after 84 iterations - its fit is measured by indices shown in Table 2 – and 129 complete entries (of 142 collected) were used in the assessment of the model.

When assessing these correlations, the Spearman rank was considered most suitable as it avoids the assumption made in the Pearson model - namely, normally distributed, linearly correlated variables (see Figure 5). The authors consider there to be insufficient existing research in the field of AR and WT to assume the presence (or absence) of linearity.

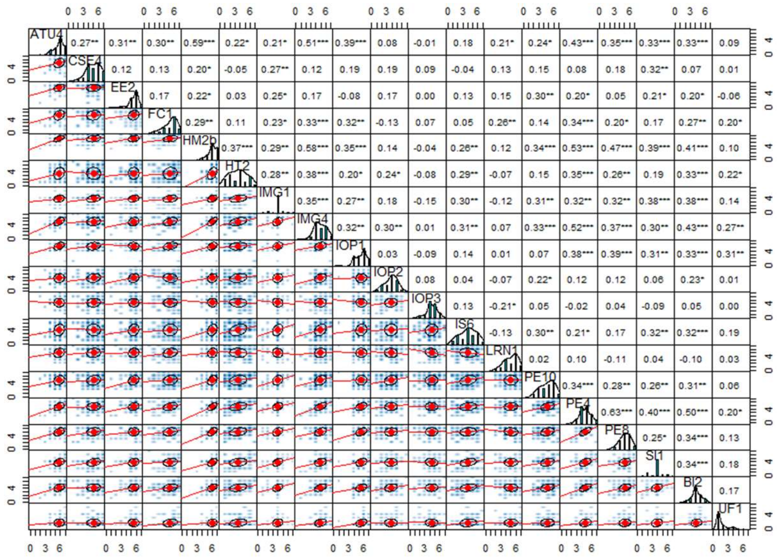

Figure 5: Spearman Rank Correlation of Observed Variables

This work will report both absolute and relative fit indices (FIs). The first group consists of the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the McDonald fit index (MFI). The second consists of the Chi-squared p-value (Chi-sq.), the relative non-centrality index (RNI), the root mean square error of approximation (RMSEA) and the standardised root mean square (SRMR). Table 6 shows the results for these metric, as well as those from (adjusted) goodness-of-fit indices ([A]GFI), though these are not considered applicable in this context.

Type of Index	Absolute					Relative			
(N=129)	**GFI**[1]	**AGFI**[1]	**CFI**[3]	**TLI**[2,3]	**MFI**	**Chi-sq.**	**RNI**	**RMSEA**[3]	**SRMR**
Value	0.988	0.980	0.883	0.846	0.768	0.000	0.871	0.059	0.067
'Good fit'[4] cutoff	-	-	> 0.95	> 0.93[2]	> 0.90	> 0.05	> 0.95	< 0.06	< 0.08

Table 6: Fit indices and cutoff values used to assess goodness of fit.

5.2 Model Optimisation

Values for all fit indices, apart from the RMSEA and SRMR, suggest that the model can be significantly improved. Acceptable values for these indices of 0.059 and 0.067, respectively, may well indicate the stability of the underlying UTAUT2 structure, due to this index's sensitivity to misspecified components in the model [Hu and Bentler 1998].

What follows is a 'number of factors' analysis as well as a process guided by modification indices (MIs) and review of the observed variables (Figure 6). Analysis to determine an appropriate number of factors for a model works by perturbing one value in a model and noticing the impact on correlated factors, and then performing this for arbitrary number of factors associated to the correlations within a dataset. MIs indicate the degree to which freeing a particular factor in the model can impact the chi-square value.

The results from these instruments assist in the dissemination of model characteristics and, consequently, offer avenues for change that can be tested for their impact in determining predictive success. The MIs are taken to be informative, rather than normative, acting as sign-posts when modifying either the data frame, through the exclusion or aggregation of items, or the connections to the UTAUT2 model. An assessment of item complexity[5], together with the principle of parsimony [Preacher 2006] informs the decisions on how to improve the theoretical model.

The Bayesian Information Criterion (BIC) places a greater significance on model parsimony and is shown in Figure 6. By simulating a model with a varied number of

---

[1] A general consensus is to avoid using GFI and AGFI, as they can be skewed by sample size. For samples of this size (N=129), Sharma, et al. [2005] report a high degree of misspecified model acceptance using GFI and recommend not using it.

[2] The Tucker-Lewis index (also known as the non-normed fit index) tends to over-reject models with samples sizes of less than 250 when using a cutoff of 0.95 so the lower value of 0.93 was used [Yu 2002].

[3] TFI, CFI and RMSEA can suffer from bias related to non-normality in samples smaller than 300. The 'sample-corrected robust' values of [Brosseau-Liard and Savalei 2014] and [Brosseau-Liard, et al. 2012] are reported here.

[4] Apart from the index in footnote 2, the cutoffs for acceptable model fit are taken from Hu and Bentler [Hu and Bentler 1999].

[5] A note on complexity: Fit indices have inherently varying strengths of dependence on changes to complexity - defined as either the stability of the chi-squared value with respect to the model fit, or a ratio of the chi-squared value to the degrees of freedom in the model. The CFI shows a somewhat negative bias for increasing complexity, while the SRMR has none.

factors and comparing their residual matrices, the Very Simple Structure (VSS) gives visibility to those (only) the largest factors loading on each variable.

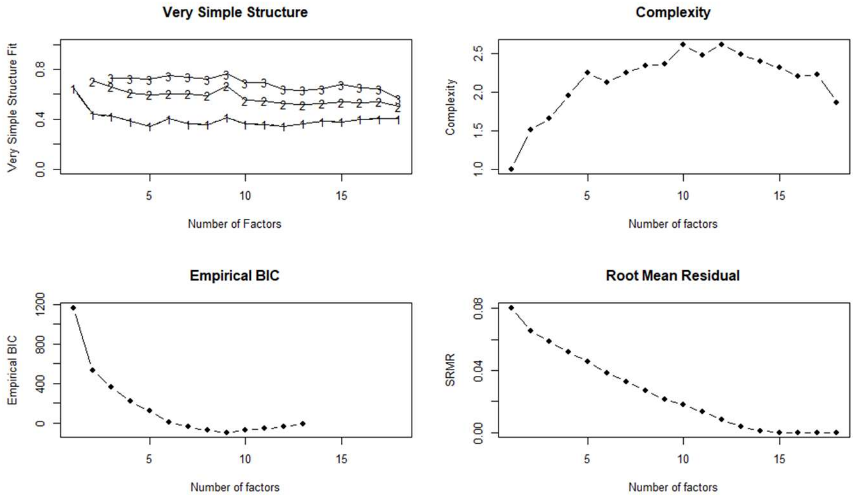

Figure 6: 'Number of Factors' Results

We can see from the above figure that the optimal number of factors, when giving importance to parsimony via the BIC index, is nine. High points in the complexity of tested cases fall at ten and twelve factors and the VSS suggests a drop off in factor loading when the number of factors exceeds nine. These results suggests that a reduction in factors (from the current number of twelve) may benefit the model's representation of reality.

Using modification indices we may highlight potential model flaws affecting model cross-loadings - the links between endogenous and exogenous variables.

A lower bound for reporting each modification index (mi) was taken to be seven and the results were put in descending order. Items under review are only those additional to the UTAUT2 model, which is thought to provide a firm foundation for this study, so only the additional constructs (described in section 4.1) are under scrutiny.

LHS	op.	RHS	MI
BI	=~	PE4	25.545
LRN	=~	PE4	9.774
LRN	=~	IOP3	8.994
LRN	=~	PE8	7.166

Table 7: Initial modification indices reported from the convergence.

The learnability construct is seen here three times and, in one instance, is linked to one of interoperability's exogenous predictors (see Table 7, shaded cells). LRN1 - "Learning curve for AR & WT is too high compared with the value they would offer" and IOP3 - “Integration costs of AR & WT with other software/ systems in use are high” are relatively complex formulations of question. It is possible that the statement from LRN1, in requiring a comparison of ease of use (learning curve) and usefulness (value offered), was hard to interpret, especially for a newcomer to the technology. Participants may have had similarly grounded confusion when responding to statement IOP3; where a limited offering of such technologies exists, integration costs are largely unknown.

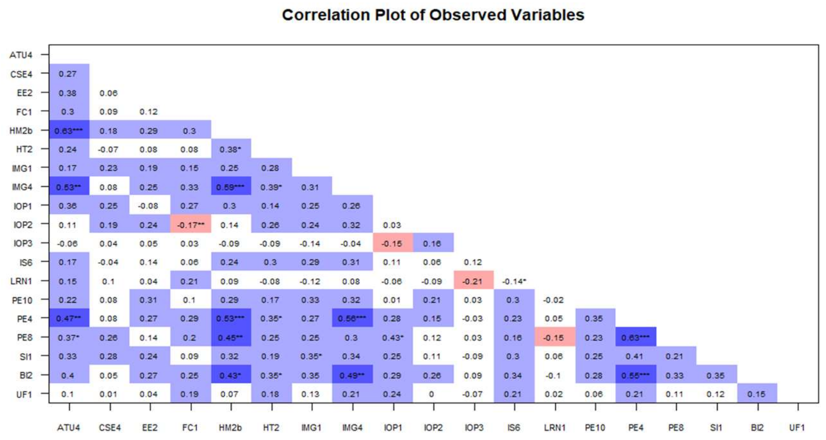

Figure 7: Correlation Plot of Observed Variables

A correlation plot of all the exogenous variables also reveals these two items as having very little in common with other items in the dataset (see Figure 7).

Excluding LRN1 and IOP3 reduces the number of factors by one, due to LRN having a single factor loading onto it. This change has a significant effect on all indices, notably reducing RMSEA to below its cutoff and drives the chi-square p-value of 0.008 above the critical threshold for significance. In 169 iterations, the modified model (N=130)[6] converges to one with the modification indices shown in Table 8.

LHS	op.	RHS	MI
BI	=~	PE4	26.06505
TF	=~	IOP2	9.233064
PE	=~	HT2	7.304016

Table 8: Modification indices after convergence in step 2.

Here, there is the inclusion of IOP2: “I am worried about vendor lock in with AR & WT”. In contrast to the previous elements, this is a more straightforward statement formulation and, assuming the participant is familiar with the concept of ‘vendor-

[6] The exclusion of the LRN1 and/or IOP3 items led to the possible inclusion of an additional participant’s responses.

lock', there is little room for misinterpretation. The strength of the covariance observed suggests that the appropriateness of the technology (for both the task and the individual) is somehow related to vendor-lock. Given that this is both a plausible link and that there are other conceivable relations between AR/WT Fit and Interoperability, a latent correlation is added to the model between these constructs.

Attention was then given to the IMG construct. Two factors load onto this item: IMG1: "People in my organization who use AR & WT have more prestige than those who do not" and IMG4: "I use AR & WT solutions, because I want to be a forerunner in technology exploitation". IMG1 implies that others in an organisation may have more importance as a corollary of their use of such technology, an idea close to that driving the Social Influence construct (see Table 2), though seen here in an organisational context. For this reason, it is suggested that this item be loaded directly onto SI.

IMG4 is a more complicated matter, due to the aforementioned novelty of technology-enhanced learning support systems. People may already feel as though they are a forerunner in technology exploitation, as indeed they are by virtue of their participation in this study. It is further argued here that there is no 'important other' in this item's wording, hence the perceived external pressure to use the technology cannot be well understood from it. Though this question may relate more to hedonic motivation, or perhaps performance expectancy, it is hence removed in order to create a more parsimonious model by allowing the IMG variable to be subsumed into SI, further reducing the number of factors to 10.

After this change, the model converges in 136 iterations, the CFI rises 0.954, while the TLI and MFI show increased in fit estimation to 0.937 and 0.908, respectively, putting all of the absolute fit indices above their cutoff value. RMSEA also drops further, to 0.033, and SRMR remains unaffected at 0.062. No further investigation of MIs is performed as all items additional to UTAUT2 have been considered (or discussed in section 4.1.4).

We may now consider how the constructs of Interoperability and Fit may relate to the UTAUT2 model. At present, IOP is thought to be predictive of PE (H5) and EE (H6), and TF of FC (H9). Considering TF and the description of ATeF (section 4.1) it is suggested that item CSE4, "I could complete a job, if I had used similar technologies before this one to do the same job" relates to PE, as the performance of said job may be implicitly assumed to have changed. Hence a new supposition (H10), that *AR/WT Fit is positively related to Performance Expectancy*, is included. These changes leads to a convergence (in 181 iterations) with a fit quality estimation shown in the final row of the table below (Table 9).

#	Change Made	CFI	TLI	MFI	RNI	RMSEA	SRMR	Chi-Sq.	Chi-sq. p-value	Degrees of Freedom
1	n/a	0.857	0.812	0.735	0.857	0.069	0.074	198.250	0.001	130
2	--IOP3 --LRN1	0.916	0.888	0.830	0.905	0.055	0.063	150.436	0.008	102
3	IOP~~TF IMG1=>SI --IMG4	0.954	0.937	0.908	0.941	0.040	0.062	113.164	0.101	88
4	PE ~ TF	0.969	0.957	0.928	0.954	0.033	0.062	106.384	0.187	87

Table 9: Summary of Model Optimisation Steps

5.3 Final Model

Modifications were made to each of the AR/WT constructs during model improvement. Interoperability loading was reduced to two items (removing IOP3), the Learnability construct (LRN1) was removed altogether, and predictors for Image were either integrated into Social Influence (IMG1) or removed (IMG4). AR/WT Fit maintained its factor loading (CSE4 and ATU4) but is also thought to be positively related to Performance Expectancy (TF ~ PE). The final result is shown in Figure 8.

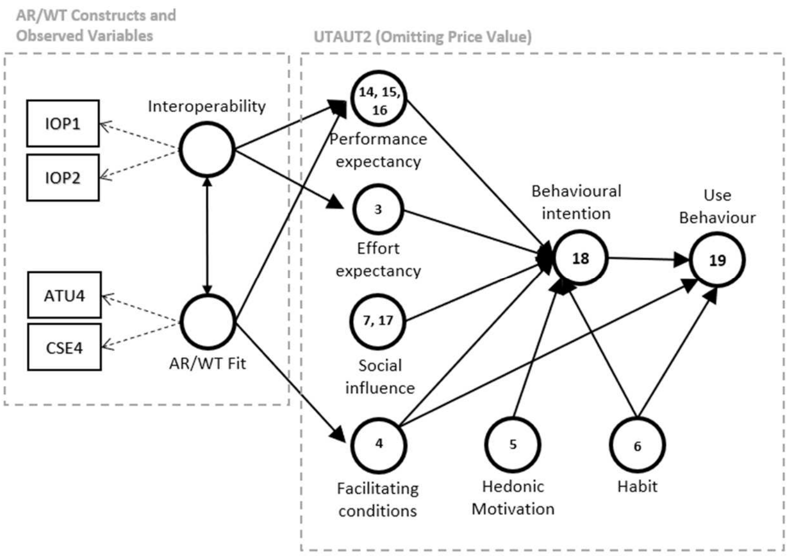

Figure 8: The Final SEM Model

This model (N=130) has a chi square p-value of 0.187 and the model has 87 degrees of freedom (giving a size-parameter ratio of 1.38) and a minimum function test statistic of 98.538. The Satorra-Bentler correction, used to understand how much the indices have been adjusted when accounting for inherent non-normality in a sample of this size, is 1.080.

5.4 Factor Loading and Residuals

The final, simplified, model is described by the factor loading and regression estimates given in tables 10 and 11, respectively. Table 12 shows the residuals in the model, which need to have a modulus of less than 0.1 to be considered significant (highlighted are those that fall outside this boundary). The RMSEA index is largely based on these values, so this gives a picture of the underlying structure leading to this estimate of fit.

Latent Factor	Indicator	Estimated Value	Standard Error	Z = Est./SE	Std. Regression Coefficient	Significance [7]
IOP	IOP2	0.518	0.252	2.054	0.118	*
TF	CSE4	0.569	0.186	3.059	0.274	**
PE	PE8	0.722	0.106	6.840	0.656	***
PE	PE10	0.473	0.122	3.869	0.373	***
SI	IMG1	0.774	0.213	3.629	0.504	***

Table 10: Significant Factor Loading

Predicted Var.	Indicative Var.	Estimated Value	Standard Error	Z = Est./SE	Std. Regression Coefficient	Significance [7]
PE	TF	0.855	0.222	3.855	0.547	***
EE	IOP	0.806	0.275	2.934	0.206	**
FC	FT	0.853	0.224	3.812	0.403	***
BI	PE	0.300	0.144	2.086	0.269	*
BI	SI	0.705	0.293	2.409	0.413	*
UB	HT	0.111	0.057	1.962	0.165	*

Table 11: Significant Regression Estimates

[7] *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$

	IOP1	IOP2	ATU4	CSE4	PE4	PE8	PE10	SI1	IS6	IMG1	EE2	HM2b	FC1	HT2	BI2	UF1
IOP1																
IOP2	-0.011															
ATU4	-0.009	-0.086														
CSE4	0.139	0.059	0.060													
PE4	-0.023	0.024	-0.017	-0.116												
PE8	0.140	0.013	-0.009	0.083	0.003											
PE10	-0.111	0.152	-0.040	0.043	-0.018	0.013										
SI1	-0.011	0.009	0.014	0.149	0.054	-0.050	0.047									
IS6	-0.096	0.037	-0.045	-0.056	0.001	0.008	0.238	-0.011								
IMG1	-0.071	0.196	-0.053	0.053	-0.025	0.044	0.206	0.018	0.013							
EE2	-0.113	0.192	0.009	-0.019	0.016	-0.019	0.163	0.049	-0.051	-0.040						
HM2b	-0.002	0.024	0.022	-0.042	-0.011	0.007	0.031	0.027	-0.018	-0.021	-0.005					
FC1	0.030	-0.136	0.036	0.006	0.040	-0.046	-0.017	0.024	-0.013	0.008	0.028	-0.013				
HT2	-0.060	0.193	-0.030	-0.055	0.047	0.049	0.072	-0.068	0.088	0.017	-0.090	0.000	-0.014			
BI2	-0.043	0.122	-0.053	-0.109	0.031	-0.044	0.022	-0.013	0.037	-0.022	-0.009	-0.003	0.006	0.007		
UF1	0.067	0.007	-0.007	0.021	0.106	0.104	0.022	0.032	0.109	0.017	-0.067	-0.065	-0.002	-0.001	0.002	

Table 12: Residuals

That more than 90% of the values in the above table are within the -0.1→ 0.1 band suggests that at least a somewhat complete picture of the data has been generated by this model. This is, of course, reflected in the final RMSEA value of 0.033, which has a 90% confidence interval of 0 - 0.062.

6 Cross Validation of Findings

The method used to determine a generalisable estimate for the structural equation model was a Monte-Carlo (repeated random subsample) validation, comprising two phases of investigation; random generation of variable size subsets and high-fold approximation. The purpose of this process is to minimise errors from overfitting and, while having a relatively small dataset makes the need for such treatment all the more necessary, the same attribute may hamper a thorough investigation, due to any split in the sample population causing noticeable degradation in the data's speculative resolution. This hindrance is, however, thought to be measurable and is tested in phase one.

6.1 Generation of Subsets

The first phase shows how the deterioration in fit corresponds with the sample size used, allowing a cutoff to be imposed (based on fit indices in Table 4), designating a minimum sample size to reliably estimate fit. The sizes to test can be simplified to a single range by using a training-test or test-training ratio (R_T) depending on the subsample. Here, a range of 1.0 - 2.5 was used, corresponding to 33-93 elements in the training set, with the remainder in the test set (from a total of 130). Changes in all fit FIs are reported, including '1-RMSEA' and '1-SRMR', in order to place the metrics on the same scale, considering that the cutoff values are similarly inverted ('1-RMSEA > 0.94'; '1-SRMR > 0.92').

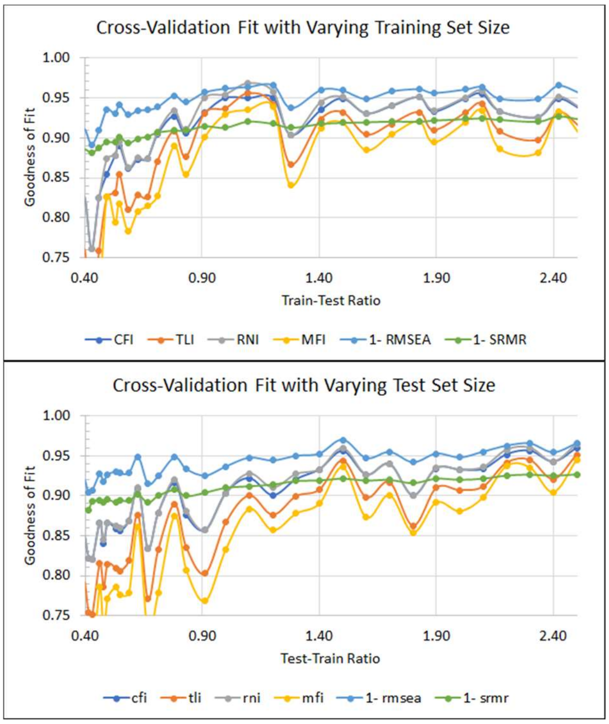

Figure 9: Changes to Fit Indices with Varying Sampling Size

More stability is seen in the FIs RMSEA and SRMR, compared to CFI, TLI, RNI and MFI. For the training set, a majority of the indices (4 of 6) are above the cutoff when N>76, both RMSEA and SRMR are indicative of a well fitted model (Table 9). In the test subsamples, the latter condition is only satisfied also when N>76. Furthermore, there is a visible trend in Figure 9 for the FIs to converge more after a test-train ratio of around 1. This may seem unsurprising, but could indicate that the dataset size is around, or slightly smaller than, the number necessary to allow for two independent populations to be fit to the data.

The fact that a case does not exist where both the RMSEA and SRMR are satisfied for both training and test subsets, together with the upward convergent trend in the other fit indices, suggests that a limited dataset size may be preventing further insight into the model. However, given the investigatory nature of this study, and the nascence of its field, it is likely that some information can be gleaned from data that is said to have moderate predictive power. By lowering the cutoff of all FIs incrementally (in steps of 0.001 to a reduction maximum of 0.020) and looking for a ratio that supports a balanced case, minimising errors to both sets caused by the small population size. When relaxing the cutoffs by 0.006, the case of N_{TRN}=65 emerges as a testable compromise for this study. In addition to the identification of this item, which will be used knowing it passes some error to the next phase, means and standard deviations are calculated for the entire set of subset size variations:

		CFI	TLI	RNI	MFI	RMSEA	SRMR
Training Set	MEAN	0.90	0.87	0.90	0.84	0.05	0.09
	STD DEV	0.04	0.06	0.03	0.06	0.02	0.01
Test Set	MEAN	0.90	0.88	0.90	0.84	0.05	0.09
	STD DEV	0.05	0.08	0.05	0.08	0.02	0.01

Table 13: Summary of Monte Carlo Cross-validation

6.2 High-fold Approximation

Since this approach relies on random sampling, it will contain Monte Carlo variation based on the potential for omission and duplication of entries. In order to more resemble an exhaustive method for validation, the data may be based on a larger number of folds, or iterations, in which the random division is made between the dataset's two groups. The above information and summary is based on a 10-fold splitting of the dataset, which is extended here to 1000 iterations (Figure 10).

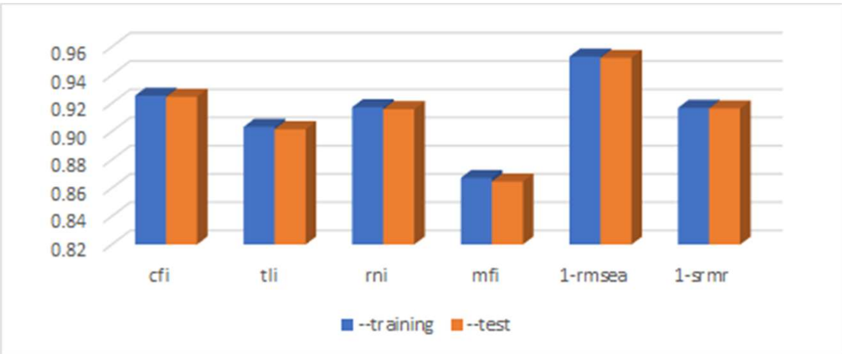

Figure 10: High-fold Cross-validation Results

Mean Values	CFI	TLI	RNI	MFI	1-RMSEA	1-SRMR	chi-sq.	p-value
Training Set	0.925	0.903	0.917	0.867	0.953	0.917	108.684	0.134
Test Set	0.925	0.902	0.916	0.865	0.952	0.916	106.834	0.151

Table 14: High-fold Mean Values for Fit Indices

7 Results Summary

The results above indicate that the dataset collected, together with the proposed structural equation model, has moderate predictive power when investigation the acceptance of AR and WT. Of the ten hypotheses posited in this work, six are testable while four are rendered null by the model simplification process (H2, H3, H7 and H8). Of the remaining six, H1 and H4 propose an overall correlation between the constructs of IOP and TF to technological acceptance and H5, H6, H9 and H10 (coined in section 5.2) refer to the predictive power of each latent variable. Based on the findings, each hypothesis can be tested, with the following outcomes as shown in Table 15.

In short, interoperability is a moderately good predictor of effort (but not performance) expectancy, while AR/WT Fit offers stronger indications that it may predict both the performance expectancy and facilitating conditions of UTAUT2. These claims are tempered by the acknowledgement that this model shows only a moderate power in representing reality. The two constructs added are also seen to be interdependent and delineating these, as well as investigating their internal structure, will be the subject of future work. Of the constructs in UTAUT2, only PE and SI are shown to have a significant bearing on BI; both of these latent variables show high factor loadings from their respective indicators, suggesting they are the main predictors of BI in contexts and situations such as the ones tested.

8 Discussion and Future Work

The results confirm work done by Venkatesh [Venkatesh 2012], showing that the UTAUT2 model seems to provide a good basis for exploration in to technology acceptance. The study also addresses limitations outlined by Williams [Williams et al. 2015], as it uses a decent size sample set, implements the technology in three different situations, and collects data from the workplace itself, rather than in a controlled setting.

#		Hypothesis	Outcome	Reason
H1		IOP is significant in predicting acceptance of AR and WT	Rejected	IOP, although showing a strong latent correlation to TF 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	Constructs removed in optimisation.
H3		IMG is significant in predicting acceptance of AR and WT	Null	Constructs removed in optimisation.
H4		TF is significant in predicting acceptance of AR and WT	Accepted	TF 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	Constructs removed in optimisation.
H8		IMG is positively related to SI	Null	Constructs removed in optimisation.
H9		TF is positively related to FC	Accepted	The model claims to predict 85% of the covariance of these items.
H10		TF is positively related to PE	Accepted	The model claims to predict 86% of the covariance of these items.

Table 15: Results of Hypothesis Testing

The notion of AR/WT Fit is seen as indispensable, in part due it's allusion to the physical comfort of worn devices. A head-mounted display, armband or other garment endowed with sensor hardware must not impede the user, either through limitations placed on their required movement during an activity, or by causing discomfort to such a degree that their attention is drawn from the task at hand. The technology seeks to minimise any such disruption that affects their competence (though some is seen as inevitable given the current technological level) and maximise the number of additional affordances, or opportunities for action, that were not previously there. For novices, a measure of affective positivity (one might say enjoyment) is shown to be indicative of a good fit between the user and the technology (in line with findings from Parkes [Parkes 2013]). More broadly, the appropriateness of the level of complexity of the augmented space offered to the user is confirmed as important when matching a wearable prototype to a particular activity, though both of these are in need of further investigation.

Taking this further, measuring an individual's competence prior to using the technology would be extremely interesting as aptitudes for a particular task, even one never before performed, are unlikely to be the same for everyone. Should there be a demonstration of increased competence shown in general, this would provide a strong basis on which to further isolate this factor, leading to a more meaningful inclusion into technology acceptance models. Studies looking at varying complexities of AR environments and their impact on cognitive load, working memory utility and other measures of (neuro-) physiological effort, would be of particular use when understanding the fit between the individual and the technology offered to them. Indeed, further iterations of the prototype described here will include biometric indicators, such as heart-rate variability and galvanic skin response.

When considering the activity-technology fit, we can consider this a departure from a traditional view of what constitutes a technology, since there is the necessary confluence of software and hardware culminating in a qualitatively different experience – that of an augmented, highly interactive workplace. This integration between programmed design and its pervasiveness (the extent to which it blends into the surrounding environment) makes the task of determining complexity far from trivial and offers many avenues for exploration and research.

Interoperability also demonstrated significance within this model, particularly in relation to the effort expectancy when using the prototype. While this data may have been somewhat skewed by a lack of additional devices that were in constant communication with the holographic display, this is likely to indicate that, by adding more wearable elements, there is a risk of making the system too difficult to use.

In the industries observed, it is clear that specific attention has to be paid to equipping users with the resources needed, in particular with the devices. Device management will be an issue, also for managing update and upgrade procedures. Integration with legacy systems is of big concern, so in the introduction of AR/WT, interfaces have to be created ensuring that any new solution fits seamlessly in into the existing pool of hard- and software. It is also important to ensure solutions meet expectations of performance gains. Hopes are high that an increase in productivity, precision and live feedback will be delivered by solutions that are sufficiently developed to make easy to use, while cater to a personal style of technology use. If these expectations are not fulfilled, acceptance will likely suffer.

The planned future iterations of this work will aim to offer less complex formulations of questions investigating Learnability, develop metrics for understand how well specific devices are incorporated in to a prototype and add additional items to capture the internal constructions found within AR/WT Fit. **These** extensions, together with a larger sample size, should lead to a stronger and more robust assessment of technology acceptance.

Acknowledgement

Valuable feedback on the structural equation model was provided by Dr. Thomas Rusch from the Wirtschaftsuniversität Wien, to whom we extend our gratitude. This work was supported by the European Commission under the Horizon 2020 Programme, as part of WEKIT (grant agreement no. 687669).

References

[Brosseau-Liard and Savalei (2014)] Brosseau-Liard. P.E., Savalei, V.: "Adjusting Incremental Fit Indices for Nonnormality";Multivariate Behavioral Research 49, 5 (2014), 460–470.

[Brosseau-Liard, et al. (2012)] Brosseau-Liard, P.E., Savalei, V., Li, L.: "An Investigation of the Sample Performance of Two Nonnormality Corrections for RMSEA";Multivariate Behavioral Research 47, 6 (2012), 904–930.

[Chang (2012)] Chang, A.: "UTAUT and UTAUT 2: A Review and Agenda for Future Research";The Winners 13, 2 (2012), 106–114.

[Cooper et al. (2007)] Cooper, A., Reimann, R., Cronin, D.: "About face 3: the essentials of interaction design"; John Wiley & Sons, Inc., New York (2007).

[Diekmann (2007)] Diekmann, A.: "Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen"; Rowohlt-Taschenbuch-Verl., Reinbek bei Hamburg (2007).

[Emspak (2017)] Emspak, J.: "Are Face Scans Resurrecting the False Science of Physiognomy?"; Atlantic, Washington, DC (2017).

[Fishbein and Ajzen (1975)] Fishbein, M., Ajzen, I.: "Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research"; Addison-Wesley, Reading, MA (1975).

[Goodhue and Thompson (1995)] Goodhue, D.L., Thompson, R.L.: "Task-Technology Fit and Individual Performance";Management Information Systems Quarterly 19, 2 (1995), 213–236.

[Guest et al. (2017)] Guest, W., Wild, F., Vovk, A., Fominykh, M., Limbu, B., Klemke, R., Sharma, P., Karjalainen, J., Smith, C., Rasool, J., Aswat, S., Helin, K., Di Mitri D., Schneider, J.: "Affordances for Capturing and Re-enacting Expert Performance with Wearables"; Springer International Publishing, Cham (2017).

[Hu and Bentler (1998)] Hu, L.-T., Bentler, P. M.: "Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification"; Psychological Methods 3, 4 (1998), 424–453.

[Hu and Bentler (1999)] Hu, L. T., Bentler, P. M.: "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives";Structural Equation Modeling: A Multidisciplinary Journal 6, 1 (1999), 1–55.

[Moore and Benbasat (1991)] Moore, G. C., Benbasat, I.: "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation";Information Systems Research 2, 3 (1991), 192–222.

[Parkes (2013)] Parkes, A.: "The effect of task–individual–technology fit on user attitude and performance: An experimental investigation";Decision Support Systems 54, 2 (2013), 997–1009.

[Preacher (2006)] Preacher, K. J.: "Quantifying Parsimony in Structural Equation Modeling";Multivariate Behavioral Research 41, 3 (2006), 227–259.

[Rogers (1983)] Rogers, E. M.: "Diffusion of Innovations"; The Free Press, New York (1983).

[Schnabl and Zenker (2013)] Schnabl, E., Zenker, A.: "Statistical Classification of Knowledge-Intensive Business Services (KIBS) with NACE Rev. 2"; Fraunhofer ISI, 1264 Karlsruhe, Germany (2013).

[Sharma, et al. (2005)] Sharma, S., Mukherjee, S., Kumar, A., Dillon, W. R.: "A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models"; Journal of Business Research 58, 7 (2005), 935–943.

[Sobernig, et al. (2006)] Sobernig, S., Danielewska-Tulecka, A., Wild, F., Kusiak, J.: "Interoperability and Patterns in Technology-Enhanced Learning"; Polish Information Processing Society (PTI), Szczyrk, Poland (2006).

[Venkatesh, et al. (2003)] Venkatesh, V., Morris, M. G., Davis, G. B., Davis, F. D.: "User Acceptance of Information Technology: Toward a Unified View";Management Information Systems Quarterly 27, 3 (2003), 425–478.

[Venkatesh, et al. (2012)] Venkatesh, V., Thong J. Y. L., Xu, X.: "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology";MIS Q. 36, 1 (2012), 157–178.

[Wild, et al. (2017)] Wild, F., Klemke, R., Lefrere, P., Fominykh, M., Kuula, T.: "Technology Acceptance of Augmented Reality and Wearable Technologies"; Proc. Immersive Learning Research Network: Third International Conference (iLRN 2017) (2017), Springer International Publishing, Coimbra, Portugal, 129–141.

[Williams, et al. (2015)] Williams, M. D., Rana, N. P., Dwivedi, Y. K.: "The unified theory of acceptance and use of technology (UTAUT): a literature review"; Journal of Enterprise Information Management 28, 3 (2015), 443–488.

[Yu (2002)] Yu, C. Y.: "Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes"; Los Angeles: University of California, Los Angeles (2002).