The Impact of Trust, Usability and Satisfaction on the Intent to Use Digital Epidemiology

Benjamin Shanahan

MSc in User Experience Design

Institute of Art, Design and Technology

Dún Laoghaire

Supervisor

Andrew Errity

2021

DECLARATION STATEMENT

I hereby certify that the material, which I now submit for assessment on the programme of study leading to the award of Master of Science, is entirely my own work and has not been taken from the work of others except to the extent of such work which has been cited and acknowledged within the text of my own work. No portion of the work contained in this research project has been submitted in support of an application for another degree or qualification to this or any other institute.

Benjamin Shanahan

3rd May 2021

The Impact of Trust, Usability and Satisfaction on the Intent to Use Digital Epidemiology

Trust in Digital Epidemiology

BENJAMIN SHANAHAN

Institute of Art, Design & Technology, Dublin, Ireland, bnshanahan@gmail.com

ANDREW ERRITY

Institute of Art, Design & Technology, Dublin, Ireland, andrew.errity@iadt.ie

ABSTRACT

Following the SARS-CoV-2 outbreak, digital epidemiology is being used worldwide for the first time in the form of contact tracing apps. Many countries have developed their own with varying levels of success. However, one of the factors that is consistent across every contact tracing app is a level of public distrust towards using the apps and a concern regarding the misuse of private data. In this study 7 Trust Focused Design Principles are developed for use in Difficult to Trust scenarios, aiming to increase the trust, perceived use, satisfaction, and intent of users. A Difficult to Trust scenario is when a user's interaction with a product has been prefixed by a sense of distrust or threat leading users to being distrustful or suspicious about the product before using it, such is the case for many users' first interaction with contact tracing apps. These principles aim to proactively pursue the trust of users, rather than passively gain it over time as many current trust models attempt to do. The results of this user-centred study which included research, ideation, and testing, showed that the implemented principles significantly increased the participants' trust, perceived usability and intent while insignificantly increasing their satisfaction. Although further testing of these principles is required, the results of this study show that they are a promising method for developing in Difficult to Trust scenarios.

1. INTRODUCTION

During the SARS-CoV-2 pandemic many countries have turned to digital epidemiology, mostly in the form of contact tracing apps, in an attempt to trace, test and treat infected individuals attempting to minimize spread while avoiding complete societal lockdown as much as possible.

Opportunistic Networking has become commonplace for the mapping of individuals' movements and their interactions with others through smart phone tracing applications (Ferretti, 2020). The technical capability of Opportunistic Networking was proved early in the pandemic with Korea's 'Self Check' app (Ministry of Public Administration and Security, 2020, March 5) contributing to the country flattening the curve of cumulative cases within a month of first testing (The Government of South Korea, 2020).

However, for South Korea and many other countries, controversy came along with the success of contact tracing as the required data for the app to function included passport information, travel history, phone number and signature as part of the onboarding process (Ministry of Public Administration, 2020, March 3). Privacy concerns surrounding South Korea's tracking app were also heightened due to the fact that the data used would in normal circumstances be deemed illegal if it were not for the overwriting authority of the Contagious Disease Prevention and Control Act (CDPCA) (Ko et al, 2020).

With the introduction of large-scale contact tracing being rife with controversy and privacy concerns, the new technology became difficult to trust for users, before they had even installed as an app. Because of this, trust becomes a major factor in convincing the population to use such an app. In a survey of the Irish population, although 82% of participants said they would be willing to download a contact tracing app, 59% expressed some form of reservation to downloading an app, the most common being *"I worry technology companies will use this as an excuse for greater surveillance after the pandemic".* (O'Callaghan et al., 2020, p. 4). A similar consensus was found within the UK and the USA populations where, in the UK, the most common reason given for not installing a tracker app was *"an increased risk of government surveillance after the pandemic"* (Abeler, 2020, p. 7), while in the USA, in a Pew Research Centre survey (2020), where out of over 10,000 US citizens 50% stated that they were not at all or not too comfortable sharing location data with a public health official.

Successful suppression of the spread of infection through contact tracing, unless the use is mandatory and enforced as seen in Singapore (The Straits Times, 2020, October 20), must be achieved through ethical, trustworthy, locally rooted and adaptive implementation (Science Journal, 2020, November 13). The planned work aims to understand how the trust, intent to use, usability and satisfaction of a contact tracing can be increased

through designing a contact tracing app which adheres to a set of Trust Focused Design principles for use in Difficult to Trust scenarios which will be developed during this project.

2. LITERATURE AND PRACTICE REVIEW

This literature review aims to discover existing research and technology in the areas of trust, usability, satisfaction, and intent to use. The focus of this study is increasing trust of users, and so this topic will be researched in further detail.

2.1 Trust

Interpersonal trust is a vital aspect of human relationship, it is essential for relationship stability and vital for maintaining cooperation (Misztal, 1996). Rotter (1967) defines interpersonal trust as *'an expectancy held by an individual or group that the word, promise, verbal or written statement of another individual or group can be relied upon'* (p. 651). Trust is reliance that people we deal with relationally and practically will be capable of honouring their word. Trust allows us to live in uncertain situations and is dictated by 3 major factors: Ability, Benevolence, and Integrity (Mayer, 1995). For trust to be established, the trustee needs to display a level of competency, the desire for the wellbeing or benefit of the trustor and prior history of following a set of principles that the trustor believes to be acceptable. Mayer also states that if there is no risk present in a relationship, there is no need for trust (Mayer, 1995).

2.1.1 Digital Trust

Our interactions with technology are not interpersonal, however as there is always a level for risk when interacting with a website or application, trust is an important factor. However, with digital or online trust, the trustee/trustor relationship is replaced with an object of trust/trustor relationship (Corritore, 2003). Although user/computer trust relationships aren't interchangeable with interpersonal trust relationships, a user's computer experience is 'fundamentally social' which doesn't come from a place of ignorance or dysfunction but is commonplace (Nass, 1994). Users will readily form dependency to computers in a similar way to that of human team relationships (Nass, 1996). To achieve the interpersonal trust characteristics of Ability, Benevolence and Integrity in a website, Schlosser (2006) states that a high level of investment in the website's technology and visual design elements, coupled with a strong privacy/security statement are needed. Additionally, there are specific factors that contribute to the level of trust a user has towards technology. Fogg (2001) states that the factors contributing to Web Credibility are: Real World Feel, Ease of Use, Expertise, Trustworthiness and Tailoring and that the factors of Commercial Implications and Amateurism decrease Web Credibility. Corritore (2003), defines online trust to be the perception of Web Credibility, Ease of Use and Risk which are affected by external factor that an individual user experience (See Figure 1).

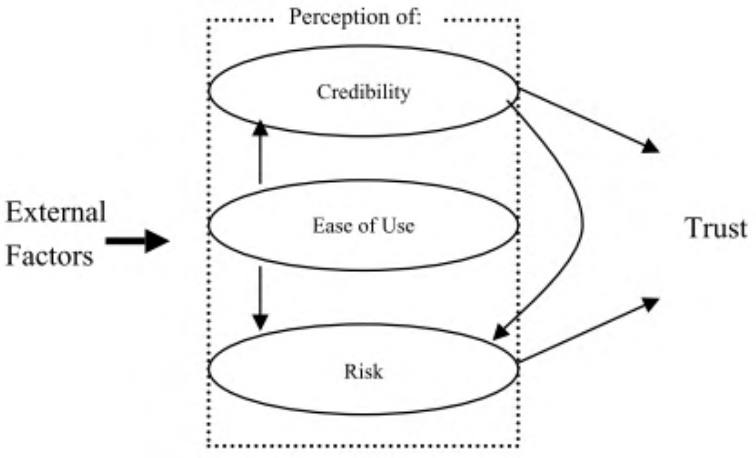

Figure 1. Corritore's Model of Online Trust

2.1.2 Measuring Trust

To measure interpersonal trust, (McKnight, 2000) developed a Trust Typology listing, scoring and grouping trust-related characteristics in order to achieve complete trust.

Trust-related Characteristic	Second Order Conceptual Category	Count	% of Total
1. Competent		14	
2. Expert		3	
3. Dynamic		3	
	COMPETENCE	20	20.4
4. Predictable	PREDICTABILITY	6	6.1
5. Good, Moral		6	
6. Good will		10	
7. Benevolent, Caring		18	
8. Responsive		4	
	BENEVOLENCE	38	38.8
9. Honest		11	
10. Credible		1	
11. Reliable		8	
12. Dependable		6	
	INTEGRITY	26	26.5
13. Open		3	
14. Careful, Safe		3	
15. Shared Understanding		1	
16. Personally Attractive		1	
	OTHER	8	8.2
	Grand Total	98	100.0

Table 1. McKnight's Trust Characteristic-based Definition Categories

These characteristics can be used to measure a trustor's level of trust in a trustee by discovering their opinion of each of the Trust-related Characteristics. In Table 1., McKnight groups the characteristics into five categorises of: Competence, Predictability, Benevolence, Integrity and Other. These findings strongly mirror the major trust factors of (Mayer, 1995) which are Ability, Benevolence, and Integrity.

To measure online trust, McKnight et al. (2002) proposed and validated a measure for a multi-disciplinary, multidimensional model of trust in e-commerce which included disposition to trust, institution-based trust, trusting beliefs and trusting intentions. In conducting this test, data was collected through a questionnaire in which participants responded to how much they agreed with a statement given to them in relation to their prior interaction with a website for legal advice. The questions were categorized under 4 trust constructs to quantify the participant's disposition to trust, institution-based trust, trusting beliefs and trusting intentions within the context of the legal site, these constructs were further divided into 16 subconstructs. Additionally, Personal Innovativeness, General Web Experience and Perceived Site Quality was also measured.

Similarly, Corritore et al. (2005) conducted a test in which participants filled in a Likert-type questionnaire of 34 component items under the constructs of Honesty, Reputation, Predictability, Perceived ease of use, Risk and Trust.

2.1.3 Designing for Trust

Designing for trust is heavily dependent on user base and content being displayed, however there are basic methods for achieving general trustworthy design. Nielsen (1999, March 6) states 4 methods of communicating trustworthiness:

- Design quality – appears professional, clear, and respectful
- Up-front disclosure – transparency, no surprises (e.g. hidden costs)
- Comprehensive, correct, current – content/product selection fell solid
- Connected to the Web – links in and out, association to third party sites

Additionally, Fogg (2001) states 7 factors that contribute to the credibility of a site, 5 which increase perceived credibility:

- Real World Feel
- Ease of Use
- Expertise
- Trustworthiness
- Tailoring

and 2 that decrease it:

- Commercial Implications
- Amateurism

To assist in the designing of trust focused infection tracing apps and attempt to reduce the threat of the introduction of unethical contact tracers, Gillmor (2020) developed 14 principles of contact tracing trust design:

- Not displacing non-technical measures
- Voluntary
- Non-punitive
- Built with public health professionals
- Privacy-preserving
- Non-discriminatory
- Minimal reliance on central authorities
- Data minimization everywhere
- No data leakage
- Measurable impact
- Have an exit strategy
- Narrowly tailored to target a specific epidemic
- Auditable and fixable
- Sustainably maintained

2.1.4 Trust in Health Apps

Trust in Health or Medical apps is different to trust in conventional apps as there is often additional risk involved when using Health or Medical apps. Lewis and Wyatt (2014) state that risk of harm when using medical apps depends on three dimensions: probability and severity of harm, the inherent complexity of the app and external factors specific to users. Their ‘App-space’ for risk assessment, seen in Figure 5 maps the level of risk of a medical app based on chance of harm and app complexity.

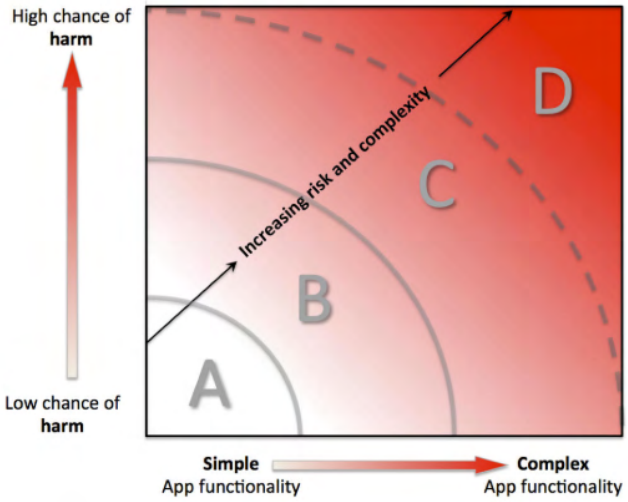

Figure 3: Lewis and Wyatt (2014) Risk Assessment of mobile medical apps

As level of risk is a key factor in digital trust, chance of harm and app complexity should be considered when attempting to design for trust within a medical app.

Additional factors required to be included in the development of medical/health apps are to be uncomplicated and not burdensome to use.

For health apps to retain users it needs to not only provide adequate functionality but also for that functionality to be provided in an engaging manner. The Mobile App Rating Scale (MARS) for health app design contains 5 subscales to consider when designing a medical app to ensure a high level of quality. The subscales are: Engagement, Functionality, Aesthetics, Information and Subjective quality (Stoyanov, 2015).

Fitness apps, although less risky to users, also need to be developed to reduce complication and burden, they also be act as users expects them to. It is also important for fitness apps to adequately provide users with the functionality they desire. (Beldad, 2018).

2.2 Usability, Satisfaction, and Intent

2.2.1 Usability

Ease of use and a predictable interface are two of the main components to build online trust, along with instilling a belief of safety and that there is no benefit for the vendor to cheat. The collaboration of trust and usability are crucial to increase the purchasing intentions of a user (Gefen, 2003).

The usability of a system is heavily influenced by the degree to which users believes it will enhance their performance as well as the level to which users believes the system will be free of effort, these determinants are defined as ‘Perceived Usefulness’ and ‘Perceived Ease of Use’ (Davis, 1989).

To measure these determinants, Davis (1989) developed two scales with which a user could define the usability of a product.

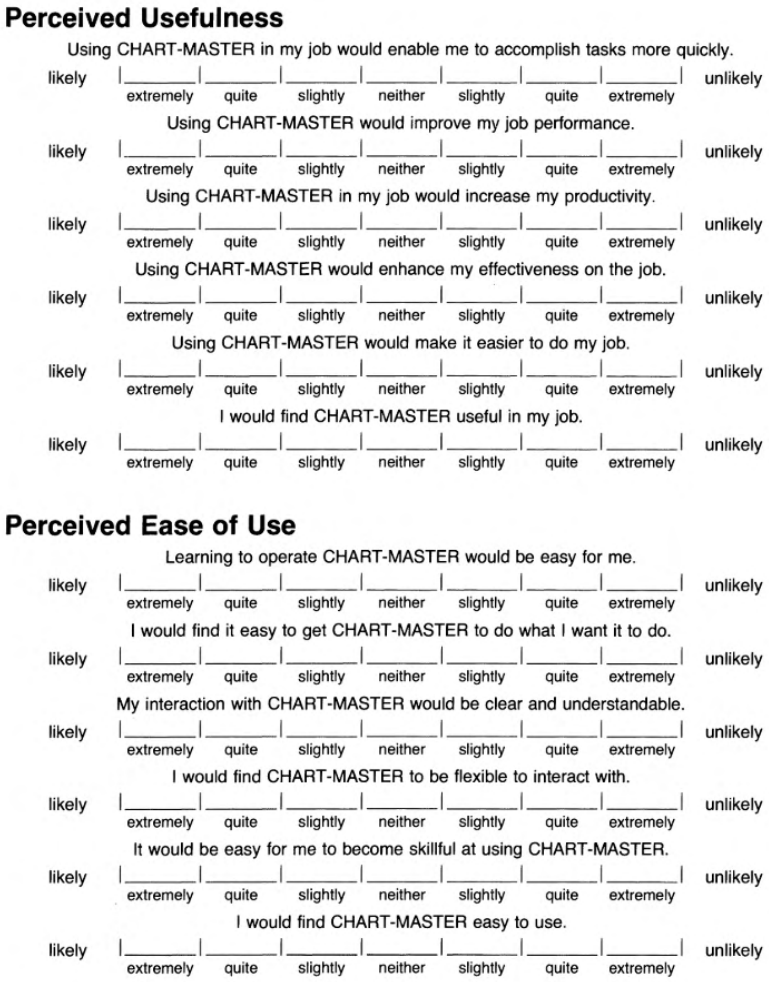

Figure 4. Measurement Scales for Perceived Usefulness and Perceived Ease of Use (Davis, 1989).

Nielsen (1990) also developed a method of usability analysis in the form of usability heuristics which forms results not based off the perception of users. Nielsen does state however that the heuristic evaluation of one individual should not be relied upon but rather the separate evaluation of 5 to 10 individuals for valid results. The heuristics to be used in the assessment are:

- Simple and natural dialogue
- Speak the user's language
- Minimize user memory load
- Be consistent
- Provide feedback
- Provide clearly marked exits
- Provide shortcuts
- Good error messages
- Prevent errors

Another popular measurement of usability is the System Usability Scale, which offers a simple 10 question Likert scale for use after a participant has interacted with the system to be tested. The SUS is described as a 'quick and dirty' method of usability assessment due to its reliability, low cost and its ability to be used for global assessments of systems (Brooke, 1995).

2.2.2 Satisfaction

Satisfaction is also an important consideration but is closely tied to the usability of a product. User satisfaction is defined by 9 variables (Figure 5) and these factors are grouped into the areas of Perceived benefits, Organizational

support, and User background the factors of perceived usefulness and perceived ease of use are especially influential (Mahmood et al., 2000).

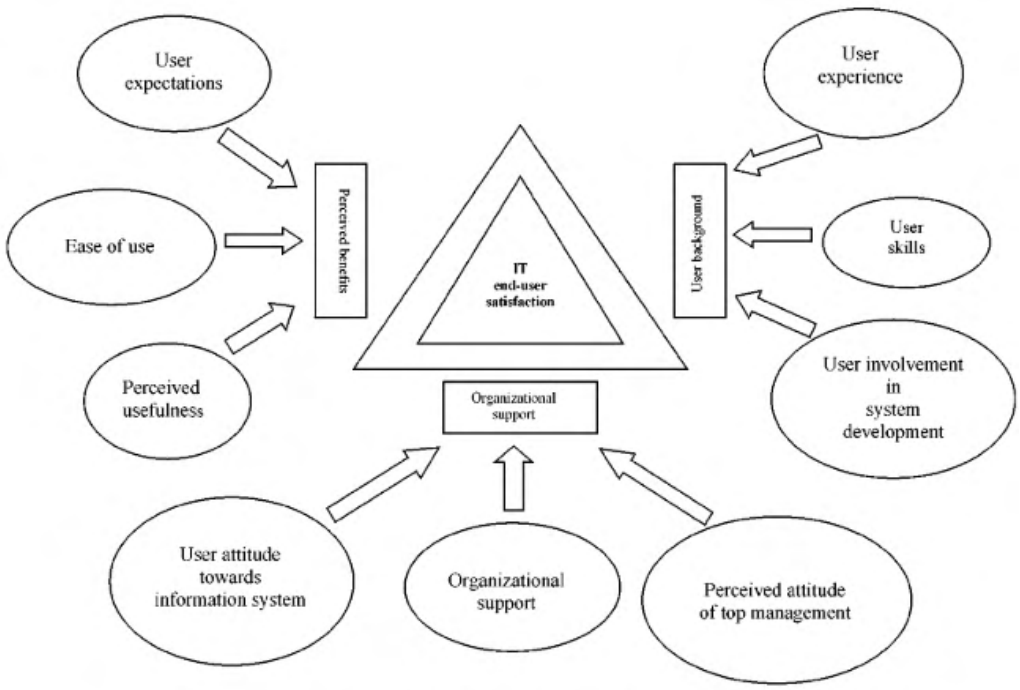

Figure 5. Mahmood's Model of IT End-User Satisfaction (Mahmood et al., 2000).

Ditsa and MacGregor (1996) list additional factors which contribute to user satisfaction, after examining a range of satisfaction models they identified 5 influential factors which are: Quality of Information, User Interface Features, Provided Support, User Involvement and User Attitude.

Satisfaction is often measured as part of overall usability, as it is considered a component of usability (Bevan et al., 1991) however there are a series of stand-alone tests that can be used to assess satisfaction. Some of the most prominent satisfaction scales include IBM's Post-Study System Usability Questionnaire (PSSUQ), Computer System Usability Questionnaire (CSUQ), After-Scenario Questionnaire (ASQ) and Printer Scenario Questionnaire (PSQ) (Lewis, 1991). Of these four, the ASQ has been the most widely used being accredited with strong generalizability of results and wide applicability (Lewis, 1995). As well as this, the ASQ has been accredited with satisfying the Standardized satisfaction measurements of Objectivity, Quantification, Communication, Economy and Scientific generalization (Nunnally, 1975).

2.2.3 Intent to Use

Intent, in the context of online transactions, is defined as '*the buyer's intention to engage in online exchange relationships with the community of sellers*' (Gefen et al., 2003).

Intent is directly influenced by how much a user trusts the supplier (Gefen, 2000). The higher perception of risk a user has towards a system, the more their consumer transaction intention is reduced with the user's willingness to purchase being directly related to the consumer's trust in the supplier (Kim et al., 2009).

Intent is deeply entwined with trust, usability, and ease of use with the three directly impacting the level in with a user intends to transact with an online system. This is seen in Pavlou's (2001) model of the factors that impact Intention.

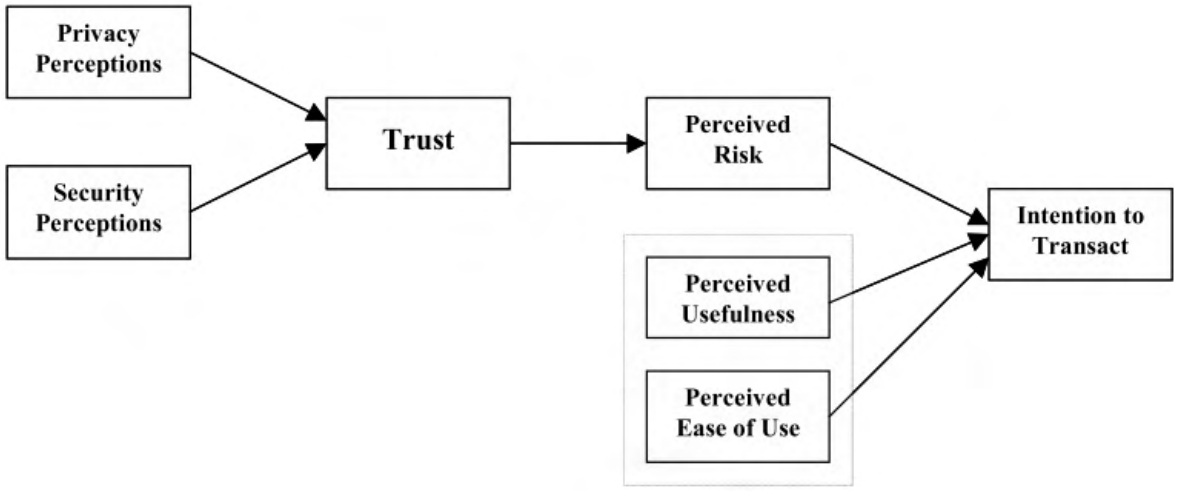

Figure 6. Pavlou's (2001) Model of Intention to Transact.

To measure intent, Pavlou and Gefen (2004), developed a subscale, as part of a scale to test the effectiveness of an online marketplace. This is made up of 3 simple questions which assess the user's prediction of whether they will interact in the future, the user's future likelihood or interacting and the user's current intention to interact in the future.

2.3 Practice Review

To assess how similar apps, including contact tracing apps, exercise apps and navigation apps, attempt to implement trust, a series of factors and principles included in established studies (Nielsen, 1999; Fogg, 2001) were used to assess the difference of features between trusted tracker apps and contact tracing apps. The result of this practice review displays how different apps practically achieve different trust principles through features and shows that several contact tracing apps contain features attempting to achieve trust in users. These findings showed that even with the inclusion of several trust principles, participants still are cautious about contact tracing apps and ultimately prompted the emphasis for designing in a Difficult to Trust scenario.

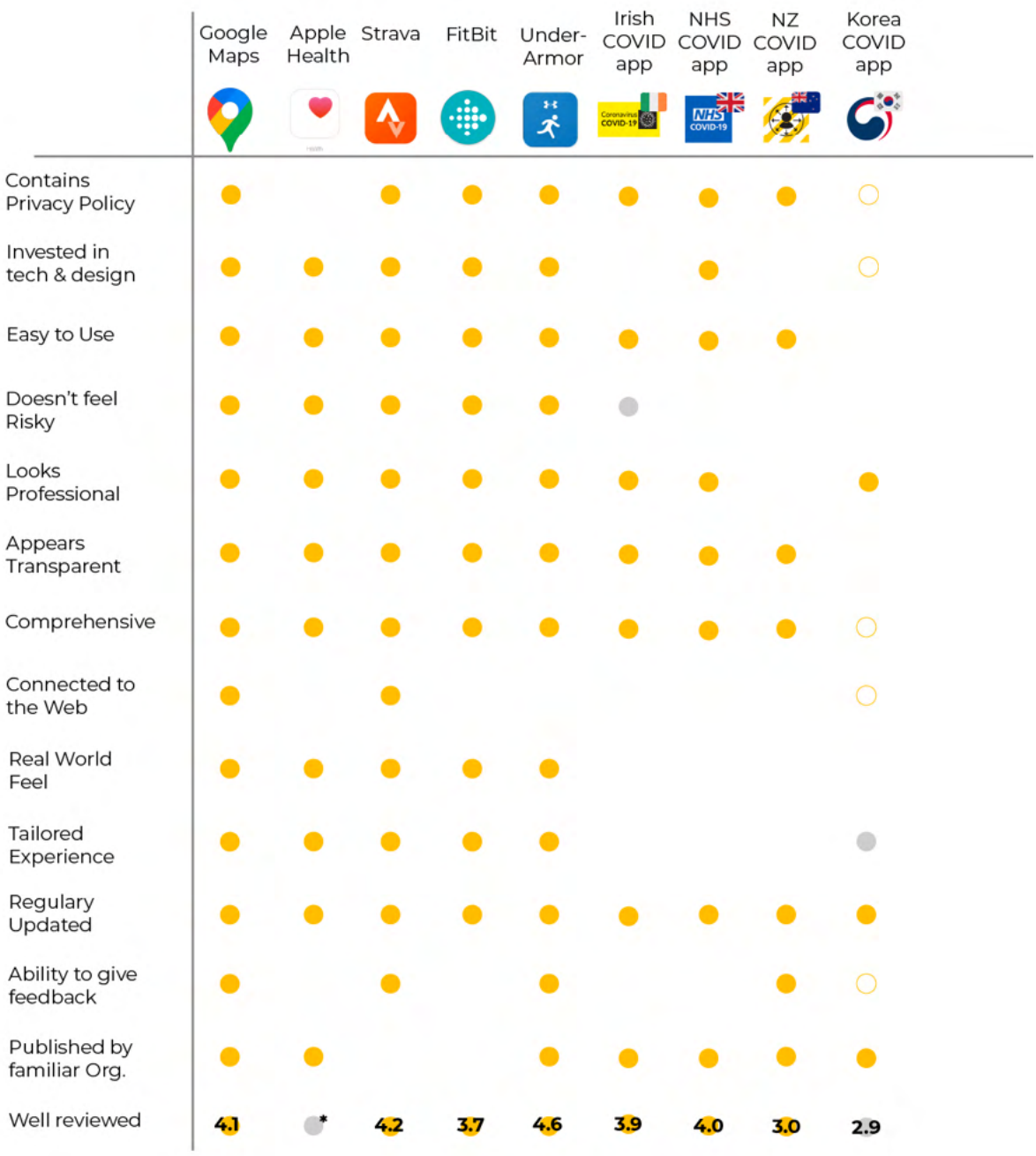

Figure 7. Competitor Analysis

2.4 Research Question and Hypotheses

Primary research of this study noted that designing for trust is a well-researched topic for use when designing systems. It was also found that usability and satisfaction are intertwined with achieving trust of users and ultimately convincing them to use the system. It was found that, although many contact tracing apps attempt to adhere to established design principles, the public's trust in them remains low. To attempt to overcome this, Trust Focused Design Principle were created and implementing into a prototype with the main research question being:

> **What is the impact of Trust Focused Design Principles on a user's Trust, Usability, Satisfaction, and Intent to Use when using a contact tracing app?**

Along with aiming to answer the main Research Question, this paper will aim to test four separate Null Hypotheses:

1. **There will be no significant difference in the level of trust between an app with implemented Trust Focused Design Principles and an app without.**

2. **There will be no significant difference in the level of perceived usability between an app with implemented Trust Focused Design Principles and an app without.**

3. **There will be no significant difference in the level of user satisfaction between an app with implemented Trust Focused Design Principles and an app without.**

4. **There will be no significant difference in the level of intent to use between an app with implemented Trust Focused Design Principles and an app without.**

3. METHODOLOGY

The aim of this project is to create a set of Trust Focused Design Principles based off literature and primary research. These principles will be implemented into a contact tracing prototype and will be tested against a similar prototype absent of the principles. The process of this project followed the Stanford Design Model (Stanford d.school, 2021) and was carried out in 3 phases. Firstly, the Empathize & Define phase, followed by the Ideate & Prototype phase and finishing with the Test & Assess phase, as seen in the figure below:

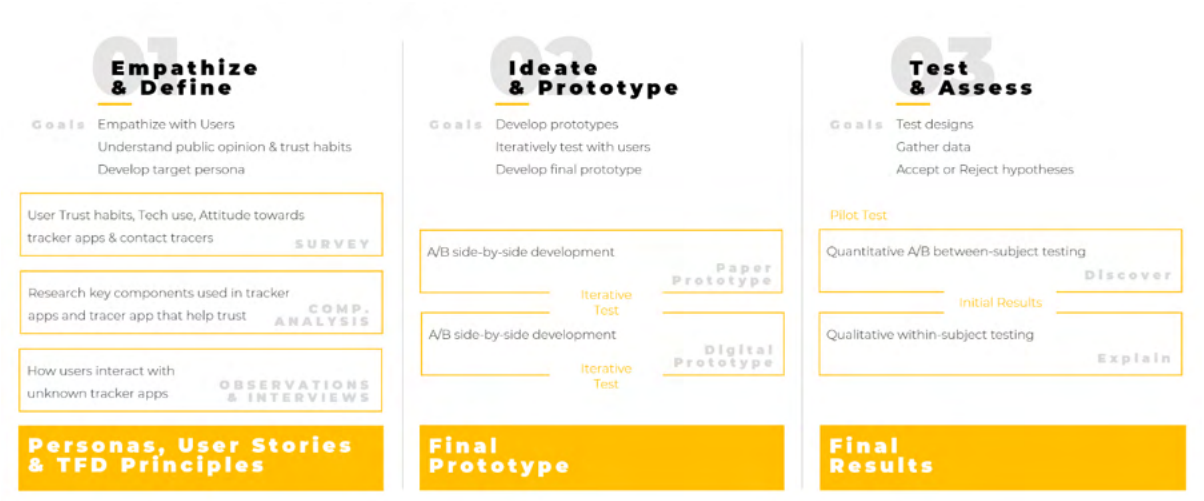

Figure 8. Project Methodology

3.1 Phase 1

The Empathize & Define phase of the project aimed to understand the target user, understand public opinion regarding tracking apps, understand public trust habits and finally, developing personas and user stories derived from the three forms of research; a public survey, a competitor analysis and qualitative observations and interviews.

A public survey was used to discover the public's trust habits, disposition to trust, technical literacy, using of tracking app and attitudes towards contact tracing apps. Secondly, a competitor analysis of various tracking apps (including contact tracing, navigation and exercise apps) was conducted, comparing the apps against a list of trust increasing features based off established trust models (Nielsen, 1999; Fogg, 2001; Schlosser, 2006). This was done to assess

if more popular apps contain more trust features. Finally, qualitative interviews and observations were conducted to gain in-depth understanding of a user's experience with tracking apps and contact tracing apps. Participants were asked questions based on personal app preferences and experience with tracking apps. They were subsequently asked to complete a series of tasks using the Irish COVID tracker app and answered questions based on their experience, followed by a questionnaire derived from the McKnight (2002) trust scale.

The results of this research were used to create user personas and user stories to help empathize with the defined user base. By identifying successful trust features in the competitor analysis, highlighting trust trends from the public survey, and listening to qualitative feedback, paired with literature research, 7 Trust Focused Design Principles were created for designing in Difficult to Trust scenarios.

3.2 Phase 2

In the Ideate & Prototype phase, two designs of a COVID contact tracing app were iteratively created. The first being a Trust Focused app, designed in adherence with the Trust Focused Design Principles. The second app is visually similar and attempts to represent current contact tracing apps while being absent of any of the Trust Focused Design Principles. To emulate an established Difficult to Trust scenario, both designs take visual and structural components from the current Irish COVID tracker app. The design process of these apps was a side by side development of paper prototypes, wireframes, basic digital prototypes and finally, finished digital prototypes.

3.3 Phase 3

The final phase was to test the two prototype designs and assess if the presence of Trust Focused Design Principles can have a significant impact on the level of trust, usability, satisfaction, and intent to use a user experiences when using an app within a Difficult to Trust scenario.

The testing was conducted in two stages using a Discover-Explain approach, where a Quantitative A/B between-subject testing would be used to Discover the results of using Trust Focused Design Principles followed by a Qualitative within-subject testing which would be used to Explain the results of the first test stage. For each stage of testing, participants were asked to confirm they were over 18 years old and to give consent for their responses to be used anonymously in the project.

Firstly, the Quantitative A/B between-subject test was conducted online where participants were asked to complete a series of tasks using one of the prototypes and then asked to complete a questionnaire which assessed their trust disposition and level of trust, perceived usability, satisfaction and intent to use relating to the prototype. The questionnaires were created using the following scales: the McKnight trust scale (McKnight, 2002), the After-Scenario Questionnaire (Lewis, 1995), the System Usability Scale (Brooke, 1995) and an Intent to Use scale derived from Pavlou and Gefen (2004). These scales were chosen based on their established reliability and concise focus on a specific topic. A/B between-subject testing was used as it was expected that a user, if testing both prototypes, would be more familiar and more confident when using the second prototype due to their experience with the first which could alter their responses, particularly trust.

The second stage of testing was a Qualitative within-subject test where participants were selected from the results of the exploratory questionnaire to represent the userbase, particularly trust disposition. Participants were asked to complete a series of tasks using prototype A, answer questions based on their experience and then were asked to complete a similar series of tasks with prototype B and to answer questions based on that experience. Using within-subject testing allowed participants to comparatively identify features and components which affected their propensity to trust and their responses used to explain the statistics gathered in the first stage of testing. The responses of the second testing stage were collected, coded and where possible the coded experience was attributed to a feature, the absence of a feature or directly to one of the Trust Focused Design principles.

4. DESIGN, PROTOTYPING AND EVALUATION

4.1 Exploratory Research

The Exploratory Research for this project consisted of 3 components, a quantitative survey, qualitative interviews and observations and a competitor analysis.

4.1.1 Survey

The qualitative survey was conducted online and assessed participants' opinions on technology use, disposition to trust, tracking app usage and opinions on current COVID contact tracing apps. The trust section of the survey was derived from the McKnight trust scale (McKnight, 2002).

The survey received 66 submissions recruited through social media, online public forums and through friends and family. 55 participants were aged 18-34 with the remainder being 35-65, 36 of the participants were female, 28

were male and 2 identified as other, 47 participants were residents of Ireland. 51 participants had achieved some form of 3rd level education certificate.

Some key findings from this survey are as follows:
In terms of technology comfort and use, over 90% of participants agreed or strongly agreed that they were very comfortable using a computer, navigating the internet, using a smart phone, and using new apps. Over 70% of participants used smart phones socially for 8+ hours a week.

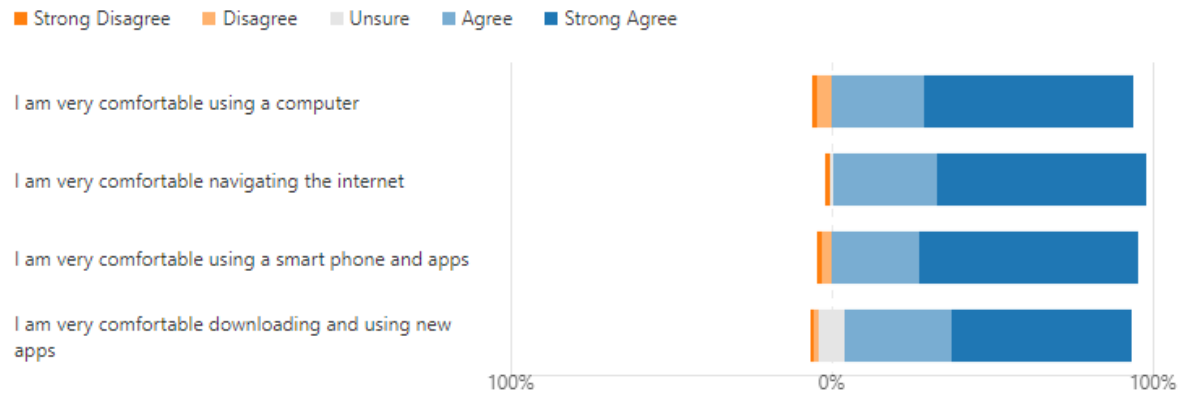

Figure 9. Survey Technology Use results

In relation to general trust, 79% of participants had a positive trust stance towards people they haven't met before and 61% had a thought positively to the structural assurances of safeguards and environments on the internet.

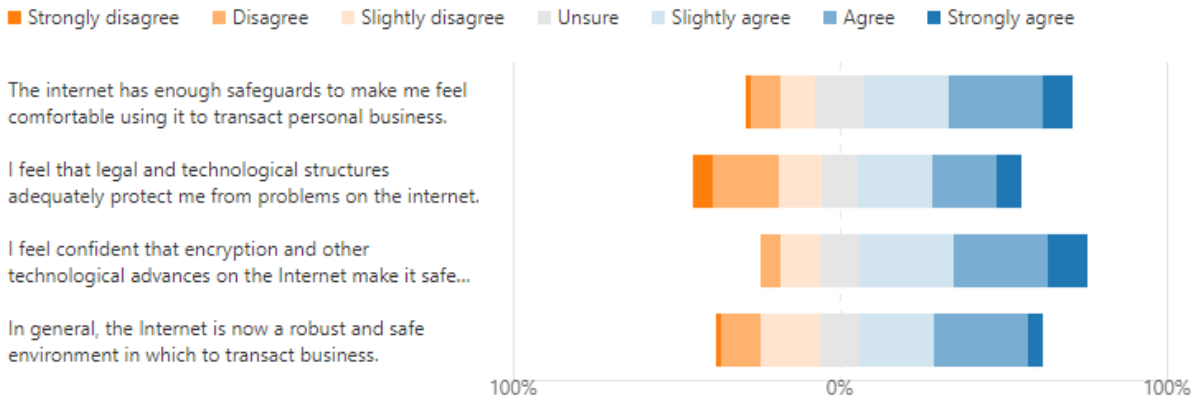

Figure 10. Survey Structural Assurance results

67% of participants used one or more tracking apps, predominantly for navigation and exercise with the main deciding factors on why they chose their used app being Ease of Use and Convenience. Additionally, some features that participants related to an app being trustworthy included: professional look, familiar publisher and the app containing a privacy policy.
Although the majority of participants were comfortable with technology, used smart phones frequently and were generally trusting of people they didn't know and towards the internet, 46% of participants who downloaded a contact tracing app rarely or never updated it while 72% of participants were unsure or didn't believe contact tracing apps were sufficiently transparent.

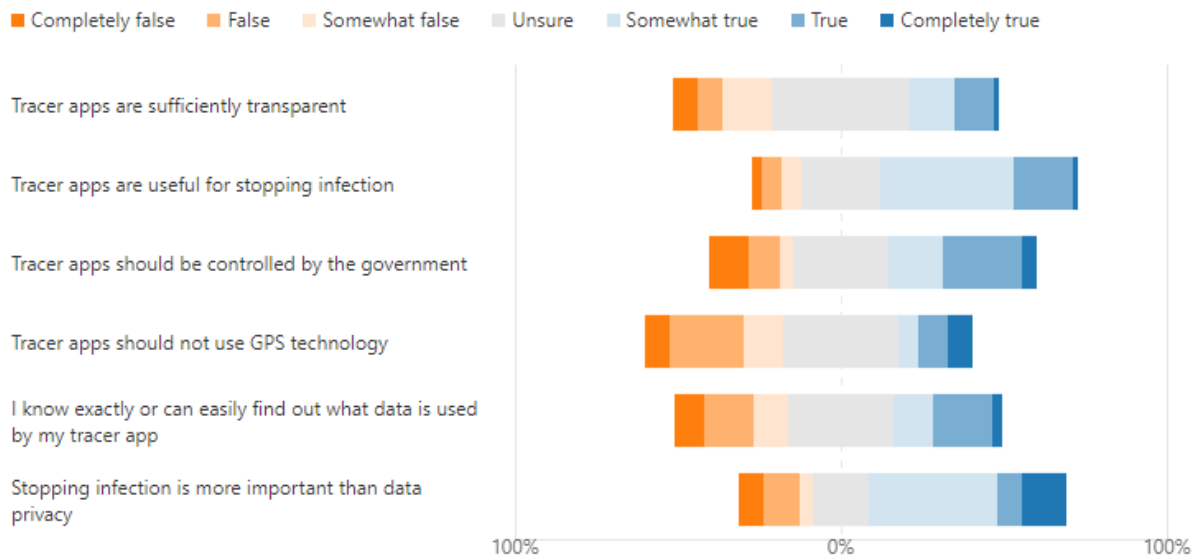

Figure 11. Survey Perception of COVID Tracing Apps results

4.1.2 Interviews and Observations

The second stage of research for the project involved 5 qualitative interviews with members of the Irish public who were chosen to represent the demographics of the survey, the main deciding factor when choosing participants was technical literacy. The interview questions covered app usage, specific preferences and levels of trust towards a frequently used tracker app by the participant, a walkthrough of the Irish COVID tracker app, questions relating to their experience and an online questionnaire derived and adapted from the McKnight trust scale (McKnight, 2002).
The 5 participants were made up of 2 females and 3 males with ages ranging from 19 to 59. The participants' varied occupations were: History Student, Health and Performance Student, Software Engineer, Software Developer and Accountant. These varying backgrounds facilitated the desired variety in technical literacy.

Some of the main findings of these interviews are listed below:
For all the participants there was an acknowledgement of general distrust of tracking apps with functionality, convenience and a sense of control being the key factors which cause them to overcome their sense of distrust. In relation to the Irish COVID app some key findings are that all participants said they found the app trustworthy, however one participant also stated:

'I trust the intentions of the people behind the app, but good intentions don't always mean good results. I don't have much trust in the technical capability of the app'

Some factors of the app which were critiqued by participants include: size and structuring of the Data Protection page, the display of figures and statistics, lack of proof of success and lack of updates.
3 of the participants compared the Irish COVID tracker app to other contact tracing app in other countries. A critique of the Irish design, unlike countries such as South Korea and Singapore, was it was not enforced, mandatory and relied too much on the individual to be accurate. A merit of the Irish model was that it was open source, public, used exposure notifications and generally not invasive, highlighting the importance of balancing public trust with functional technology.

4.1.3 Personas

From the results of the Qualitative Interviews, 2 personas were developed representing the 2 major identities of the participants:

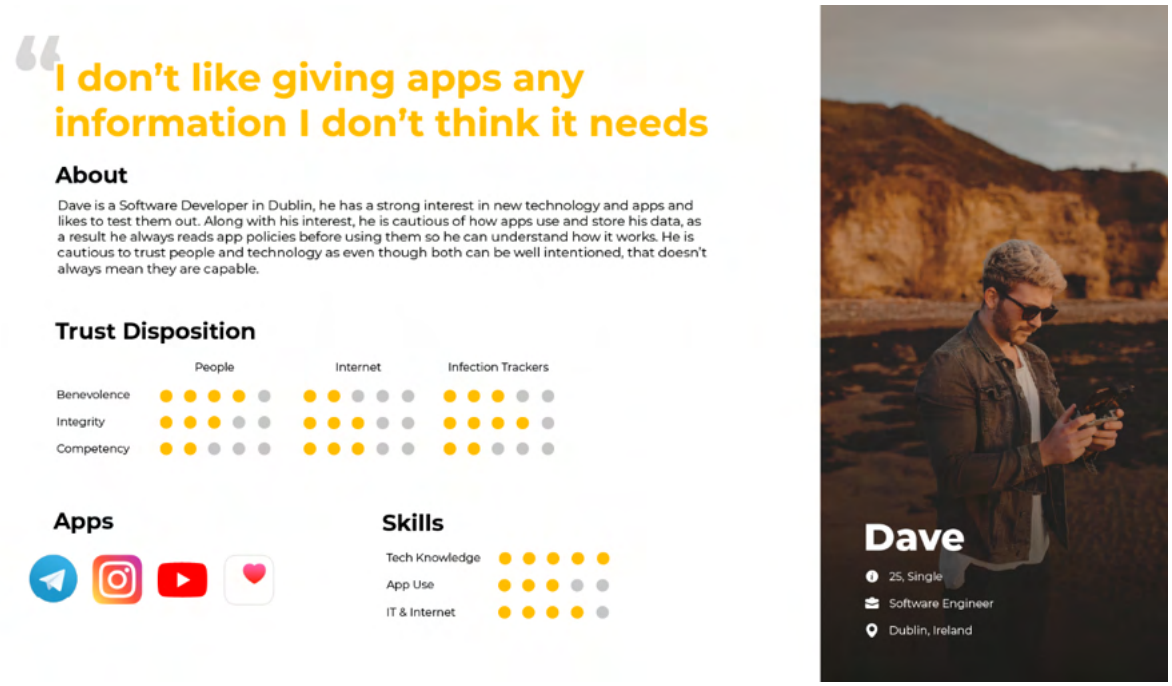

Figure 12. Persona 1. The Engineer

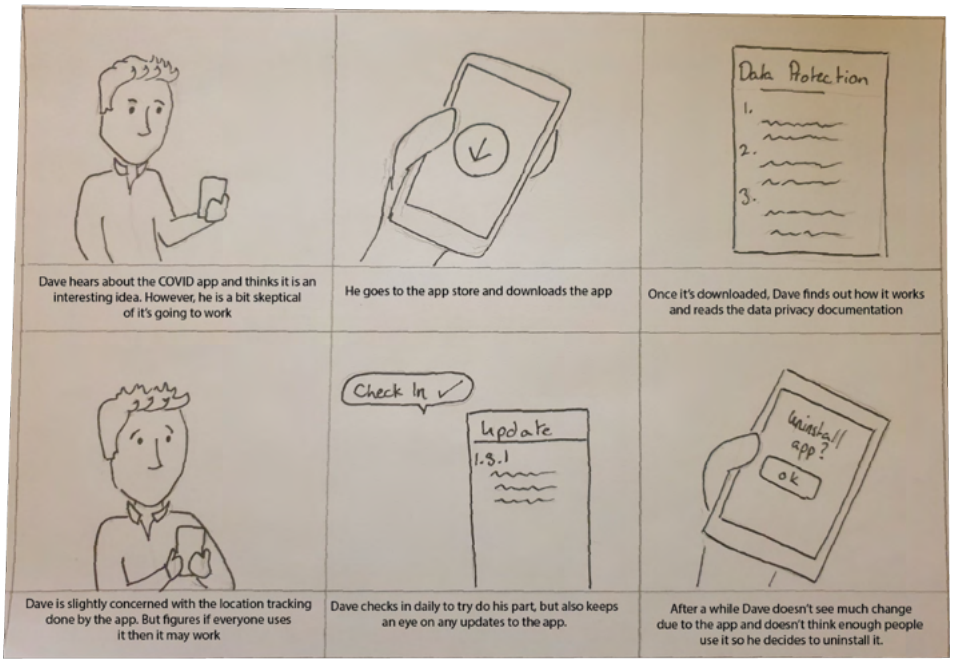

Figure 13. User Story 1. The Engineer

The first persona represents the technically literate user who has a stronger understanding of how the apps they use work, is more likely to distrust the intentions behind an app and is more likely to want to read statements and policies before committing to trusting an app.

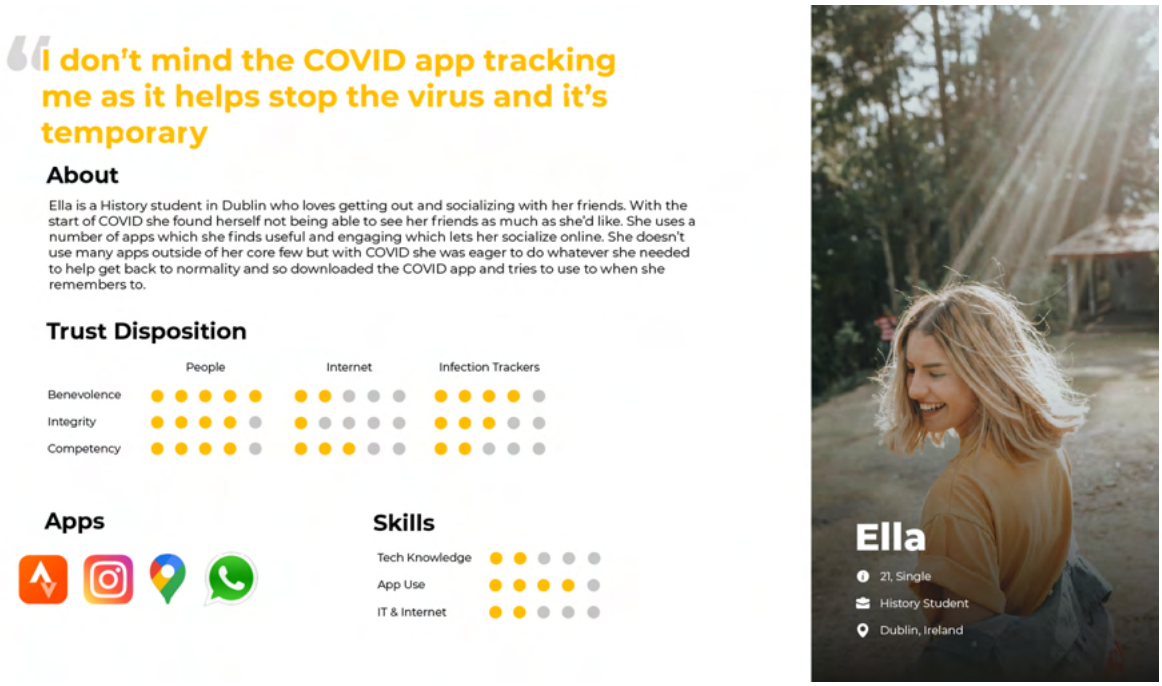

Figure 14. Persona 2. The Socialite

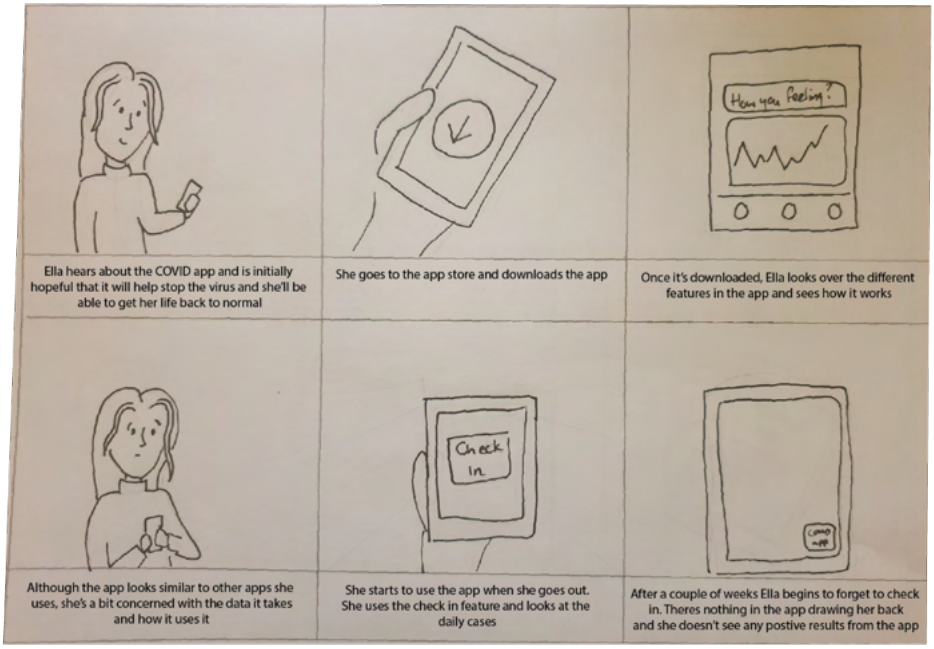

Figure 15. User Story 2. The Socialite

The second persona represents the less technically literate user who is comfortable with the technology she uses often but doesn't like to venture too far from what's familiar. Trust in an app for this user comes more from familiar components and the existence of, rather than the contents of, perceived safeguards, such as security policies.

4.2 Trust Focused Design Principles

Using the results of the exploratory research and findings from the literature review, 7 Trust Focused Design Principles were developed with the goal to increase trust, usability, satisfaction, and intent to use when a user is in a Difficult to Trust scenario.

The process of creating the Trust Focused Design Principles was carried out in three parts. Firstly, the main concerns participants had regarding tracking apps and contact tracing apps were identified by common trends from the survey results, qualitative feedback from participants being interviewed and observed. Following this, solutions were created to attempt to remove the concerns of users. This involved assessing the literature for solutions to similar problems, listening to participant feedback and critiques and conducting specific competitor analyses on apps who have combatted similar user concerns well. Once the main concerns were highlighted and solutions for these concerns were established, the final step was to group similar solutions together and form a list of principles. These principles are:

1. Accessible Privacy Policy

The presence of a privacy policy has been shown to be an establishing factor of credibility and trust for a user (Schlosser, 2006), the importance of explaining the use and collection of data is also highlighted as an important factor when developing a contact tracing app (WHO, 2020). This importance of privacy policies as reiterated in the survey results where 45% of participants stated that a policy's presence increases their confidence in an app. The decision to specify accessibility of a privacy policy came from the qualitative interviews where participants exclaimed the size and inaccessibility of the current Irish COVID tracker app caused suspicion and a lack of trust in the app.

2. Explain Yourself

Like Nielsen's Credibility Factor of 'Upfront Disclosure' (Nielsen, 1999), 'Explain Yourself' expresses the importance of avoiding hidden information and making your intentions transparent. This has been longstanding factor of increasing user trust however this only becomes more important when dealing in Difficult to Trust scenarios, such as with contract tracing apps (O'Callaghan, 2020; Abeler, 2020; McClain, 2020). In the primary research, control was seen as an important factor to participants where 'my data could be misused' was the second highest concern of those who had installed the a contact tracer and was a reason for over a quarter of participants not installing one.

3. Affirm Usefulness

The usefulness of an app is often self-explanatory and the reason a user chooses to install an app. However, when the usefulness is not clearly apparent and/or not fully understood, usefulness may need to be presented in an alternate form to instill confidence in the user. Disbelief in the usefulness of contact tracing apps was seen in exploratory research results where 'I don't think it will help' was the most common reason for participants who hadn't installed the app. Additively, many participants in the qualitative interviews felt that the Irish COVID tracker app wasn't useful or not as useful as it could be with one participant stating their reasoning for this opinion being the 'lack of success stories'.

4. Engage Your Audience

Engagement is an important tool for increasing a user's retention, trust, and satisfaction in a system (Stoyanov, 2015). Engagement features, such as gamification and leaderboards have been shown to increase follow up periods of self-management health apps compared against similar apps without engagement, even when assistive potential of both as been shown (Wu et al., 2015; Morton, 2005). From the qualitative interview results, engaging features which enhanced their experience was the main reason that most participants had picked their chosen tracker app over competitors.

5. Ensure Control

Assurance of Control is a crucial factor for retaining a user's trust in a system. This becomes more important in Difficult to Trust scenarios, in the context on contract tracing apps retaining trust is significantly more difficult in countries with compulsory installation (Ministry of Public Administration and Security, 2020; Technology Review, 2021). This is stated in previous guidelines for developing contact tracing apps where Voluntariness of the user is a requirement for ethical development (Gillmor, 2020; WHO, 2020). Although data misuse was a concern for the majority of participants in relation to tracker apps, results from the qualitative interviews showed that participants were willing to allow use of almost any form of data if it was explained to them why it was needed.

'I'm willing to give an app access to almost any form of data, if makes sense to why it needs it. But if an app requires data I don't think it needs, I won't use it.' - Participant 3

6. Look Professional

Professionalism and good visual design have been longstanding, important factors when designing for credibility and trust (Nielsen, 1999; Fogg, 2001). This was reinforced in the primary research survey where 'Visually Appealing' and 'Professional Look and Feel' were rated as some of the major attributes for increasing confidence

in an app. The same was found in the qualitative interviews where 'look and feel' were key for most participants finding an app trustworthy while buggy features and poor page structuring contributed to participants distrusting an app.

7. Easy to Use

Ease of Use is a crucial component when designing for trust and credibility (Fogg, 2001; Corritore, 2003). Features and navigation should be intuitive to the user with the design never restricting their ability to use a system. In the primary research survey, over a 3rd of the participants stated 'Ease of Use' as a major factor when choosing between functionally similar apps. This was also seen in the qualitative interviews where ease of use was highlighted as a key factor as to why participants had chosen their preferred app over competitors, and was detailed as an important feature contributing to their trust in apps.

4.3 The Design Process

After creating the Trust Focused Design Principles, they were implemented into a prototype design and tested against a design absent of the Principles.

In order to simulate a Difficult to Trust scenario, it was decided to implement the Principles into a contact tracing app, visually similar to the Irish COVID tracker app as this is a Difficult to Trust scenario that many of the participants that would be testing the prototype would have experienced. The second prototype, which the first would be tested against, was designed to be visually similar to the first and to the Irish COVID tracker app, however with the absence of any features or components which might fulfill any of the Trust Focused Design Principles.

The decision was made at the beginning of the design phase to not make the second prototype intentionally unprofessional or uneasy to use as these principles have already been shown to increase trust and usability while intentionally not including them might have sabotaged the testing of the effectiveness of the other principles.

4.3.1 Prototype A Design

Before beginning the development of the prototypes, illustrations, typography and colour schemes were selected for use in the A Prototype. Although the decision was made to not intentionally make the B Prototype visually unappealing, no specific attention was taken to increase it with the colour scheme and design choices (such as lack of illustrations) being adapted from existing contact tracing apps.

COLOUR

- Comforting, credible
- Warmer, more comforting yellow compared to 'pandemic yellow' used in current tracing apps
- Calming blue to contract confident yellow

Figure 16. Prototype A Colour Scheme

TYPOGRAPHY

- Confident
- Secure, Reliable
- Modern

Montserrat Black

Montserrat Regular

Figure 17. Prototype A Typography

ILLUSTRATIONS

- Friendly
- Familiar
- Minimal

Figure 18. Prototype A Illustrations (Freepick icons from flaticon.com)

4.3.2 Paper Prototyping

It was decided that Prototype B would be designed to emulate current COVID contract tracing apps, excluding any features that may adhere to the first 5 Trust Principles. For Prototype A, features were ideated that would help adhere to the 5 Trust Focused Design Principles. The paper prototyping stage of this project was focused on defining what features would fulfill the principles through iterative design and testing. Below are the first 5 Trust Principles, what feature was chosen to fulfill them and the initial design for that respective feature.

Have an Accessible Privacy Policy:

To create an accessible privacy policy, two main design choices were made to achieve ‘accessibility’.

Firstly, the privacy policy linked at the top of the ‘Control Page’ allowed the privacy policy to be ‘two-taps’ away from the home page allowing for quick discovery by those using the app.

Secondly, the ‘accessible policy’ is broken into 3 sections, rather than an detailed ‘wall of text’. The summary sections allows any user to quickly view and understand the most important pieces from the privacy policy.

Following this, a contact information section offering the user connection to a ‘Data Protection Officer’ who would be able to answer questions relating to the privacy policy. This section will also put the user at ease, ‘humanizing’ the app and ensuring them that any worries they might have with the app are easily solved.

The final section of this page is the full privacy policy, containing any additional information that a user might want to know. The full policy should be as short as possible while not being absent of any information a user might want.

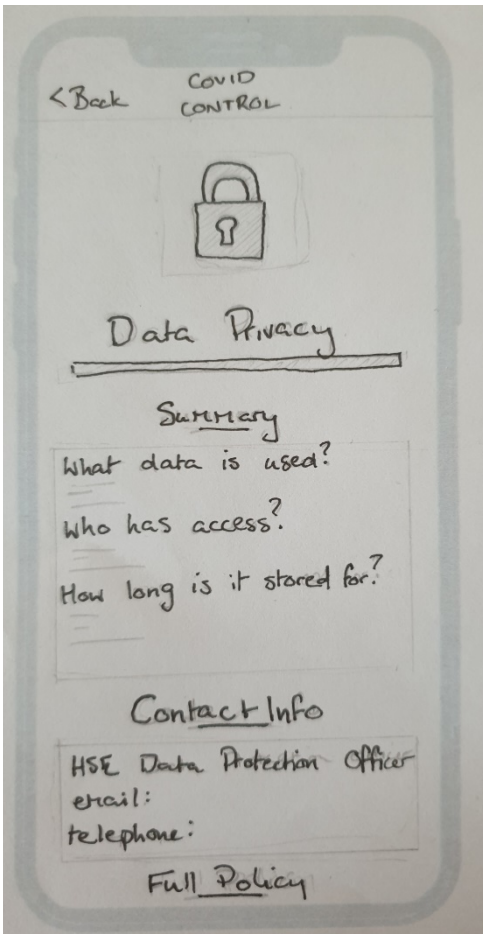

Figure 19. Privacy page paper prototype

Explain Yourself & Ensure Control:

‘Explain Yourself’ is achieved with a simple paragraph reassuring the user that they own and are in control of all the data the app requests to use and explains where to find information about the data used and how they can control the use of it.

For ‘Ensuring Control’, the app offers users the ability to turn off access to specific data, ability to pause tracking and the ability to request data being stored about them. Additionally on the profile page, users have the option to delete and leave the app, assuring users that all of their data is deleted.

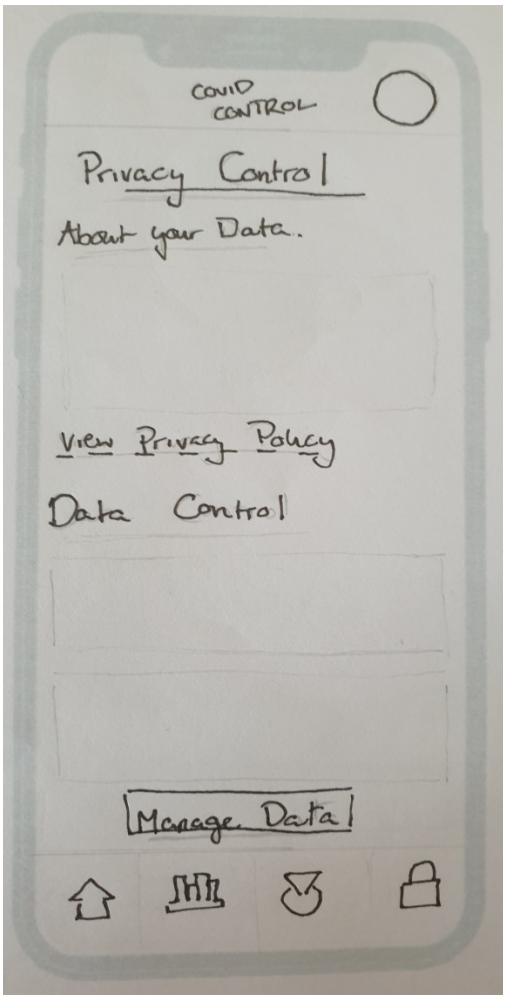

Figure 20. Control page paper prototype

Affirm Usefulness:

Displaying the usefulness of a contact tracer to users can be achieved by the app being used as a ‘check in’ token when entering public areas (TraceTogether, 2021, February 8). However, as this usefulness requires a large infrastructure outside of the app, the decision was made to achieve ‘Affirming Usefulness’ by displaying tangible information to users, helps them to quantify the app’s value. The main affirmation of usefulness is the ‘Potential Contacts Notified’ tab on the home page, giving an up to date number and explaination of potentially infected people the app has notified with an estimated number of people who will have avoided catching COVID-19 if all potential contacts adhere to the app’s recommendations. The app also contains a stats page detailing up to date pandemic information.

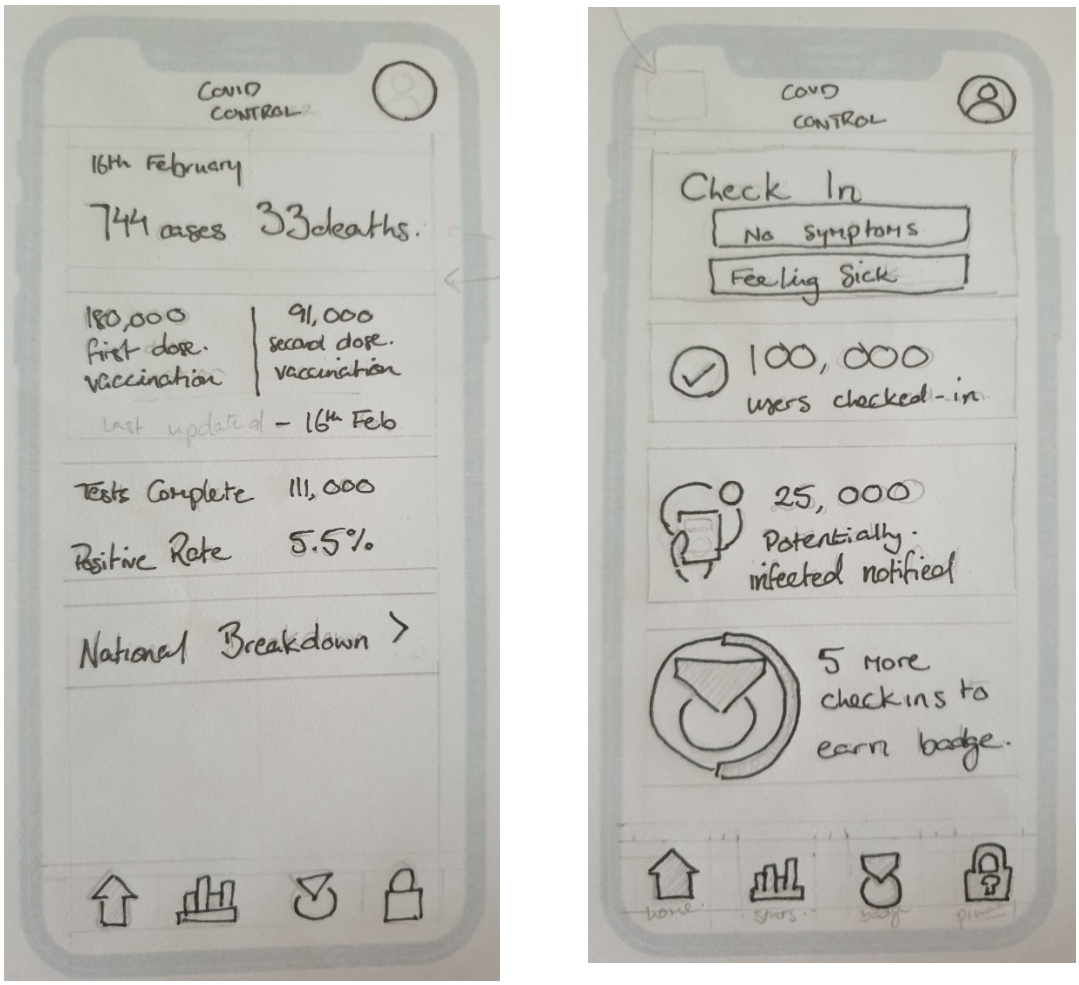

Figure 21. Home & Stats pages paper prototype

Engage Your Audience:

The final principle is achieved using two features. Firstly, the ability to complete challenges and earn badges. Users will be able to work towards collecting badges and can use the badges as profile avatars. The second feature is the ability to compete with friends and the nation and climb the leaderboards by earning points via challenges. The aim of introducing gamification and competition is to create a ‘drawback factor’ which brings users back to the app for more than just to check in.

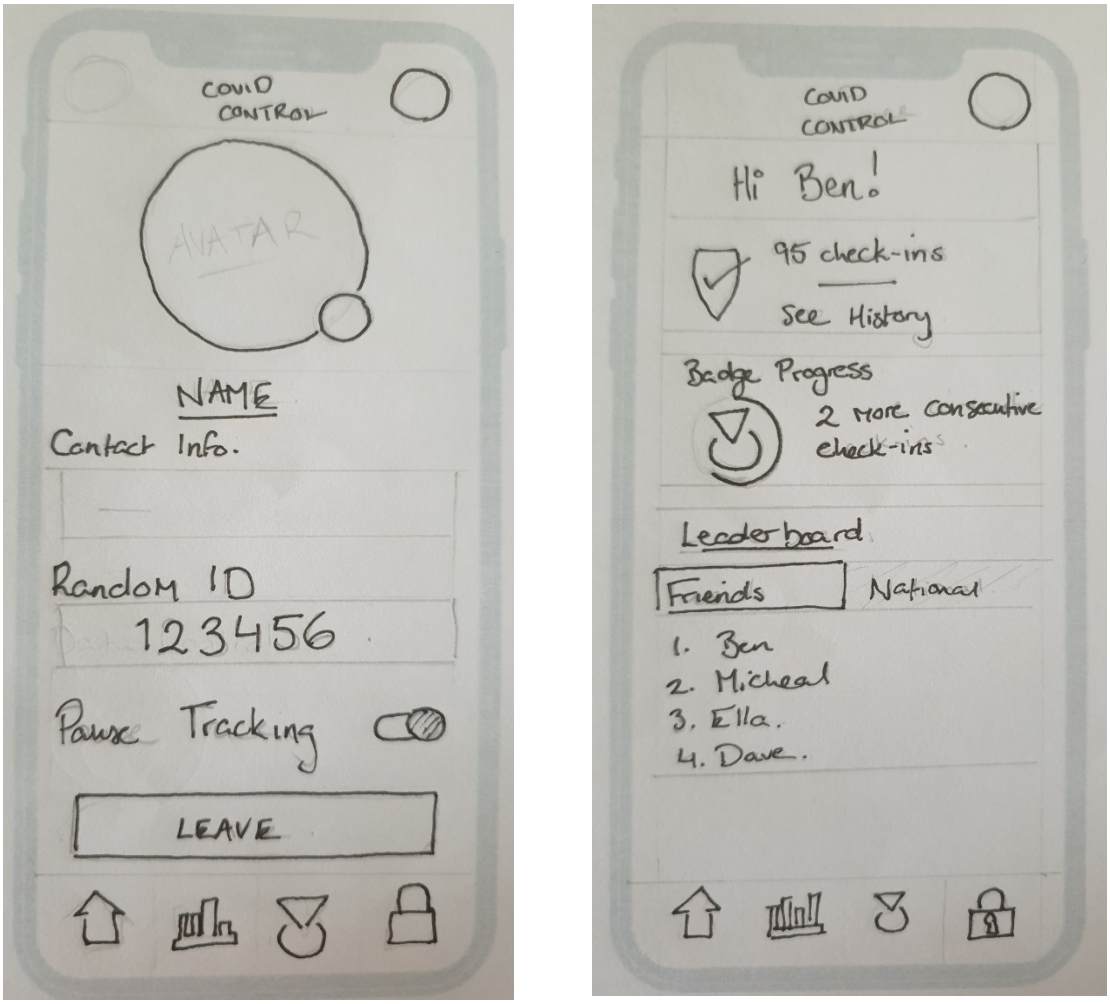

Figure 22. Social & Profile pages paper prototype

4.3.3 Pilot Testing and Iterations

Once completing initial paper prototypes, designs were coverted to navigable versions using Marvel (marvelapp.com). Using this app two participants were recruited to complete tasks taking them through all pages. Key findings from this include one participant stating a critique of the privacy policy button:

'I would've assumed the policy button was an external link and I wouldn't bother checking it'

Another useful comment from the second pariticpant was a comparision to the gamification and competition factors of Duolingo:

'I think some people might use the badge feature, but I wouldn't. Its similar on Duolingo, I know some people like the badges, but I don't bother. The leaderboard progress helps me to keep doing it'

The findings from this testing, along with additional competitor analysis of Duolingo, detailed wireframing for both prototypes were created, seen in Figure 23 and 24.

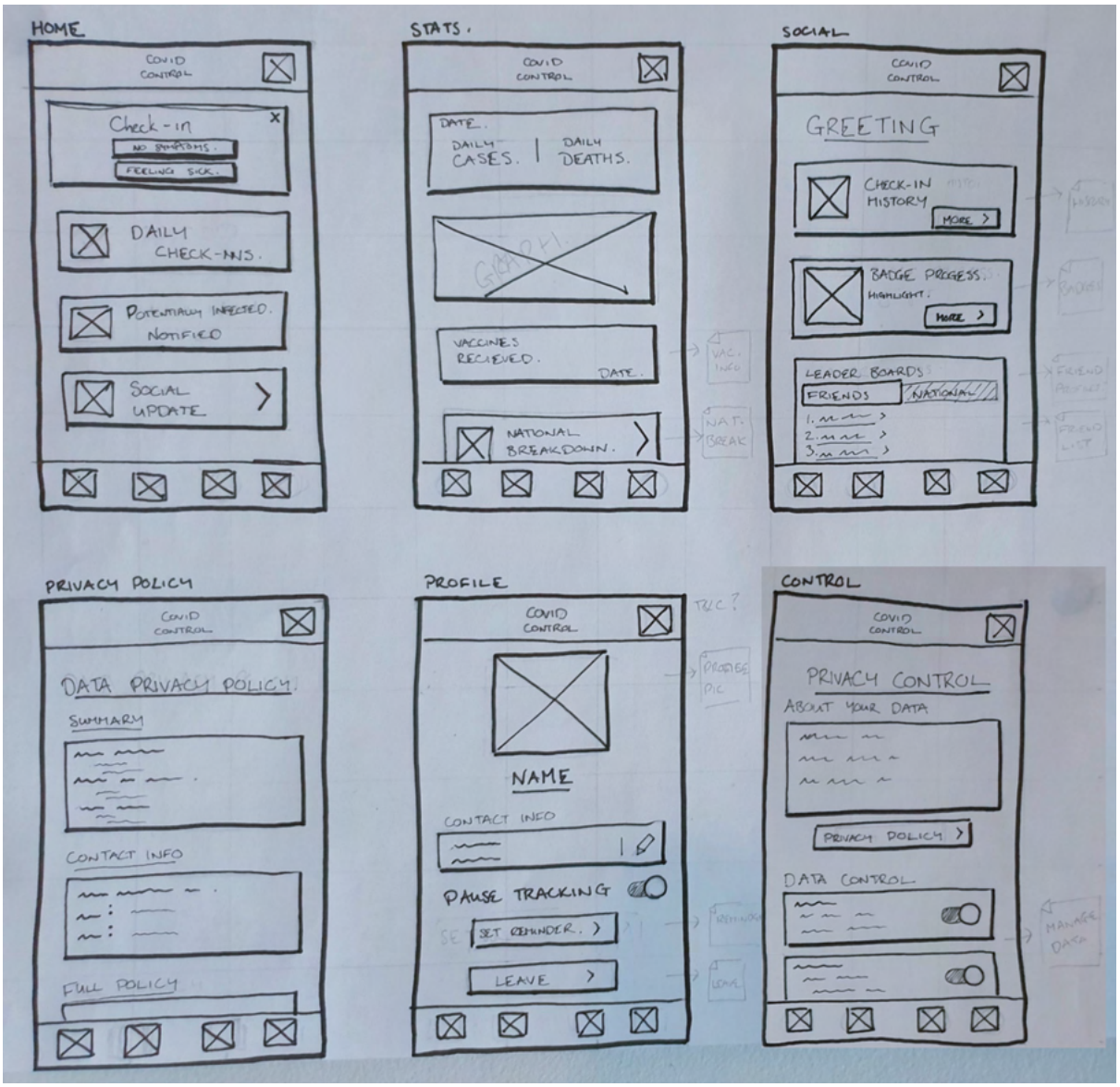

Figure 23. Prototype A wireframes

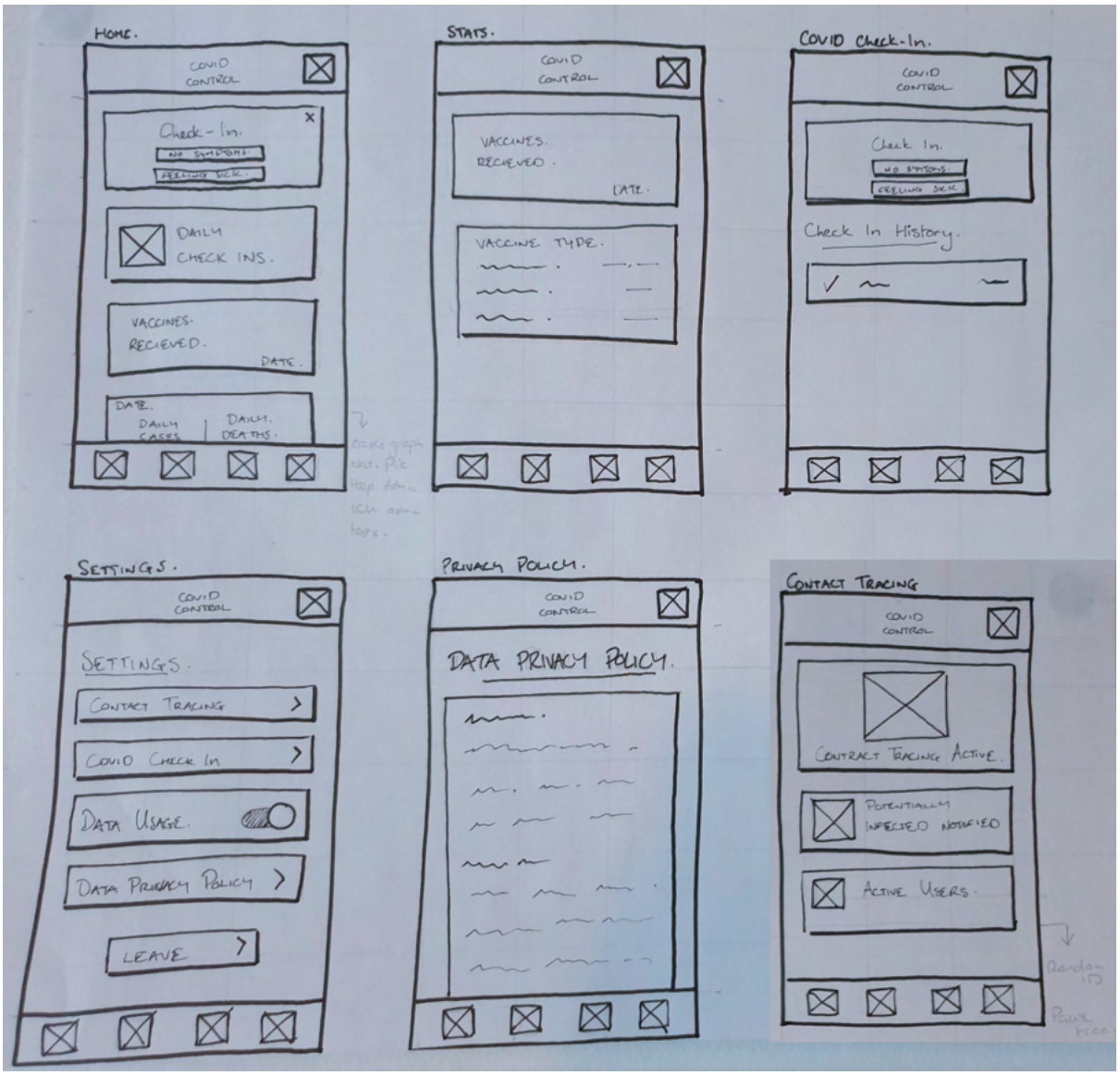

Figure 24. Prototype B wireframes

4.3.4 Digital Prototyping

After defining the main structure and design choices of the two prototypes, digital prototypes were created using Figma (www.figma.com).

Below are some of the key pages of both prototypes, detailing the differences with explainations of some of the design choices.

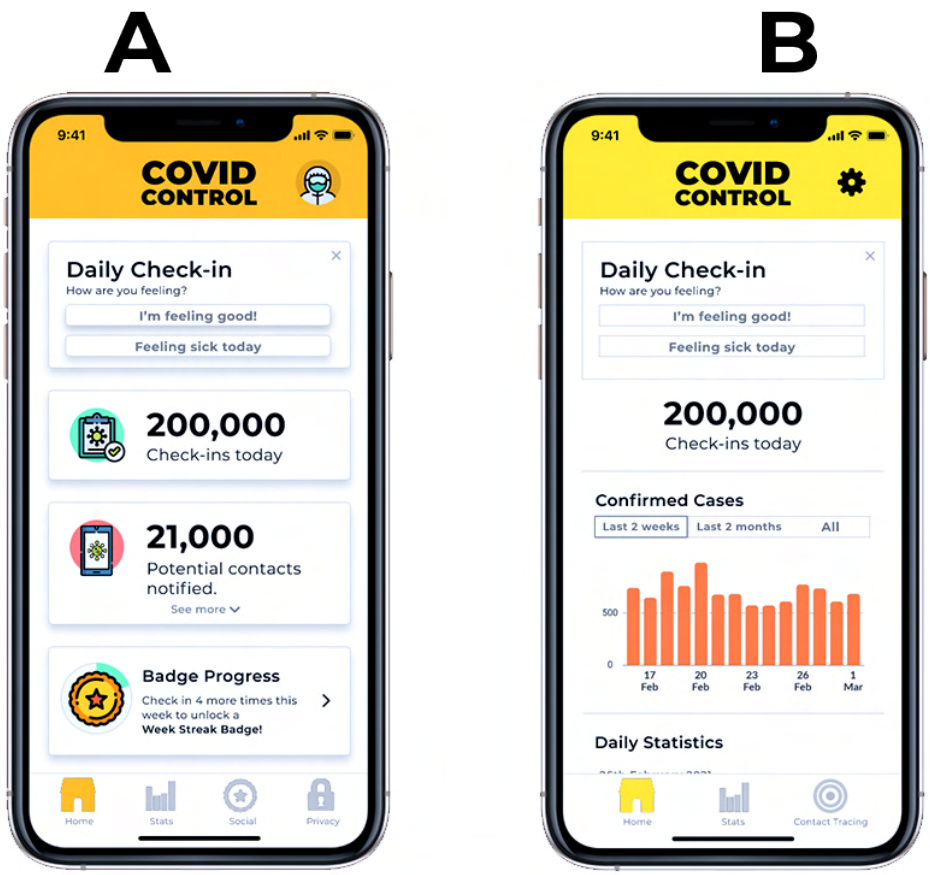

Figure 25. Home pages comparison

Home pages of both prototypes are similar with the core check in feature on both with the daily updates listed below. The core differences are, instead of repeating statistics on the home page as in prototype B as well as in many current contact tracing apps, prototype A contains the recent potential contacts notified with a drop down explaination as well as badge progress.

Figure 26. Stats pages comparison

As discussed previously, it was decided that prototype B's stats page would contain a similar level of content compared to prototype B as this feature is included across most contact tracing apps reviewed during this project. As a result these pages have many similarities with the main change being an enhanced level of visualization and clearer use of language which were common complaints participants had in the primary research interviews about the Irish COVID tracker app stats page.

Figure 27. Control pages comparison

Unlike the first 2 screens, the control page of the two prototypes varies quite differently, prototype B mirrors many current contact tracing apps where control of the app tracing and data use is divided and/or nested across the app. Prototype A's control page details the 'Explain Yourself' message followed by a clear link to the privacy page and lastly, all the data controls in one place.

Figure 28. Privacy policy pages comparison

The main difference between the two prototypes' privacy policies is that prototype B's, accessed through the 'settings' page, is a *'huge wall of text'* as one participant described the Irish COVID tracker app's as while prototype A's, accessed from the 'control' page is segmented into four sections of: Summary, Data Used, Contact Info and Full Policy, aiming to achieve the 'Accessible Privacy Policy' principle.

Figure 29. Profile & Social pages

The final pages to be mentioned in the digital prototypes are the addition of a profile page and social page, as these additions are directly attempting to achieve the 'Engage Your Audience' principle, there is no similar page on prototype B. From the profile page users have the option to add contact information, add friends and customize their avatar. During the design stage, it was decided to use avatars in the app instead of the option to upload a profile picture, although profile pictures are more personal and more recognizable, it was thought that using avatars further limits the use of personal data while still allowing customization.

4.3.5 Pilot Testing and Final Iterations

During pilot tests 4 participants were asked to complete a series of tasks which would be used in the final testing phase. Following this, they were asked to complete the post-task survey to be used in the final testing.

From these tests several minor issues with the prototype and the phrasing of questions were identified, for example when asked to *'Find out what 'Request Data' is and then request your data'* all participants clicked 'Request Data' before clicking the 'What is this?' link.

Figure 30. Original 'Request Data' design

This issue was solved by changing the 'Request Data' information to be displayed whenever either button was selected.

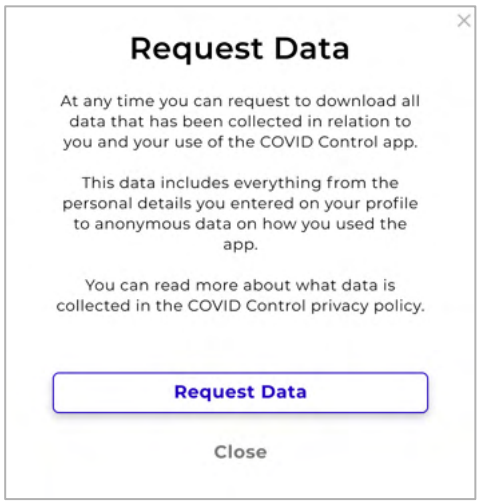

Figure 31. Updated 'Request Data' design

Figure 32. Onboarding screens

One participant stated that their trust in the app was suppressed by the fact that they couldn't see who the developer of the app was. The information they desired had been present in the 'Contact Info' tab of the privacy policy. However, it was clear that nesting that information in the app caused reduced trust in users.

This finding, paired with all participants misunderstanding the purpose of at least one feature in the app, highlighted the need for an onboarding experience in prototype A.

Following this, four onboarding screens were developed. The process introduces users to the app and app developer before taking them through the main components of the app. Page two details the functionality of the app as well as the value of using the app strengthening the 'Affirm Usefulness' principle. Page three iterates that the app requires personal data to function but emphasizes that users can easily see and control what data is used, supporting the 'Explain Yourself' principle. The final page introduces users to the social features of the app by briefly explaining badges, leaderboards, and their Control Score.

4.4 Testing

Testing of prototypes was done in two stages. Firstly, a quantitative A/B between-subject testing of the two prototypes where participants carried out a series of tasks with one of the prototypes and completed a survey based their experience with the prototype. The second stage of testing was a qualitative within-subject testing of both prototypes where participants were asked to complete a series of tasks with prototype A, answer questions on their experience, complete similar tasks with prototype B and answer the similar questions based on that experience.

Participants for the quantitative test were recruited through online forums, social media, friends, and family, half of the participants used prototype A and half using prototype B. Prototype test was conducted using Maze and linked to the survey conducted using Microsoft Forums.

Participants from the exploratory research phase who accessed to be part of further testing were recruited for the qualitative tests. Participants were asked to use the Figma prototype walkthrough to complete the tasks, user observations were carried out as the participants completed the tasks and, on competition, user interviews were conducted assessing their experience with both prototypes.

5. RESULTS

5.1 Quantitative Results

From the first stage of testing, 45 participants completed the questionnaire, 21 of which had used Prototype A and 24 using Prototype B. 20 participants identified as female with 25 identifying as male, the age range of participants was 18- 65 with over half being in the range of 25-34 years old.

All results gathered as part of the survey were done through 7-point Likert scales except for the System Usability Scale which was done through a 5-point Likert scale.

The results of the Satisfaction, Usability, and Intent to Use scales and two Trust sub scales which directly assessed the participant's opinions of the prototype were statistically assessed to measure the results' reliability, normality, and significance.

5.1.1 Reliability and normality

A Cronbach's alpha reliability test was run on the scale results to test their reliability. All results had a Cronbach's alpha higher than 0.70 meaning all data could be deemed reliable. Result of tests are summarized below:

	Cronbach's alpha	N of items
Trust Belief Subscale	.959	10
Trust Intention Subscale	.950	13
Intent to Use Scale	.890	3
System Usability Scale	.892	10
Satisfaction Scale	.852	3

Table 2. Reliability scores

A Shapiro-Wilk Normality test was run on each of the sets of data to test their reliability. For the inputted data to be deemed normal, the data set is required to return a p => 0.05. As can be seen in the table below, the System Usability Scale data set was deemed normal, while the other four data sets were deemed not normal by the test.

	Statistic	df	Sig. (p value)
Trust Belief Subscale	.921	45	.005
Trust Intention Subscale	.931	45	.010
Intent to Use Scale	.919	45	.004
System Usability Scale	.958	45	.105
Satisfaction Scale	.900	45	.001

Table 3. Normality scores

5.1.2 Trust

Since the results of the Trusting Belief and Trusting Intentions subscales are part of the same scale and are measuring subsections of the main topic or trust, there is an increased chance of observing a rare event thus increasing the likelihood of returning significant results. To counteract this potential problem, the Bonferroni adjustment will be applied to the results of both Mann-Whitney U tests being run on each of these subscales where the adjusted significance score will equal *p* / 2.

The trust disposition of the participants was generally positive with the mean score for Benevolence, Integrity and Competency in relation to people in general were: 4.6, 4.7 and 4.8 respectively. However, in relation to Trust Stance, a measure of a participant's neutral trusting intentions towards strangers, the mean was significantly higher at 5.5, the distribution of which can be seen below:

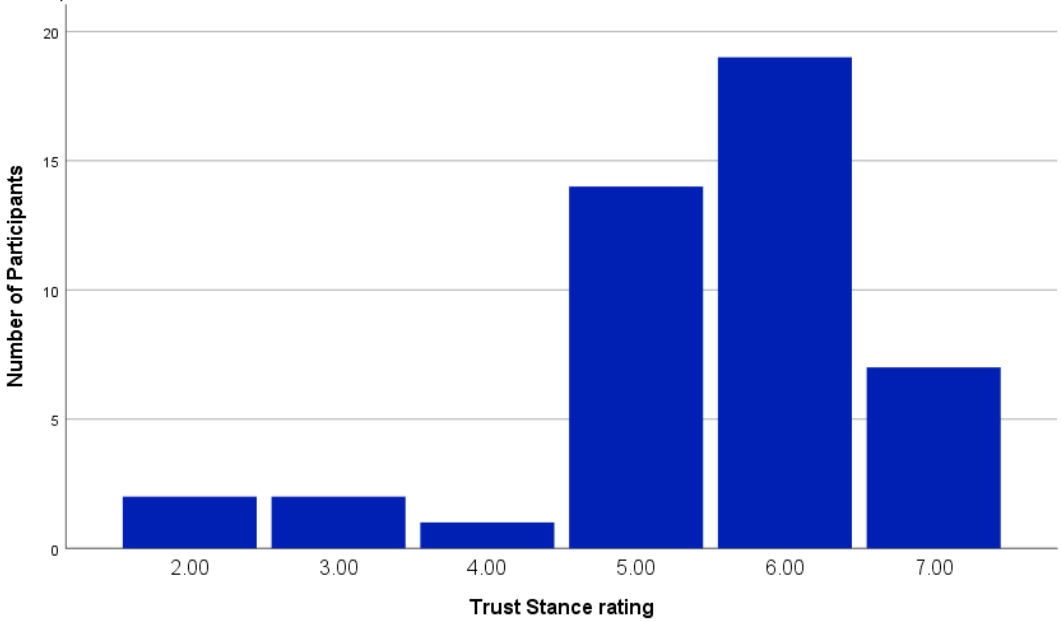

Figure 33. Trust Stance results

By contrast, the measurement of Structural Assurance, trust in the internet, was shown to be the lowest measurement in the trust disposition assessment with a mean of 4.2, the distribution of this was also more spread as can be seen below:

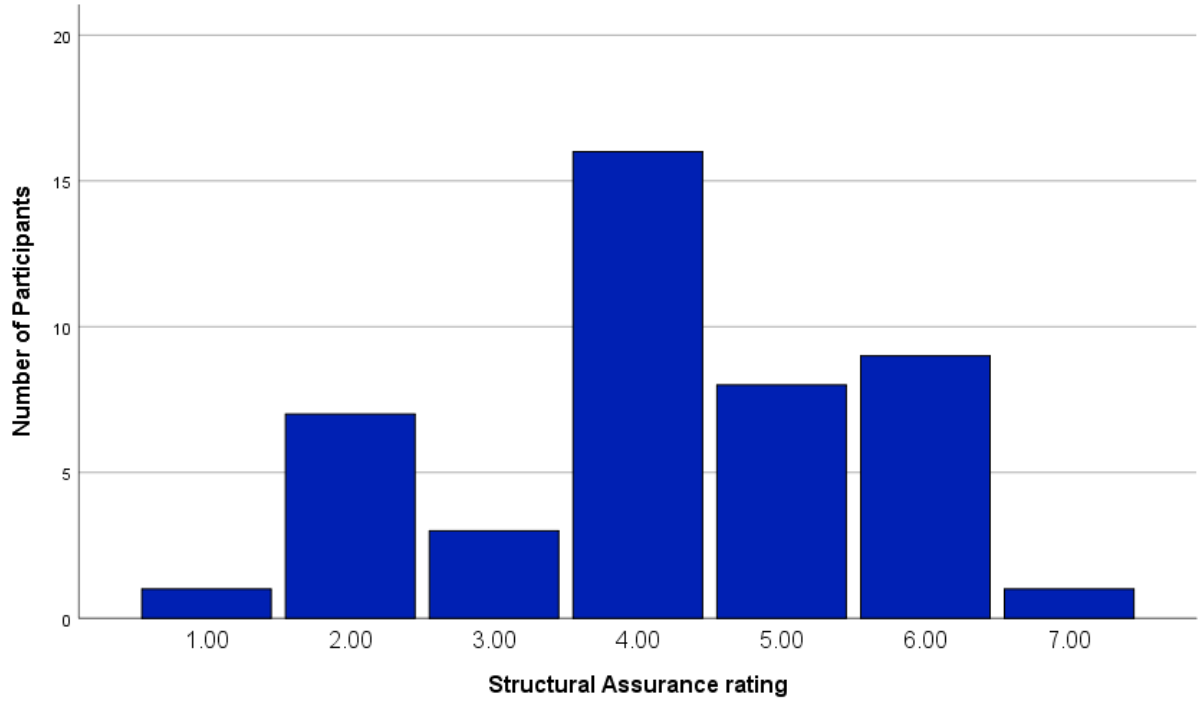

Figure 34. Structural Assurance results

Trusting Belief:

To test the significance of the Trusting Belief data set, a Mann-Whitney U test was run and showed that there was a significant difference (U = 134.5, p = 0.007) between Prototype A with Trust Focused Design Principles compared to Prototype B without. The mean rank was 28.60 for the Trust Focus Design compared to 18.10 suggesting that the implemented Trust Focused Design Principles were effective at increasing the user's Trusting Belief in the prototype. The difference can be viewed on the box plot below:

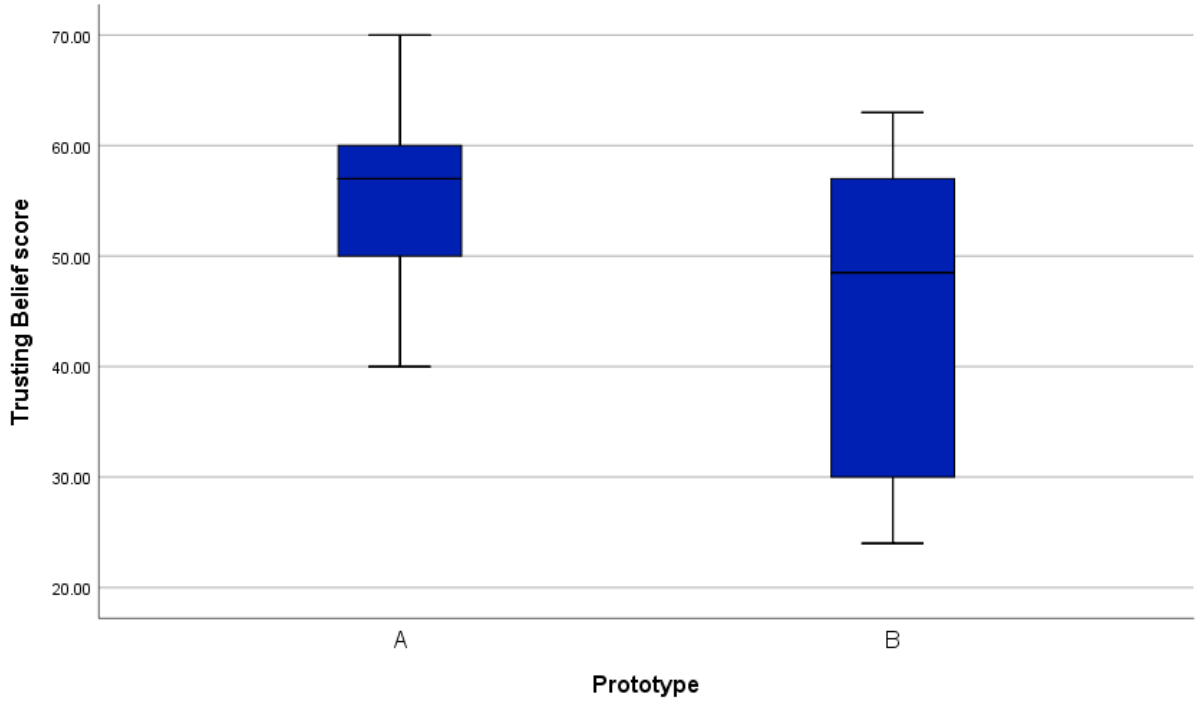

Figure 35. Trusting Belief comparison

Trusting Intention:

To test the significance of the Trusting Intention data set, a Mann-Whitney U test was run and showed that there was a significant difference (U = 109, p = 0.001) between Prototype A with Trust Focused Design Principles compared to Prototype B without. The mean rank was 25.88 for the Trust Focus Design compared to 20.48 suggesting that the implemented Trust Focused Design Principles were effective at increasing the user's Trusting Intention in the prototype. This difference is seen in the box plot below:

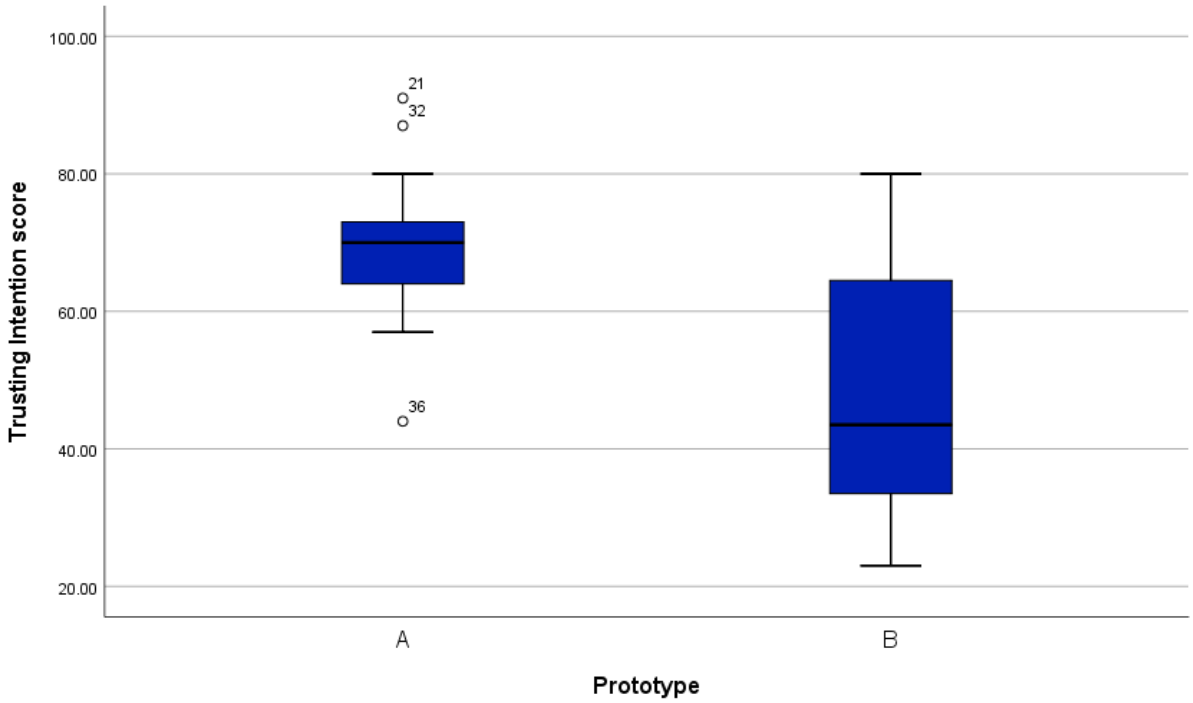

Figure 36. Trusting Intention comparison

With the results of both the Trusting Belief data and the Trusting Intention data showing significance, it can clearly be stated that the first null hypothesis of this paper, *'There will be no significant difference in the level of trust between an app with trust focused design and an app without'* can be rejected, and the alternative hypothesis, that there is a significant difference, can be accepted.

5.1.3 Satisfaction:

To test the significance of the Satisfaction data set, a Mann-Whitney U test was run and showed that there was no significant difference (U = 191.5, p = 0.164) between Prototype A with Trust Focused Design Principles compared to Prototype B without. The mean rank was 28.60 for the Trust Focus Design compared to 18.10 suggesting that the implemented Trust Focused Design Principles were not effective at increasing the user's Satisfaction with the prototype, this is the only scale which returned non-significant results. This is seen in the box plot below:

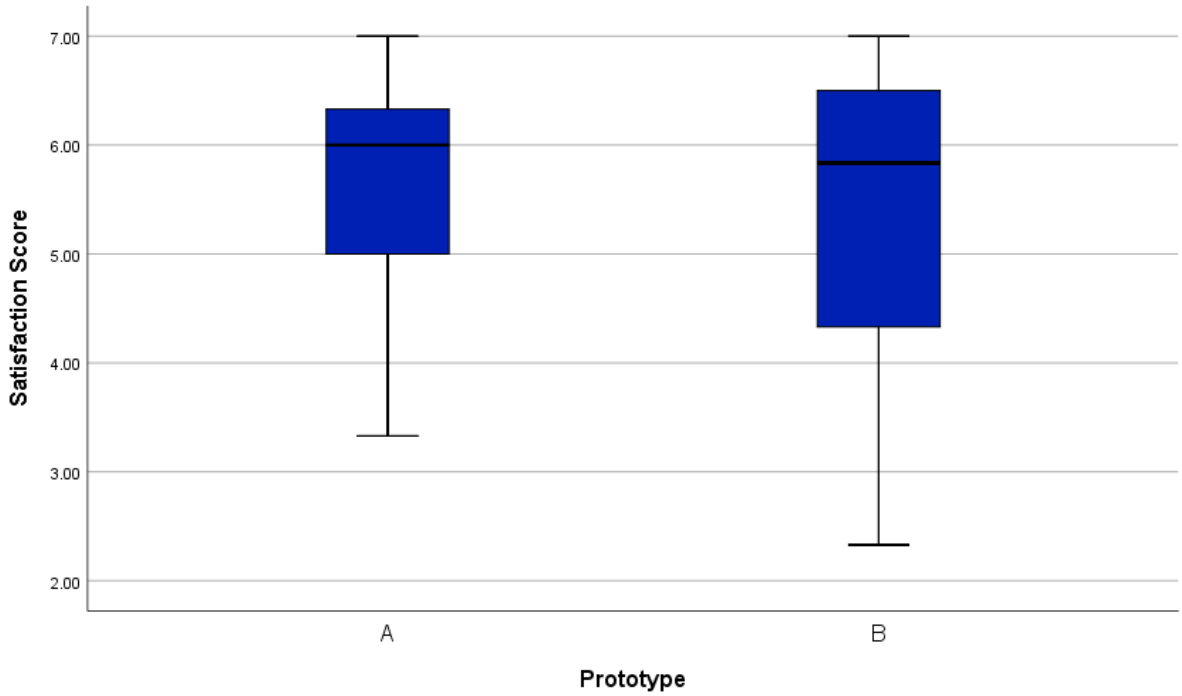

Figure 37. Satisfaction comparison

Using the results of the Mann-Whitney U test on the Satisfaction data set, the second null hypothesis of this paper, *'There will be no significant difference in the level of satisfaction between an app with trust focused design and an app without'*, has been retained as there was not enough of a difference between the two prototypes to be significant.

5.1.4 Intent to Use

To test the significance of the Intent to Use data set, a Mann-Whitney U test was run and showed that there was a significant difference (U = 119.5, p = 0.002) between Prototype A with Trust Focused Design Principles compared to Prototype B without. The mean rank was 29.31 for the Trust Focus Design compared to 17.47 suggesting that the implemented Trust Focused Design Principles were effective at increasing the user's Intent to Use the prototype in the future, this is seen in the box plot below:

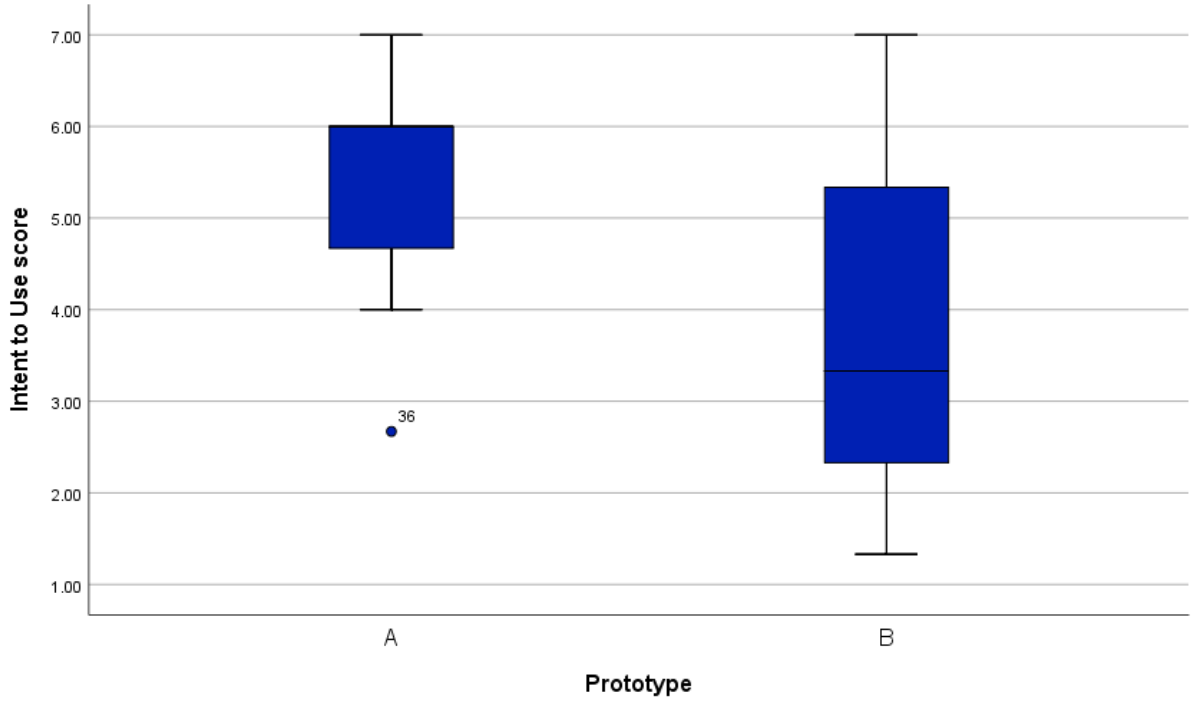

Figure 38. Intent to Use comparison

Using the results of the Mann-Whitney U test on the Intent to Use data set, the third null hypothesis of this paper, *'There will be no significant difference in the level of intent to use between an app with trust focused design and an app without'*, can be rejected, and the alternative hypothesis, that there is a significant difference, can be accepted.

5.1.5 Usability

Finally, to test the significance of the System Usability Scale data set, as it data was Normal, an Independent T-Test was run and showed that there was a significant difference ($t(42.017) = 2.368$, $p = 0.023$) between Prototype A with Trust Focused Design Principles compared to Prototype B without. The mean was 75.476 for the Trust Focused Design compared to 64.271 suggesting that the implemented Trust Focused Design Principles were effective at increasing the user's Perceived Usability of the prototype, this is seen in the box plot below:

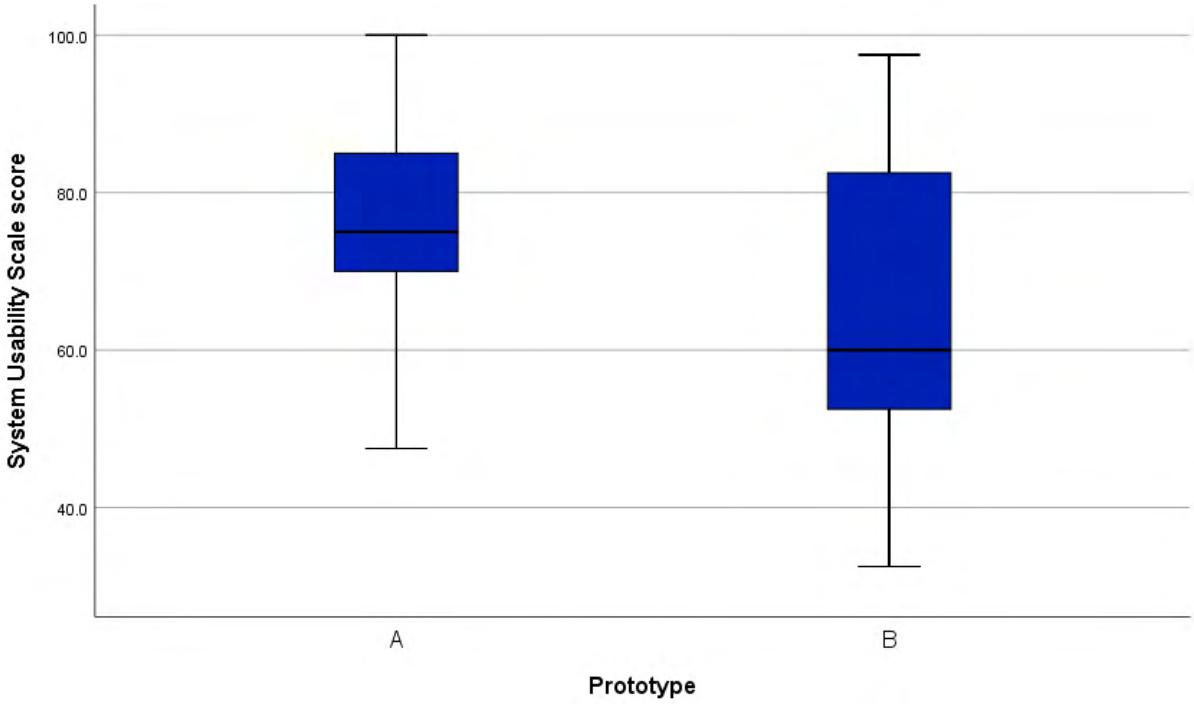

Figure 39. System Usability Scale comparison

Using the results of the Independent T-Test on the System Usability Scale data set, the final null hypothesis of this paper, *'There will be no significant difference in the level of perceived usability between an app with trust focused design and an app without'*, can be rejected, and the alternative hypothesis, that there is a significant difference, can be accepted.

5.2 Qualitative Results

For the qualitative testing, 5 participants were asked to conduct a series of tasks using both prototypes, A , followed by B. After each test, participants answered questions based on their experience with the prototype. Additionally, participants were asked questions in relation to comments made in the feedback section of the quantitative testing.

Below are tables highlighting key user comments. All the user's comments are coded based on themed that participants frequently brought up (Saldaña, Saldaña & Miles, 2013). Included in these are the 7 principles and specific features, such as the Onboarding, that were brought up by participants:

Prototype A

Topic	Comment	Tags
Did you find the app trustworthy?	I found the app very trustworthy; privacy of data is important to me. Found the app was respectful. (P.1)	Privacy
Did you find the app trustworthy?	Definitely trustworthy. Particularly the Onboarding and Control section helped me trust. App felt transparent. (P.5)	Data control, Explain Yourself, Onboarding
Did you find the app easy to use?	Very easy to navigate, clear layout, Onboarding explanations were helpful. (P.2)	Easy to Use, Onboarding
Did you find the prototype satisfying?	Yes, it was useful for what I needed; the feedback confirmation was reaffirming. (P.4)	Affirm Usefulness
Would you use the app?	I would over the current app, yes. Its clearer, I feel more in control and its more accessible (P.2)	Data control
Would you use the app?	Yes, it's more trustworthy and it makes me feel like I'm contributing. (P.5)	Trustworthy, Affirm Usefulness

Table 4. Prototype A qualitative test results

Prototype B

Topic	Comment	Tags
Did you find the app trustworthy?	Comparatively, no. It's fine, I'm just not comfortable with it. The introduction screens were important. (P.1)	Comfort, Onboarding
Did you find the app trustworthy?	Less so compared to prototype A, I'd trust it on its own, but its less organized and more layered. Turning off the different features were more spread out. (P.4)	Easy to use, Data control
Did you find the app easy to use?	Not as intuitive, more time consuming. On its own I might've just felt that I was about at using the app. (P.1)	Easy to use, User confidence
Did you find the app satisfying?	It's fine, I'd be less inclined to check it every day. (P.1)	Affirm Usefulness
Did you find the app satisfying?	No, no graphics, I didn't feel looked after while using the app. (P.2)	Look professional, User confidence
Would you use the app?	Probably not, I think using the first app first has made me less likely to use this one. (P.5)	Affirm Usefulness

Table 5. Prototype B qualitative test results

In the table below are the tags used in the coding of the qualitative data along with how many times that tag was highlighted in participants' comments. Usage of the tag in the context of Prototype B marks noted absence of the tag:

Tag	Prototype A Usage	Prototype B Usage
Privacy of data (Principle 1)	5	-
Information & Explanation (Principle 2)	4	2
Useful (Principle 3)	2	1
User Control (Principle 5)	2	1
Professional Look (Principle 6)	4	1
Easy to Use (Principle 7)	7	2
Trustworthy	3	-
Comfort	-	1
User Confidence	1	2

Table 6. Qualitative tests coded results

6. DISCUSSION

6.1 Quantitative Results

The results of the quantitative research show that the implementation of the 7 Trust Focused Design Principles had a significant increase in user trusting belief, trusting intentions, intent to use and perceived usability. The principles also showed an increase in satisfaction although not significant.

At the beginning of the design phase, it was decided to not consciously make prototype B absent of the 'Look Professional' and 'Easy to Use' principles. This was decided as both principles have been widely tested as part of other trust models (Fogg, 2001; Corritore, 2003; Koufaris, 2003; Pavlou, 2001) and because intentionally creating a prototype unprofessional and difficult to use would likely sabotage the assessment of the other principles.

Because of this choice it was thought that the tests for usability and satisfaction were less likely to be significant compared to the other tests. However, the usability difference between the two prototypes was significant and, while not significant, the overall satisfaction score of prototype A was higher than prototype B. As seen in the respective boxplots, while the range between the upper and lower quartiles is larger in the results of prototype B with the lower quartile being significantly lower compared to that of prototype A, which is consistent across the results of all scales, the upper quartiles for the usability and satisfaction results, are close to being the same, while for trust and intent, prototype A's upper quartile is notably higher. From this finding, not deliberately removing 'Professional Look' and 'Ease of Use' from prototype B increased the upper quartile, however, the influence of other factors contributing to achieving these principles significantly narrowed the range between the upper and lower quartiles in prototype A's usability and satisfaction results.

6.2 Qualitative Results

From the qualitative results, it was clear that participants' opinions of the prototype became a lot more polarizing and defined when comparing both against each other. Apart from further emphasizing the significance between the two prototypes, the qualitative results also begin to show which Trust Focused Design Principles had a larger impact on the participants' trust, perceived usability, satisfaction, and intent. Some of the reoccurring themes which were mentioned in comments include: Ease of Use, Data Control and Affirmation of Usefulness. Although these results don't prove that these themes were the most impactful, it does show that they are some of the most obvious to the participants, however further testing will be required on the individual principles.

Another finding from the qualitative testing was the significance of the onboarding experience for the participants. 3 of the 5 participants mentioned the Onboarding experience as a key factor which increased their trust and understanding of the app or commented that the absence of onboarding in prototype B damaged their perception of the app.

The qualitative results identified some important factors which should influence future work related to this study. Firstly, as mentioned, these results show that some of the Trust Focused Design Principles have a stronger influence on user perceptions compared to others and individual testing is something to be conducted in the future, this was seen by the fact that no participants mentioned any 'Engage Your Audience' features as influencing them, this could be due to the short length of testing and that this principle is intended to increase user retention over a period of time. Further long-term testing of this principle may be required. Secondly, the results have highlighted

the fact that user opinion is heightened when they compare both prototypes rather than conducting between-subject testing. Future testing should consider this finding when conducting quantitative tests.

6.3 Design Issues

A consideration that come from the results of the testing was the validity of the attempt to create a Difficult to Trust scenario. At the beginning of the project and during the design phase, it was thought that, as the majority of participants in the test would be living in Ireland, a relevant, real world Difficult to Trust scenario that many participants would have experienced would be with the current Irish COVID tracker app, as it was initially met with suspicion and reservations on release. It was thought that by basing the design off the current app that the initial feelings of suspicion and reservation could be emulated while using the prototypes.

However, during the pilot test one participants expressed their familiarity with the app upon first use, this was also seen in comments received after the quantitative tests stating that their familiarity with the current contact tracing app assisted in their navigation of the tasks.

After receiving this feedback, it was thought that the connection to the current app and the prototypes made through the design had, not only failed to create a Difficult to Trust scenario, but also inadvertently made both prototypes more familiar to many participants, which directly reduces the perceived risk of a system (Mayer, 1995; Romita, 2001).

To test if this was the case, as part of the qualitative interviews, participants were asked *'Has using the Irish COVID tracker app made this prototype more familiar?'* to which three participants thought their prior experience helped make prototype A slightly familiar with one participant stating:

'Maybe. They're similar, but I find this one more user friendly'

-Participant 2

and two participants stated that prototype B would've been hard to use without prior experience with one stating:

'Yes, using the current app everyday made using this one a lot easier'

-Participant 3

Although prior experience with the Irish COVID tracker app helped participants with both prototypes, comments in relation to prototype A were phrased with the prototype being an upgrade to the current while comments relating to prototype B being phrased as the current app being a crutch to their prototype experience. It is a positive result however the fact remains that the context of prototype tests was not 'Difficult to Trust', in order to achieve this the designs would need to be recreated or the participants would need to not have a past experience of the Irish COVID tracker app.

7. CONCLUSIONS & FUTURE WORK

7.1 Summary

This paper details the development and evaluation of 7 Trust Focused Design Principles which aimed to increase a user's trust, usability, satisfaction, and intent to use of an application within a Difficult to Trust scenario. During the study, the principles were developed based on literature and primary research, implemented into prototype A, a contact tracing app and compared against prototype B, a contact tracing app without the principles, and tested using both quantitative and qualitative techniques. The results of these tests were gathered using a series of proven scales for separately assessing trust, usability, satisfaction, and intent to use. The study had a main research question of '*What is the impact of Trust Focused Design Principles when using a contact tracing app?*' and four null hypotheses for each of the variables being measured. The results of the study were that the null hypotheses for trust, usability and intent to use were rejected, showing that the Trust Principles did have a significant impact on the level of trust, usability and intent to use the app compared to an app without the principles while the satisfaction null hypothesis was retained showing the principles did not have a significant impact compared to an app without the principles.

7.2 Key Contributions

This study includes the creation of 7 Trust Focused Design Principles and shows that, when implemented in an app, are capable of significantly increasing the trust, perceived usability, and intent to use of a user compared to an app absent of the principles.

Strengths of the study include the data from the quantitative and qualitative testing of prototype A were very positive, the results from the Intent to Use scale were particularly promising. Assuming that prototype B is a fair representation of current COVID tracker apps, the significant increase of intent from participants using prototype A displays well over 60% of participants positively intending to using a future version of the app. This result would cover the minimum population percentage for effective smartphone contact tracing (Hernandez-Orallo et al., 2020). If the results are representative, the changes to the current COVID tracker app would increase the public use to a level where the contact tracing would be able to contribute to the reduction of infection spread.

7.3 Limitations

The main limitation of this study was, due to the parameters of the study the 7 Trust Focused Design Principles had to be tested as one unit, not individually and as a result the individual effectiveness and contributions of each of the individual principles could not be tested or defined. Additional research would be required to test the effectiveness of the individual principles as well as the effectiveness of the implementation of these principles.

Another limitation of the study was the ineffective implementation of a 'Difficult to Trust' scenario. While this study used Trust Focused Design Principles in a contact tracing app the principles were created for use in any 'Difficult to Trust' scenario. However, the attempts of creating a 'Difficult to Trust' scenario by replicating a similar experience to that of the Irish COVID tracker app on release, were not effective. Although using the Irish COVID tracker app at the beginning was difficult to trust for many users, at the time of testing the prototype, the Irish app was familiar to many which counteracted this intended design.

Finally, the sample size of this study, particularly the quantitative results, is a limitation of the study. In future work, testing should be carried out with a larger sample size.

7.4 Future Research

The first, and most important, piece of future research will need to assess the validity of each of the 7 Trust Focused Design Principles individually as well as testing how to effectively implement each individual principle. Although the implementation of the principles as a group has shown to be significantly effective, the impact of each individual principle should be proven.

Further attempts to test the principles in a 'Difficult to Trust' scenario should also be carried. Part of the aim of this study was to create the principles for use in 'Difficult to Trust' scenarios. However the results of this study have not proven these principle's effective in such a scenario. Improved attempts to create a scenario where participants are reluctant to trust in should be part of future research.

Finally, the 'Trust Focused Design Principles' should be used within a different context other than a contact tracing app, to strengthen their validity.If the 'Trust Focused Design Principles' are applicable in a variety of scenarios where users are distrustful of the system prior to installing, they will be applicable for use in multiple contexts including; banking, navigation, exercise , medical, GPS and other similarly data sensitive apps.

8 REFERENCES

Abeler, J., Altmann, S., Milsom, L. et al. (2020). Support in the UK for app-based contact tracing of COVID-19.

Al-Shorbaji, N. et al. (2020). Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing.

Adam Mahmood, M., Burn, J., Gemoets, L., & Jacquez, C. (2000). Variables affecting information technology end-user satisfaction: a meta-analysis of the empirical literature. *International Journal Of Human-Computer Studies*, *52*(4), 751-771. doi: 10.1006/ijhc.1999.0353

Bevan, N., Kirakowski, & Maissel, J. (1991). What is usability? In Proceedings of the Fourth International Conference on Human Computer Interaction (pp. 651-655). Stuttgart, Germany: ElsevierBeldad, 2018)

Blasimme, A., & Vayena, E. (2020, November 13). What's next for COVID-19 apps? Governance and oversight. Retrieved January 11, 2021, from https://science.sciencemag.org/content/370/6518/760

Broken promises: How Singapore lost trust on contact tracing privacy. (2021). Retrieved 8 February 2021, from https://www.technologyreview.com/2021/01/11/1016004/singapore-tracetogether-contact-tracing-police/?truid=186e6760f5a7dccd005b0dffdde4d9e5&utm_source=the_download&utm_medium=email&utm_campaign=the_download.unpaid.engagement&utm_term=Active%20Qualified&utm_content=01-25-2021&mc_cid=9acbda47ef

Brooke, J. (1996). SUS: a „quick and dirty" usability scale. In P.W.Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.) Usability Evaluation in Industry (189-194). London: Taylor and Francis.

Chodosh, Joshua & Morton, Sally & Mojica, Walter & Maglione, Margaret & Booth, Marika & Hilton, Lara & Rhodes, Shannon & Shekelle, Paul. (2005). Meta-Analysis: Chronic Disease Self-Management Programs for Older Adults. Annals of internal medicine. 143. 427-38.

Corritore, C., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. International Journal Of Human-Computer Studies, 58(6), 737-758. doi: 10.1016/s1071-5819(03)00041-7

Corritore, C.L., Marble, R., Kracher, B., & Wiedenbeck, S. (2005). Measuring Online Trust of Websites: Credibility, Perceived Ease of Use, and Risk. In Proceedings of the 11th Americas Conference on Information Systems (2005), Omaha, NE, USA, 2419- 2427.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. - MIS Quarterly: Management Information Systems

Denise Christine Rieser and Orlando Bernhard. (2016). Measuring Trust: The Simpler the Better? In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16), 2940--2946.

Ditsa, G. and MacGregor, G. E. (1995). "Models of User Perceptions, Expectations and Acceptance of Information Systems", Sixth International Conference of the Information Resource Management Association, pp. 221-229.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., & Abeler-Dörner, L. et al. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science, 368(6491), eabb6936. doi: 10.1126/science.abb6936

Fogg, BJ., Marshall, J., Laraki, O., Osipovich, A., Varma, C., & Fang, N. et al. (2001). What Makes Web Sites Credible? A Report on a Large Quantitative Study.

Gefen, D. (2000). E-commerce: The role of familiarity and trust. Omega 28(5) 725–737.

Gefen, D., E. Karahanna, and D. Straub (2003) "Trust and TAM in Online Shopping: An Integrated Model," MIS Quarterly (27) 1, March, pp. 51-90.

Gillmor, K. Daniel, B. (2020). Principles for Technology-Assisted Contact-Tracing [White paper]

Hernandez-Orallo, E., Manzoni, P., Calafate, C., & Cano, J. (2020). Evaluating How Smartphone Contact Tracing Technology Can Reduce the Spread of Infectious Diseases: The Case of COVID-19. *IEEE Access*, *8*, 99083-99097. doi: 10.1109/access.2020.2998042

Kim, D., Ferrin, D., & Rao, H. (2009). Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration. *Information Systems Research*, *20*(2), 237-257. doi: 10.1287/isre.1080.0188

Ko, H., Leitner, J., Kim, E., & Jeong, J. (2017). Structure and enforcement of data privacy law in South Korea. International Data Privacy Law, 7(2), 100-114. doi: 10.1093/idpl/ipx004

Koufaris, M. and Hampton-Sosa, W. (2004). 'The development of initial trust in an online company by new customers', Information & Management, 41(3), 377-397.

Lewis, J. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal Of Human-Computer Interaction*, *7*(1),57-78. doi: 10.1080/10447319509526110

Lewis, J. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies. *ACM SIGCHI Bulletin*, *23*(1), 78-81. doi: 10.1145/122672.122692

Lewis, T., & Wyatt, J. (2014). mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use. *Journal Of Medical Internet Research*, *16*(9), e210. doi: 10.2196/jmir.3133

Mayer, R., Davis, J., & Schoorman, F. (1995). An Integrative Model of Organizational Trust. The Academy of Management Review, 20(3), 709-734. Retrieved December 8, 2020, from http://www.jstor.org/stable/258792

McClain, C. (2020, November 13). Key findings about Americans' views on COVID-19 contact tracing. Retrieved January 11, 2021, from https://www.pewresearch.org/fact-tank/2020/10/30/key-findings-about-americans-views-on-covid-19-contact-tracing/

McKnight, D. Chervany, N. (2000). What is Trust ? A Conceptual Analysis and an Interdisciplinary Model. - Proceedings of the 2000 Americas Conference on Information Systems AMCI2000 AIS Long Beach CA August 2000

McKnight, D. Chervany, N. (2002). What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology. - International Journal of Electronic Commerce, 6, 35-60.

Ministry of Public Administration and Security (2020, March 5). Self-quarantine safety protection app. Retrieved from https://www.mois.go.kr/frt/bbs/type002/commonSelectBoardArticle.do;jsessionid=7bA+UtY0JOIXJytznXoyYNHR.node40?bbsId=BBSMSTR_000000000205&nttId=76155

Misztal, B. (1996). *Trust in modern societies.* Cambridge, England: Polity Press.

Nass, C., Fogg, B., & Moon, Y. (1996). Can computers be teammates?. *International Journal Of Human-Computer Studies*, *45*(6), 669-678. doi: 10.1006/ijhc.1996.0073

Nass, Clifford & Steuer, Jonathan & Siminoff, Ellen. (1994). Computer are social actors. Conference on Human Factors in Computing Systems - Proceedings. 204. 10.1145/259963.260288.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. Proc. ACMCHI'90 Conf. (Seattle, WA, April 1- 5), 249-256.

Nunnally, J. (1975). Psychometric Theory. 25 Years Ago and Now. *Educational Researcher*, *4*(10), 7. doi: 10.2307/1175619

O'Callaghan et al. (2020). A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. Research Square Preprint posted online on July 10, 2020. [FREE Full text] [doi: 10.21203/rs.3.rs-40778/v1]

Pavlou, P., & Fygenson, M. (2006). Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior. MIS Quarterly, 30(1), 115-143. doi:10.2307/25148720

P.A. Pavlou, What drives electronic commerce? A theory of
planned behavior perspective, 2001

P. Pavlou, D. Gefen, Building effective online marketplaces with institution-based trust, Information Systems Research 15 (1) (2004) 37–59.

Romita, T. “Familiarity Breeds Net Sales,” Business 2.0, February 6, 2001, p. 98. Smith, J. H., S. J. Milberg and S. J. Burke "Information Privacy: Measuring Individuals' Concerns About Corporate Practices," MIS Quarterly (20:2), 1996, pp. 167-196.

Rotter, J. (1967). A new scale for the measurement of interpersonal trust.
- Journal of Personality.

Saldaña, J., Saldaña, J., & Miles, M. (2013). *The coding manual for qualitative researchers + qualitative data analysis*. London: Sage Publications.

Schlosser, A.E., White, T.B., & Lloyd, S.M. (2006). Converting web site visitors into buyers: How web site investment increases consumer trusting beliefs and online purchase intentions. Journal of Marketing, 70(2), 133–148

Stanford d.school. (2021). Retrieved 30 April 2021, from https://dschool.stanford.edu/

Stoyanov, S., Hides, L., Kavanagh, D., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. JMIR Mhealth And Uhealth, 3(1), e27. doi: 10.2196/mhealth.3422

The Government of the Republic of Korea. (2020). Tackling COVID-19, Health, quarantine and economic measures: Korea experience.

TraceTogether. (2021). Retrieved 8 February 2021, from https://www.tracetogether.gov.sg/

Wong, L. (2020, November 12). TraceTogether check-ins to be compulsory at public venues in S'pore by end-December. Retrieved January 11, 2021, from https://www.straitstimes.com/singapore/checking-in-with-tracetogether-to-be-compulsory-at-public-venues-by-december

World Leaders in Research-Based User Experience. (n.d.). Trustworthiness in Web Design: 4 Credibility Factors. Retrieved December 08, 2020, from https://www.nngroup.com/articles/trustworthy-design/

Wu, Y., Kankanhalli, A. and Huang, K.-W. (2015). “Gamification in Fitness Apps: How Do Leaderboards Influence Exercise?” In: International Conference on Information Systems, Fort Worth, USA.