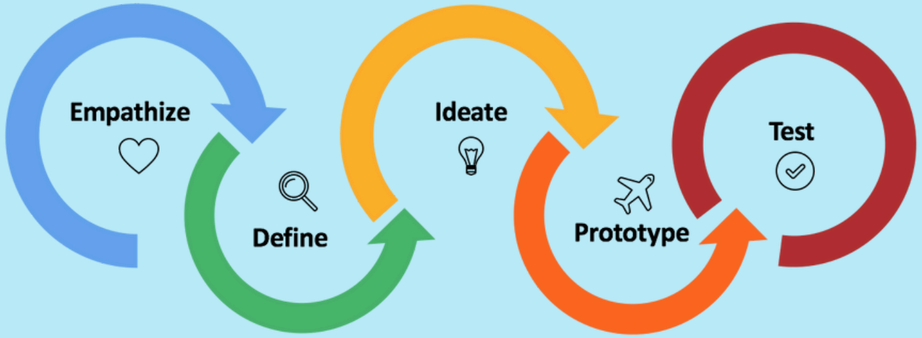
Human
Computer
Interaction

Group 33

DESIGN CHALLENGE

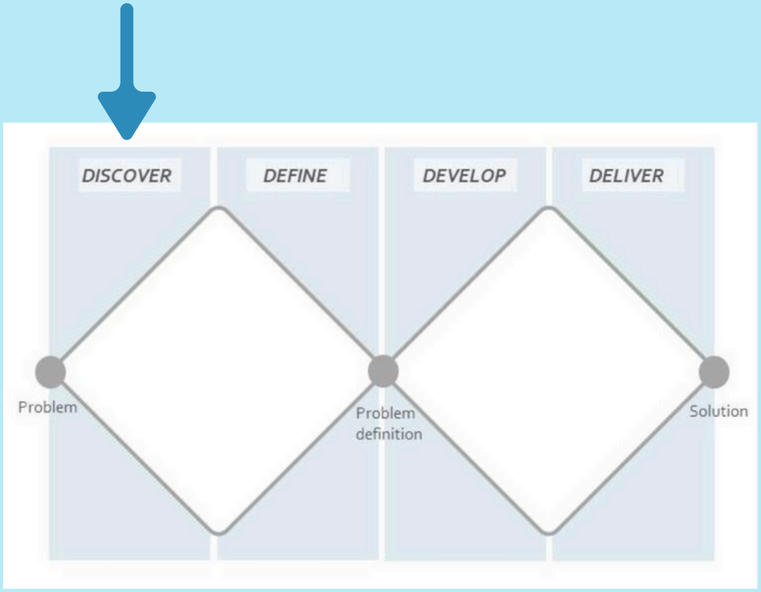
Create something to help people address a problematic behaviour

Ilya Sullivan, Danny Wayles, Asia Petrie, Stephanie Franks, Ava Boomla



Empathise

Understanding users

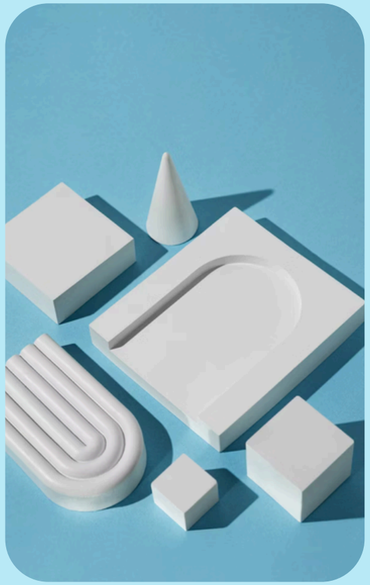


Breaking Down the Challenge

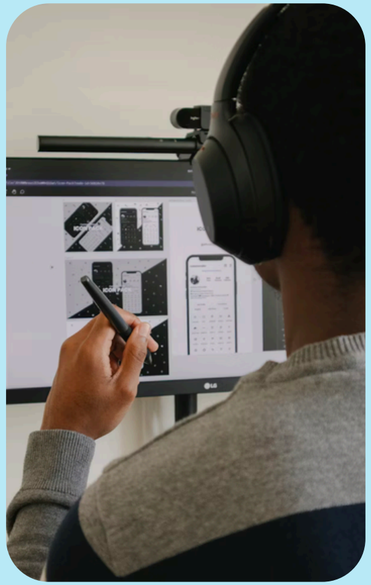
We began our process by identifying a range of categories of **problematic behaviours that impact everyday life**.
We were able to narrow down to a single category through group discussion and reasoning.

Health and wellbeing	**Chosen due to wide scope, applicability across many demographics and because it sparked the most enthusiasm within the group.**
Education and study	Ruled out as we couldn't see as wide scope of problematic behaviours in this category.
Environment and sustainability	Ruled out as a close second to health and wellbeing, with slightly less scope.
Work and productivity	Ruled out as we were less inspired initally by behaviours within this category.
Finance	Ruled out as we thought this could be included within health and wellbeing.

Intervention Type



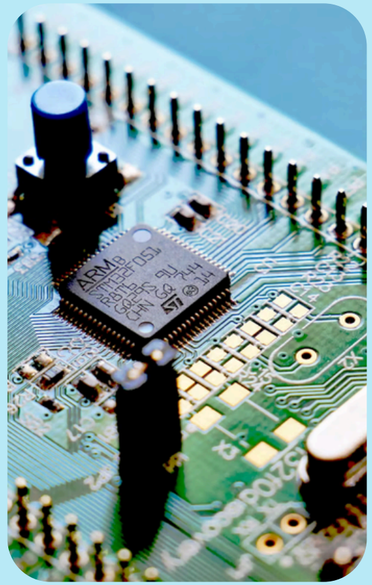
Physical Product



Digital Product

Policy

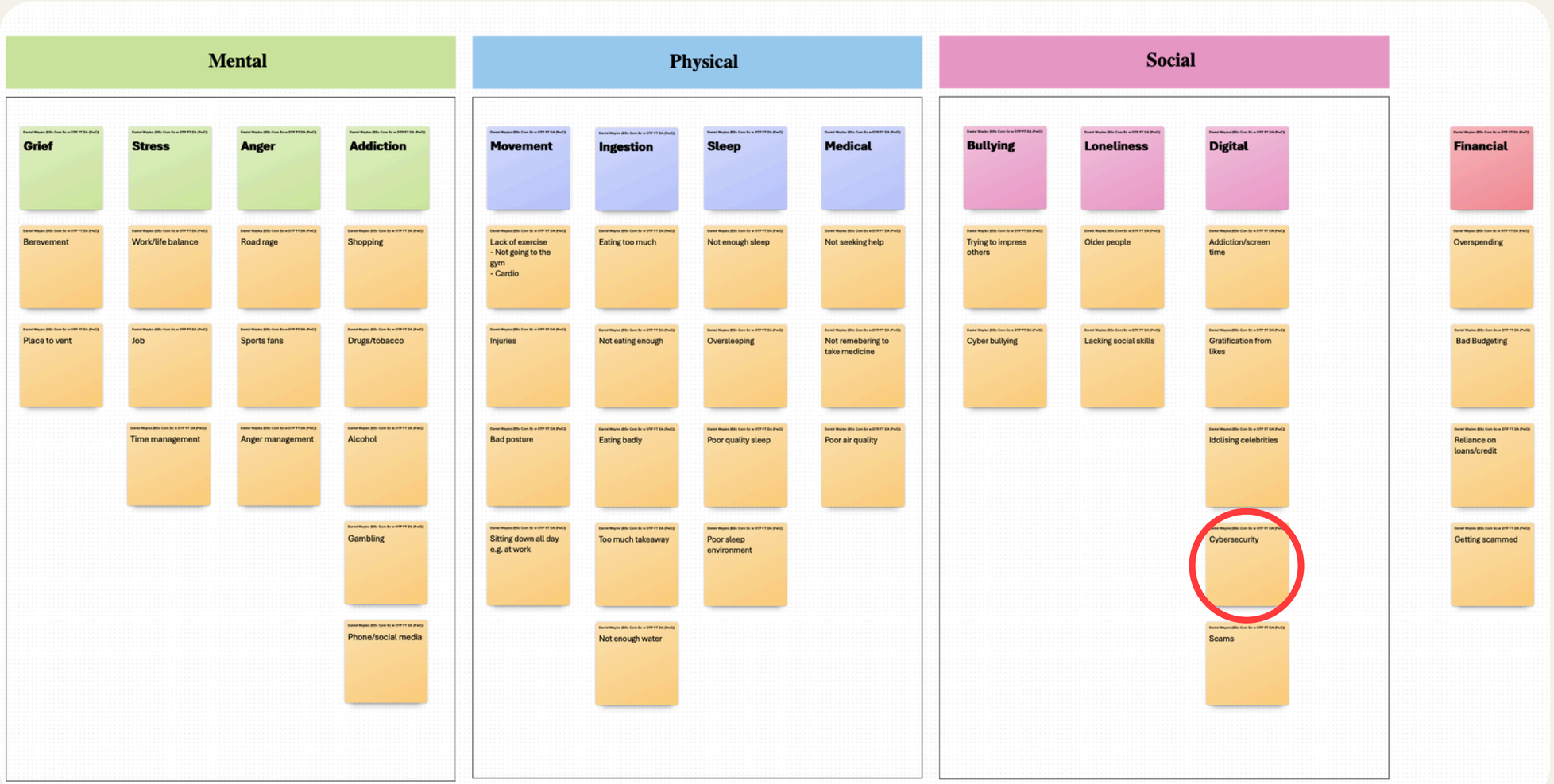
Program



Embedded System

We explored various forms of interventions to better understand which type would be **most meaningful to create change**. We **ruled out policy** as it would be harder to implement in a creative manner. We have left all other intervention types open with the **potential to combine multiple**.

Brainstorming



We **ruled out** mental behaviours such as addiction as they would be **sensitive subjects,** which could pose **challenging** when trying to **elicit responses**.

We **ruled out** these physical behaviours as we thought **a lot of existing products** came to mind that provided solutions when we thought about them.

We thought **digital and financial** wellbeing were two **interesting** categories within health and wellbeing, especially financial as we **struggled to categorise it. These two gave** us lots of **unique undesirable behaviours.**

We used a **Microsoft Whiteboard** to **brainstorm** undesirable behaviours together, then grouped them together to see any overlap in topics.

We landed on **lacking awareness of cybersecurity as our overarching theme** as our behaviour to target because it links **digital and financial wellbeing**, alongside being a prevalent and ever emerging issue.

SWOT Analysis of Themes

We came up with **4 undesireable behaviours** within cybersecurity and **conducted SWOT analysis** to understand which provides the **best scope** for creating a design to change the behaviour in a **meaningful way.**

Undesirable Behaviour within Cybersecurity	Strengths	Weaknesses	Opportunities	Threats
Forgetting passwords	• Universal issue • Automation increases usability • Encourages secure passwords	• Single point of failure • Syncing across devices is hard • User distrust	• Integration with browsers • Growing awareness of risks	• Heavily competitive landscape • Security breach risk • OS restrictions
Falling victim to scams	• Very relevant, high demand and concerning topic • Saves users from monetary loss • Lots of areas to explore within	• False-positive/negative risk on detecting scams • Requires frequent updates	• Large scope for workplace use • Prevent and pre-empt scams • Detect social engineering patterns	• Arms race of product vs scammers • Dealing with sensitive information
Losing personal documents or data	• Multiple avenues from data loss in the workplace or not backing up data personally	• Relies on individual to be motivated, hard to intervene • Distrust of saving confidential documents on an external app	• Big appeal to the corporate market, specifically with importance of confidentiality	• Data breaches • Competitive landscape within cloud file storage and sharing
Not keeping software up to date	• Reduces chance of being exposed to hackers • Could work in background	• May be hard to change actual user behaviour rather than working automatically • OS integration may be hard	• Using APIs to update different apps • Useful for businesses to check staff update status	• Highly dependent on other third parties • Forcing updates could cause issues in saving data

We concluded that **falling victim to scams** posed the **most opportunities** and strengths with the least downsides. The threat of being in an arms race with scammers could be **remedied by future proofing the design** or using updates. It is a topic that is **less explored, yet is a widespread behaviour.**

Potential Target Audience

Category	Examples	Reasons for consideration/exclusion
Age Groups	Children Teenagers Students Adults Elders	• Children were excluded due to **privacy and consent** concerns when handling sensitive data. • Students are highly **active online** with good **technical skills** but vulnerable to misinformation and data leaks. • Adults are a primary focus with frequent professional & personal use making them at risk for **phishing & financial fraud**. • Cognitive decline such as working memory contributes to **increased susceptibility** to deception and financial exploitation among elders (Ebner, 2023).
Accessibility & Inclusion	Visual Auditory Physical Cognitive Speech	• Accessibility is essential, scams target everyone regardless of **age** or **ability**. • Barriers such as **poor readability**, **unclear alerts** or reliance on visual/auditory cues increase the likelihood of falling victim to scams. • Design needs to be **inclusive**, ensuring that users with visual, auditory or cognitive impairments can recognise scam indicators as well.
Cultural & Location Contexts	Different countries & languages Settings, home, work, online	• Scams occur across home, work and online spaces and user **behaviour shifts between these environments.** Individuals are often **less protected out of the workplace** even with the same threat. • As we are in the UK, we would keep app data storage in the UK to apply to data laws including GDPR.

Ebner, N. C., Pehlivanoglu, D., & Shoenfelt, A. (2023). Financial Fraud and Deception in Aging. Advances in geriatric medicine and research, 5(3), e230007. https://doi.org/10.20900/agmr20230007

Brainstorming Evaluation

Purpose

We brainstormed types of problems to **narrow down the domain**. It allowed us to explore multiple categories of problematic behaviour across mental, physical, social and financial wellbeing. Using Microsoft Whiteboard enabled **rapid idea generation** and **grouping**. This step ensured we **did not prematurely narrow down** the problem.

Contribution

The process highlighted digital and financial wellbeing as the most interesting and unique area to explore. **SWOT analysis** revealed **"falling victim to scams"** as **relevant** and necessary. By **categorising interventions** and **audiences**, we stayed **aligned** and **inclusive**, knowing as a group which interventions we wanted/ did not want to build.

Learning

Many undesirable behaviours were **too sensitive** or **many solutions** already existed. Scams are relevant across groups and we learnt **scam susceptibility** is heavily influenced by **context**, demographics and behaviour not purely technical gaps.

Effectiveness

Brainstorming allowed us to narrow down from a broad brief but also highlighted limitations. Some of our **choices** were **based** on our own **assumptions** rather than user input, and certain **areas** may have been **dismissed too quickly** as we **lacked input**. Having earlier user input could have **challenged** our assumptions.

Influence on next stage

Brainstorming has given us a direction, but the next steps will include proper **research** to understand the types of scams out there, **why** people fall for scams, **who is affected**, and the **solutions** available. This has shaped questions for our next step for background research.

Background Research

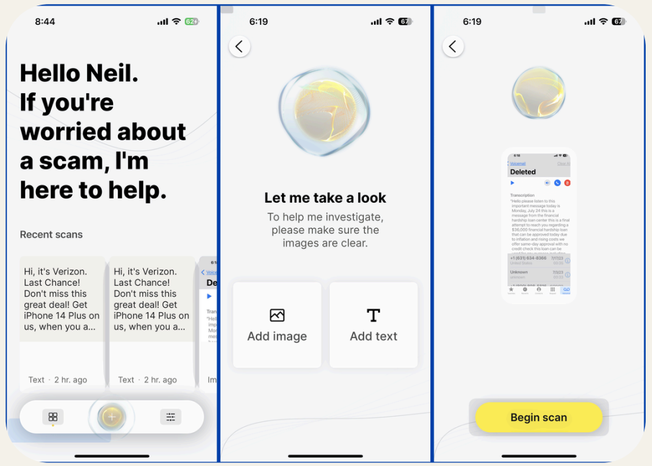
Existing Solutions Addressing Cybersecurity Scams

Norton Genie

An AI powered tool where users can upload a suspicious text/email/ URL and get a quick verdict and **advice on what to do next**.

- Fast AI verdicts
- Guidance on why content is unsafe
- Explains why a message is risky

- Users must upload items manually
- Reactive rather than preventative



Scam Guard by EE

Monitors spam calls and **scans the dark web** for exposed user data.

- Auto-blocks spam calls/texts
- Dark-web alerts for extra protection

- Limited beyond calls/SMS
- Only available for EE customers

PhishProtection

Real-time integration with databases of phishing sites, URL & attachment scanning for businesses.

- URL scanning and attachment sandboxing
- Integration with Microsoft 365 and Google Workspace

- Business focused, not personal
- Only provides protection for phishing

PhishProtection.

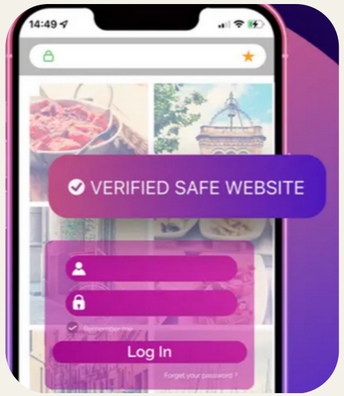
Background Research

Existing Solutions Addressing Cybersecurity Scams

NovoShield Anti-Phishing

Mobile extension that protects your phone from phishing sites and **detects fake login forms**.

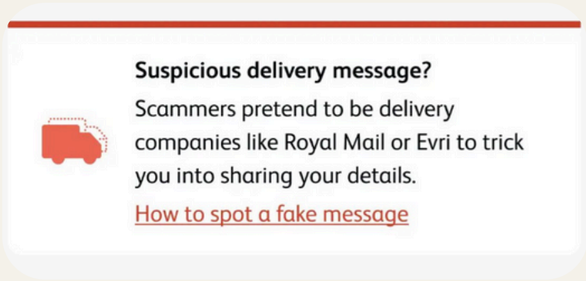
- Real-time browser protection
- Works across browsers
- Focused mainly on phishing



Banking Apps

Banking apps have **information about common scams**, for example: warning about delivery scams impersonating well-known companies.

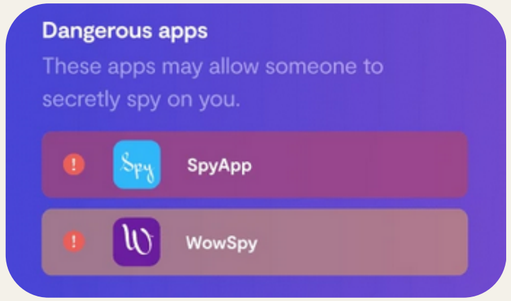
- Simple, accessible language
- Additional links providing more information and guidance.
- Informational only
- Users often ignore alerts when time pressured



Clario Anti Spy

Security app detecting spyware, hidden apps and leaked passwords. Find out what permissions each app has and **which could be spying on you**.

- Device-level protection from spyware
- 24/7 support for additional help
- Does not protect against specific scams, just a permission focused spyware tracker



Background Research

Building on Existing Solutions

Reviewing current apps revealed **clear strengths in detection** but apps often stop at surface-level intervention; warnings are provided but they **fail to change behaviour**.

Features done well	Explanation
Phising & Scam Detection	A lot of the apps had real-time detection of malicious links, websites and calls.
AI - Powered Threat Analysis	Some apps used AI and threat databases to identify scams and suspicious content.
Blocking & Filtering	Automatic blocking of spam content.
Data Privacy & Identity Alerts	Some apps include identity theft or data breech monitoring on the dark web.

Current apps protect users from scams, but **don't help users understand scams**.

Areas for development	Explanation
Reactive, not preventative	Most apps focus on one-time notifications e.g. "This message could be a scam"
Limited Personalisation	Most apps assume knowledge and skill, they should be adaptable to the persons skillset.
Poor Accessibility/ Readability	Apps may use jargon, or low colour contrast.
No reflection or follow-up	Current apps rarely include any follow-up or progress tracking. They do not measure how many scam attempts a user encounters, nor assess the user's individual risk level over time.
Static examples and outdated training	Whilst some apps use AI to detect fraudulent content, their detection models may rely on static or outdated indicators.
1-dimensional, single device	Few apps had cross-platform protection.

Background Research

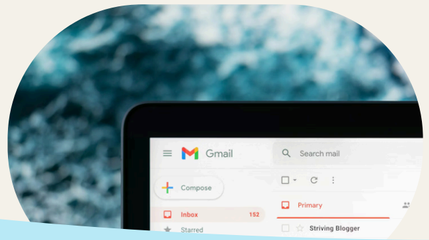
How Findings Shape Design

After identifying **gaps** in the existing solutions, we have outlines areas or **where to focus or improve design.**

Areas for development	Solutions
Reactive, not preventative	Simulate real-world scam scenarios to teach recognition early. Make sure users are learning.
Limited Personalisation	Tailor messages to each user's confidence and adapt learning to their needs.
Poor Accessibility/Readability	Use clear language, high colour contrast and add audio/visual feedback for diverse needs.
No reflection or follow-up	Encourage users to pause and think with reflective prompts. Measure user's risk and put in place interventions if a user is deemed more risky.
Static examples and outdated training	Systems need to be continuously updated with examples of current, real-world indicators. Keep training relevant and motivating.
1-dimensional, single device	Sync across multiple platforms and provide tracking across devices.

Background Research

Scams of the Present



Mid 1990s

Email Phishing

What it is: Fraudulent emails that pose as a trusted source but aim to extract personal information or money.

This type of scam has been ongoing for many decades but in recent years has been on the rise again, now combining with other scams, and becoming harder to detect and more personalised.

Early 2000s

SMS Phishing

What it is: Fraudulent SMS messages that pose as trusted source, linking to a fraudulent website in order to extract personal info and/or money.

During and post COVID, scammers have been acting as delivery companies, asking to reschedule or pay a fee for a parcel, exploiting the “right place, right time” situations.

2000s

Tech Support Scam

What it is: Popups on websites, phone calls or fake websites claim the computer is infected and need to contact support to remove it.

These scams pry on people‘s fear and wanting to avoid losing their data, willing to pay any amount to keep it. This scam is less effective today but they persist as it is cheap to run and target the most vulnerable.

Background Research

Scams of the Present

Mid 2010s

SIM Swapping

What it is: Criminals are able to trick mobile operators into transferring your phone number to a SIM they control to intercept SMS 2FA number.

This is a relatively new type of scam and is hard to prevent due to it being on the mobile operator to determine, and now with real-time deepfake voices tricks, more difficult to resist.

Early 2020s

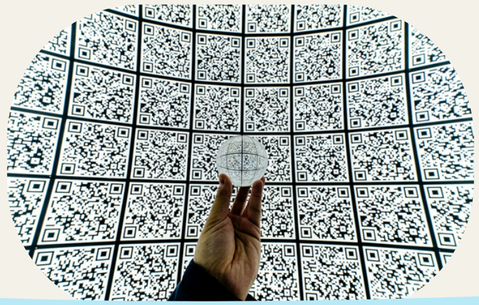
Crypto Scams

What it is: Fraudulent NFT and cryptocurrency projects advertise themselves, promising high returns and increase investment, then deliberately liquefy all assets they have in it and leaving the project.

Crypto scams take advantage of the newness of digital investing, advertising to anyone hoping to grow their money. They thrived during COVID, when many were unemployed and desperate for income. Today, with greater public awareness, scammers focus on elders and job hunters.

Background Research

Scams of the Present



2020s

Fake QR Codes

What it is: Malicious QR codes are placed on top of legitimate ones, acting like the real site but taking money and personal information out of it.

During and post COVID, scammers exploited this technique as things became more digitalised. QR codes obfuscate the link, making it harder to spot a fake site . As QR codes often appear in legitimate settings, people tend not to question them.

2020s

Deepfake Scams

What it is: Scammers make use of AI to synthesise a target's face or voice to impersonate them in order to do malicious acts.

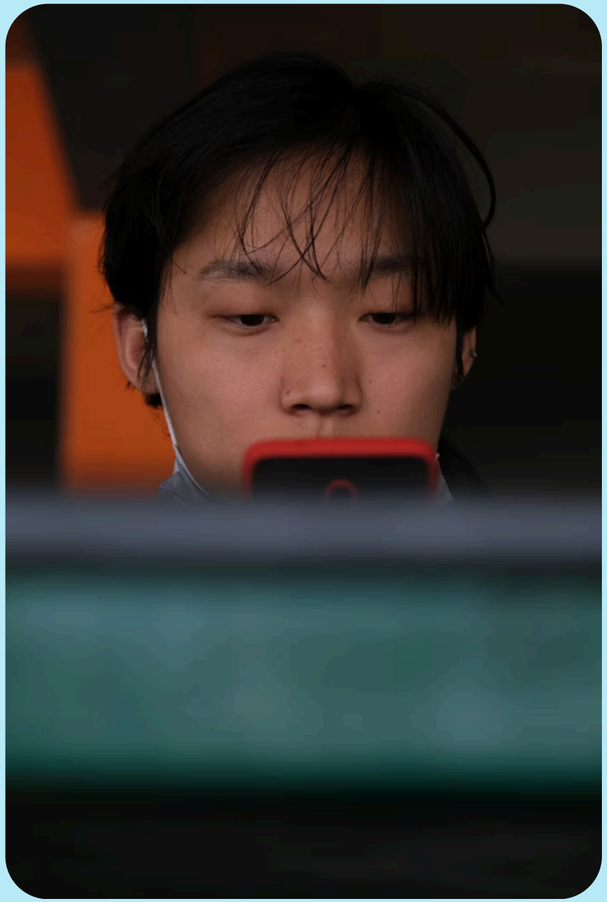
Deepfake scams are among the hardest to detect. With just a few seconds of audio or a single image, people's face and voice can be cloned. These can be harvested from social media profiles or by victims responding to malicious phone calls. From friends to families to bosses to celebrities, deepfakes can be everywhere in our lives without us knowing, this is the type of scam that is hardest to detect and easiest to slip through.

Background Research

Scams of the Present

Key learnings from scams of the present:

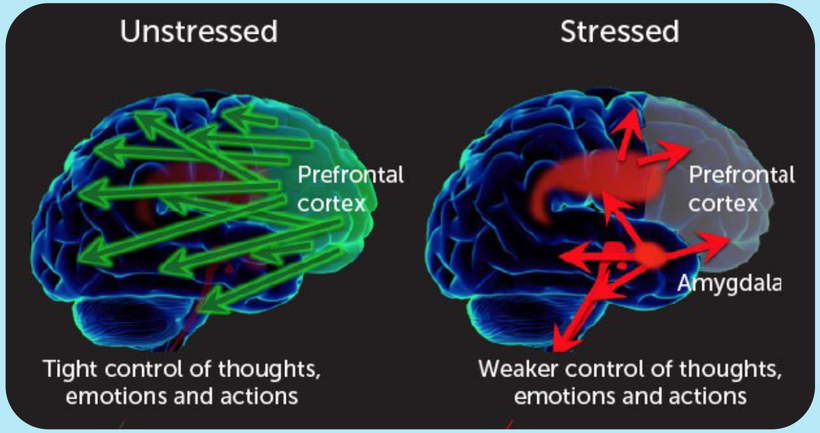
- Scams are everywhere, knowing real vs. fake is critical
- In the past 5 years, scams have risen exponentially
- AI makes scamming easier - we are one prompt away from the next threat
- Technology dependence has left everyone exposed, especially the elderly and vulnerable
- Staying alert and informed is the only defense



Kowalewski, M. et al. (2025) Scanned and Scammed: Insecurity by ObsQRity? Measuring User Susceptibility and Awareness of QR Code-Based Attacks [USENIX Security '25] Available at: usenix.org/conference/usenixsecurity25/presentation/kowalewsk
Muhly, F., Chizzonic, E. & Leo, P. (2025) 'AI-deepfake scams and the importance of a holistic communication security strategy', International Cybersecurity Law Review, 6, pp. 53–61. doi:10.1365/s43439-025-00143-7 [Online] Available at: https://link.springer.com/article/10.1365/s43439-025-00143-7
Myounghoon K, Joon S, Hunyeong K. (2022) A Study of the Emerging Trends in SIM Swapping Crime and Effective Countermeasures [Online] Available at: https://ieeexplore.ieee.org/abstract/document/9900510
Butt, U.A., Amin, R., Aldabbas, H., Mohan, S., Alouffi, B. & Ahmadian, A. (2022) 'Cloud-based email phishing attack using machine and deep learning algorithm', Complex & Intelligent Systems, 9, pp. 3043–3070 [Online] Available at: https://link.springer.com/article/10.1007/s40747-022-00760-3

Background Research

The Human Element: Why Humans Fall for Scams



Why do we react impulsively to scams?

- **Automatic threat response (fight or flight) -** when we are threatened, the brain's emotional centre the Amygdala can trigger a **rapid stress response** and releases adrenaline and cortisol. **Attention narrows** and reduces reflective thinking, causing people to act less rationally. Users are more likely to miss **red flags** or **skip cues** when rushed (Butavicius et al., 2022, BodyMindBrain, 2024).

- **Scammers exploit human heuristics -** users rely on fast **cognitive shortcuts**: trust, urgency, familiarity. Social engineering **mimics these cues** to trigger **automatic responses.** People rarely perform deep analysis unless something "feels off" (Desolda et al, 2021).

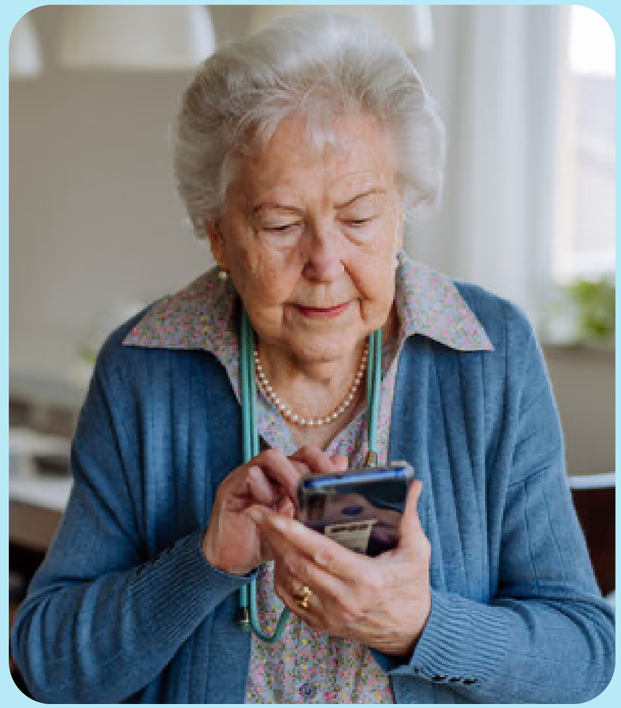
- **Cognitive Biases (Normalcy Bias, Optimism Bias & Habits) -** Normalcy bias causes people to **underestimate** the **likelihood** of an unexpected **threat**, so familiar **messages appear safe** (SCARS, 2023). Optimism bias affects scam susceptibility as people **do not believe negative events will happen to them**, so vigilance is reduces and phishing susceptibility increases (Wen, Flowerday & van der Schyff, 2024).

Butavicius, M., Parsons, K., Pattinson, M. and McCormac, A. (2022) 'Why people keep falling for phishing scams: the effects of time pressure and deception cues', Computers & Security, 120, 102937.
BodyMindBrain (2024) The psychology of stress and performance. Available at: https://www.bodymindbrain.co.uk/the-psychology-of-stress-and-performance/
SCARS (2023) The Normalcy Bias: Understanding the Cognitive Bias That Can Put You in Danger. Romance Scams Now. Available at: https://romancescamsnow.com/dating-scams/the-normalcy-bias-understanding-the-cognitive-bias-that-can-put-you-in-danger/ (Accessed: Nov 2025
Wen, M., Flowerday, S.V. & van der Schyff, K. (2024) 'Optimism bias in susceptibility to phishing attacks: an empirical study', Information and Computer Security, 32(5), pp. 656–675. doi: 10.1108/ICS-02-2023-0023.

Background Research

Who is Vulnerable to Scams?

To understand our target users, we reviewed research on scam susceptibility across demographics. Studies show a clear divide in how age groups react to cybersecurity scams:



- **Older adults (55-74)** had more trouble discriminating between genuine emails and phishing emails , due to **lower digital literacy.** (Grilli et a. 2021)
- **Younger adults** were more confident in the identification. However, more likely to be targeted by scammers. They are mostly targeted through social media and texts/emails, but are also **caught out online shopping sites**, (Mouncey & Ciobotaru , 2025) .
- **Overconfident users:** studies from Diaz et al. (2018) show that students who believe they understand phishing are actually **more likely to click phishing links** than those with little or no knowledge, highlighting **overconfidence as a key** risk factor.

Across studies, both **overconfidence** and **low digital literacy** increase vulnerability to scams. Since anyone with an email, phone number or social media can be targeted, our focus is on adults with varying digital literacy, designing proactive, real-time support rather than relying on users' judgement.

Grilli, M.D., Fermin, A.S., Bercovitz, K.E., Jarret, E.M., Garrett, K.D., Scheiderer, E.M. and Huentelman, M.J. (2021) 'Aging Influences the Difference in Trust That Individuals Place in Authoritative Information versus Social Information'
Mouncey, R. and Ciobotaru, G.-F. (2025) Investigating Online Users' Phishing Awareness on Instagram: An Exploratory Study Using Auditory and Visual Stimuli and Granular User Data Analysis.
Diaz, A., Sherman, A.T. and Joshi, A. (2018) Phishing in an Academic Community: A Study of User Susceptibility and Behavior.

Background Research Evaluation

Purpose

To understand the **types of scams** impacting people, we must understand **who is targeted** and impacted most by scams. To identify **existing solutions** for scam detection and prevention. This helped us develop a **deeper understanding** into the specific problem area we wanted to focus on. Background research ensured our problem was rooted in **real user vulnerabilities, not assumptions**.

Contribution

Our research revealed how rapidly scams have adapted due to **AI-generated content** and **social engineering**. Stress, emotion, urgency and cognitive overload play a key role. Reviewing existing apps, revealed strong detection. It clarified the intervention must support **real-time, emotionally-aware decision making**, rather than static education.

Learning

We learned vulnerability is driven by both **digital illiteracy** and **overconfidence.** Existing tools **act too late,** rely on users noticing warnings, and rarely support reflection. Behaviour under pressure, not knowledge, drives susceptibility; **scams exploit emotion and speed.** We realised our early assumptions about more training = more safety was false. Defences must intervene outside the moment of high stress.

Effectiveness

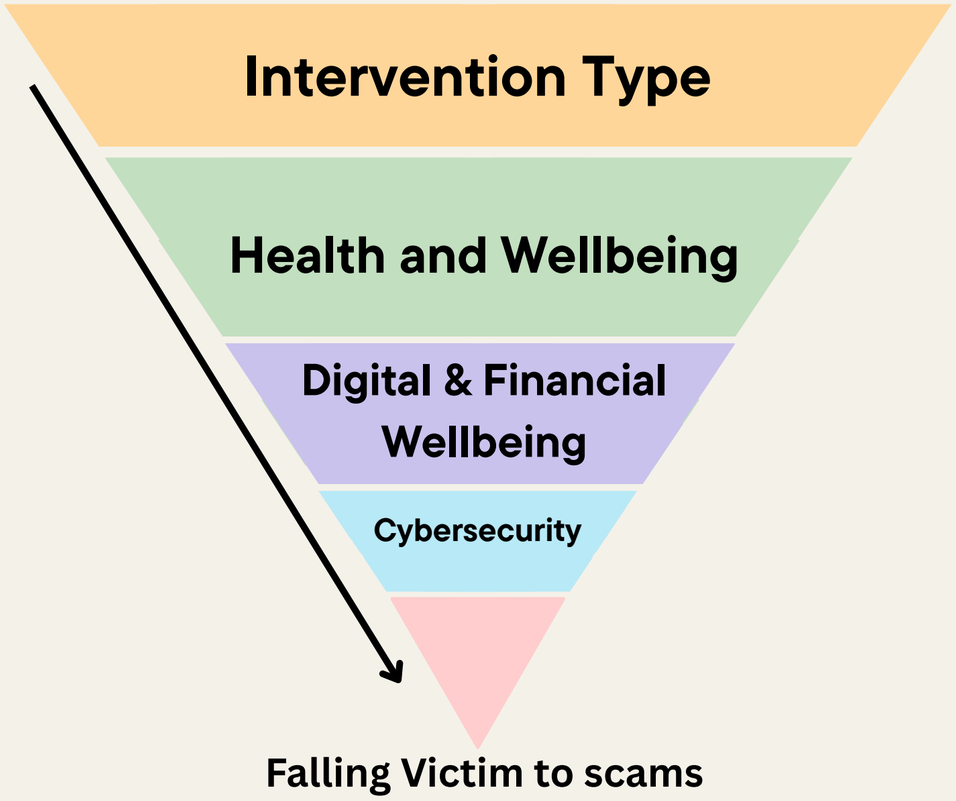
Background research was very effective in identifying the **key issues that need addressing** out the threat landscape, exposing gaps in current solutions, and developing a **cognitive understanding**. We relied heavily on **secondary literature**, and could have gone deeper into case studies, examining specific scams and the mechanisms behind them **would have strengthened** the work.

Influence on next stage

This research has informed areas to focus on in the questionnaire - **users' confidence** levels, how users stay safe online and their **current cybersecurity awareness**. It allows us to gather **qualitative insights** that secondary research cannot provide and ensures our primary research will be **built on validated evidence**. It will help shape the focus of subsequent design stages.

Early Problem Statement

From our brainstorming and demographic exploration, we identified **falling victim to cybersecurity scams** as a critical, everyday behaviour that links **digital and financial wellbeing**. We narrowed it down further with **background research.**



Context

Cybersecurity scams (including phishing, smishing, impersonation) are becoming more sophisticated and personalised with advancements in AI.

Challenge

People continue to trust and interact with fraudulent messages due to overconfidence, convenience or unclear warnings, rather than due to a lack of tools.

Statement

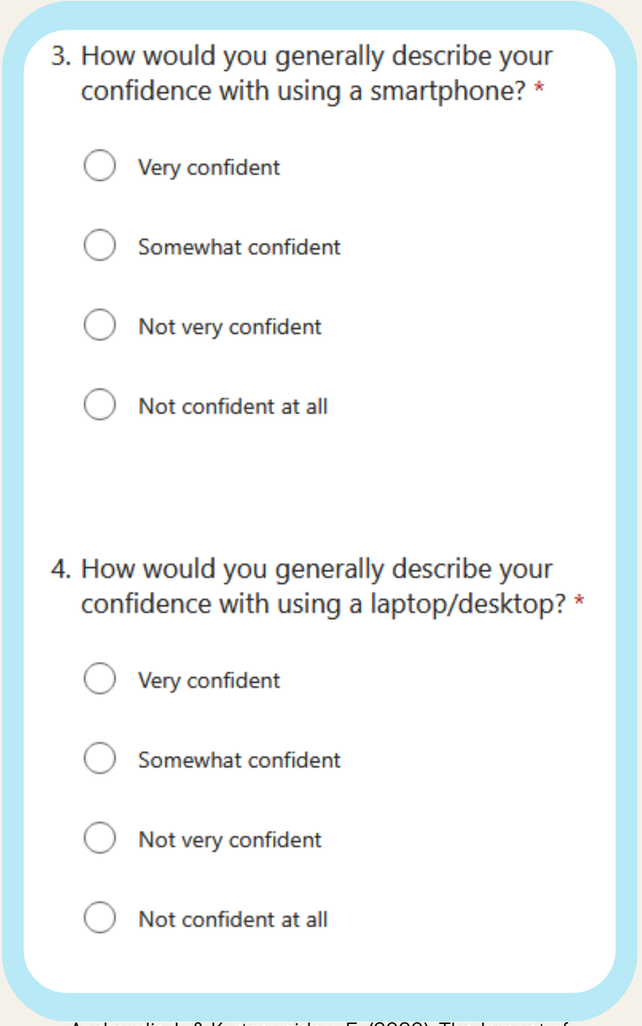
Individuals struggle to recognise and respond to online scams confidently despite a wide range of protection tools.

This statement will be refined as questionnaire, interview feedback and persona informants reveal deeper insights into the specific factors influencing individuals' responses to scams.

Questionnaire Design

The aim of our questionnaire was to narrow down and **identify the problem** as well as gain reliable, **empirical** data, and improve validity.

- We used the **Likert scale** for quick analysis of **quantitative** data. We **purposely removed the neutral option** to prevent respondents from defaulting to the middle answer. This forces them to reflect more on their decisions, **removing the "safe" option** and achieves **clearer data insights.**
- **Branched** the questionnaire so respondents only received questions relevant to them. This resulted in **more personalised answers** and **better engagement**. It means the questionnaire was not unnecessarily long, improving **data accuracy** and ensuring **validity** of results.



We trialed the questionnaire on a small **sub-section** of people initially, to test out the flow and wording of questions. This helped us to **understand logical ordering**.
We split the questionnaire into 5 sections: background information, awareness of cybersecurity, personal experiences, behaviour using technology and desired support. Splitting it into categories, **improved data quality** and **reduced respondent burden,** leading to **higher completion rates, with only 1 non completion** (Andreadis, I., & Kartsounidou, E., 2020).

Andreadis, I., & Kartsounidou, E. (2020). The Impact of Splitting a Long Online Questionnaire on Data Quality. Survey Research Methods, 14(1), 31–42. https://doi.org/10.18148/srm/2020.v14i1.7294

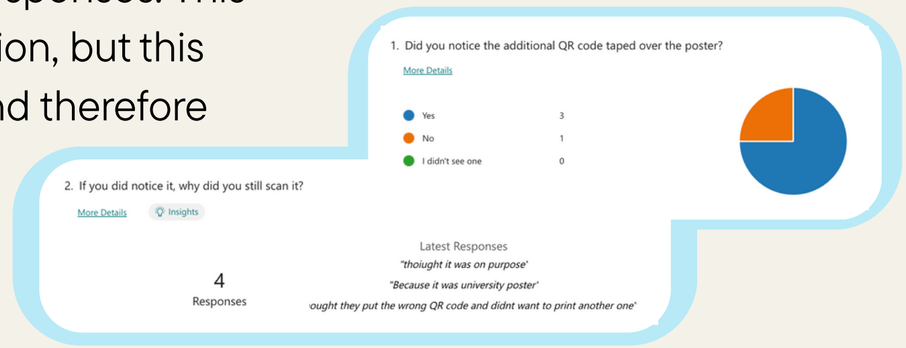
Questionnaire Distribution: Flyer

We distributed our questionnaire via two methods: forwarding a **link** to a Microsoft Form online and by sharing a physical **QR code** to the form on **flyers**.

Link sharing was effective for gathering responses, as we all knew the friends, family and colleagues we were sharing the form with, so we could **guarantee responses**.
However, we were aware that sharing the link to people we know created an **unconscious bias** in our results, as most respondents were also **computer science students, aged 18-24**.

In an attempt to shift this bias and **diversify** our results, we distributed **flyers** with a QR code to the form around various buildings on university campus. The design of the flyer contained statistically **eye catching colours** in an attempt to **gain attention**.
Our flyer featured an additional QR code taped over the flyer, as we tried to impersonate the "Fake QR code" scam.

However, this approach was **not as effective** as we received a low number of responses. This could be due to the **locations of the flyers** or the **lack of interest** in participation, but this told us that **QR codes scams were not a big issue** with our target audience and therefore **not an area we will focus on with our design**.



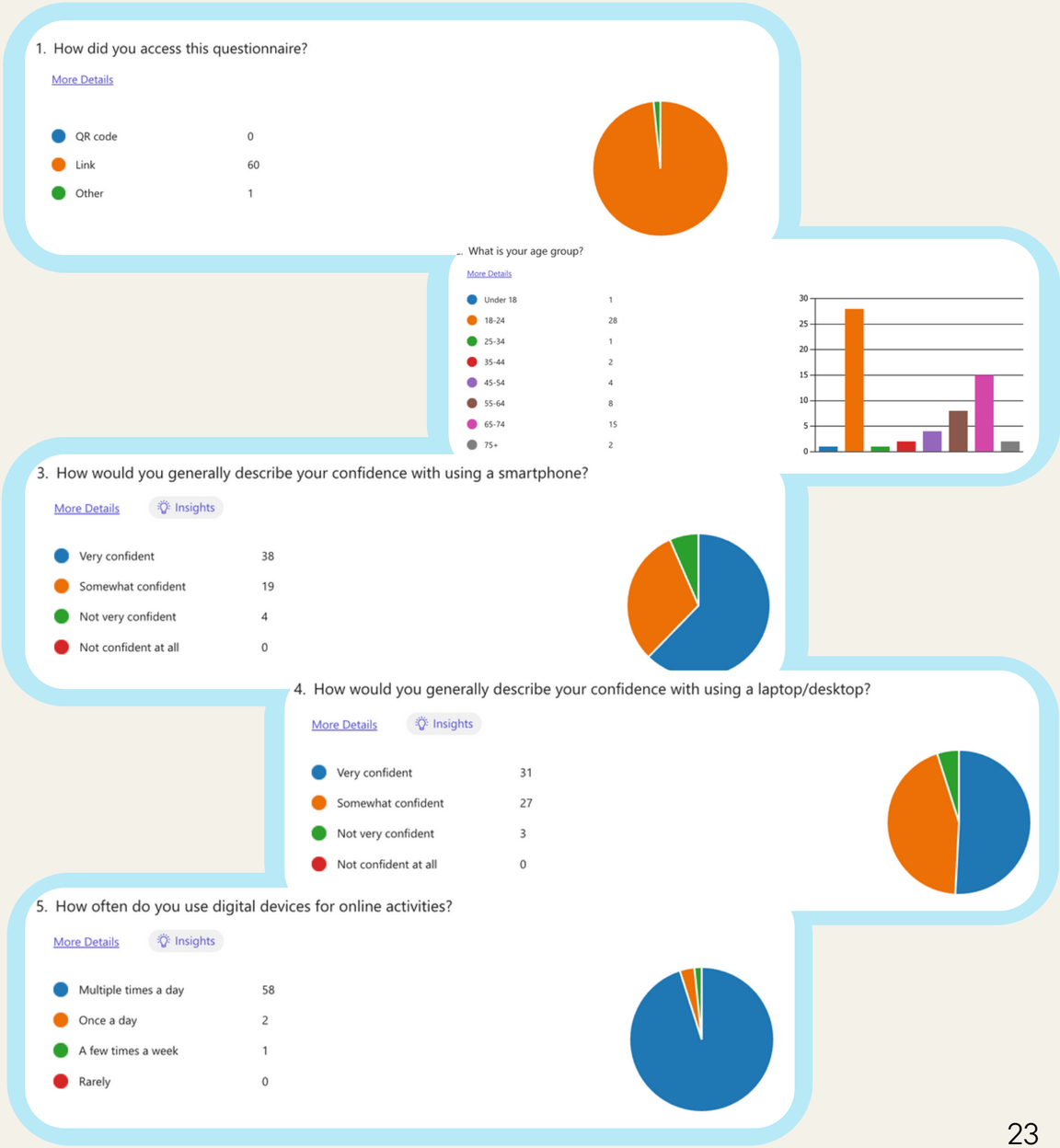
Outcome - In our fieldwork exercise, people did not fall victim to fake QR codes

Questionnaire: Background Information

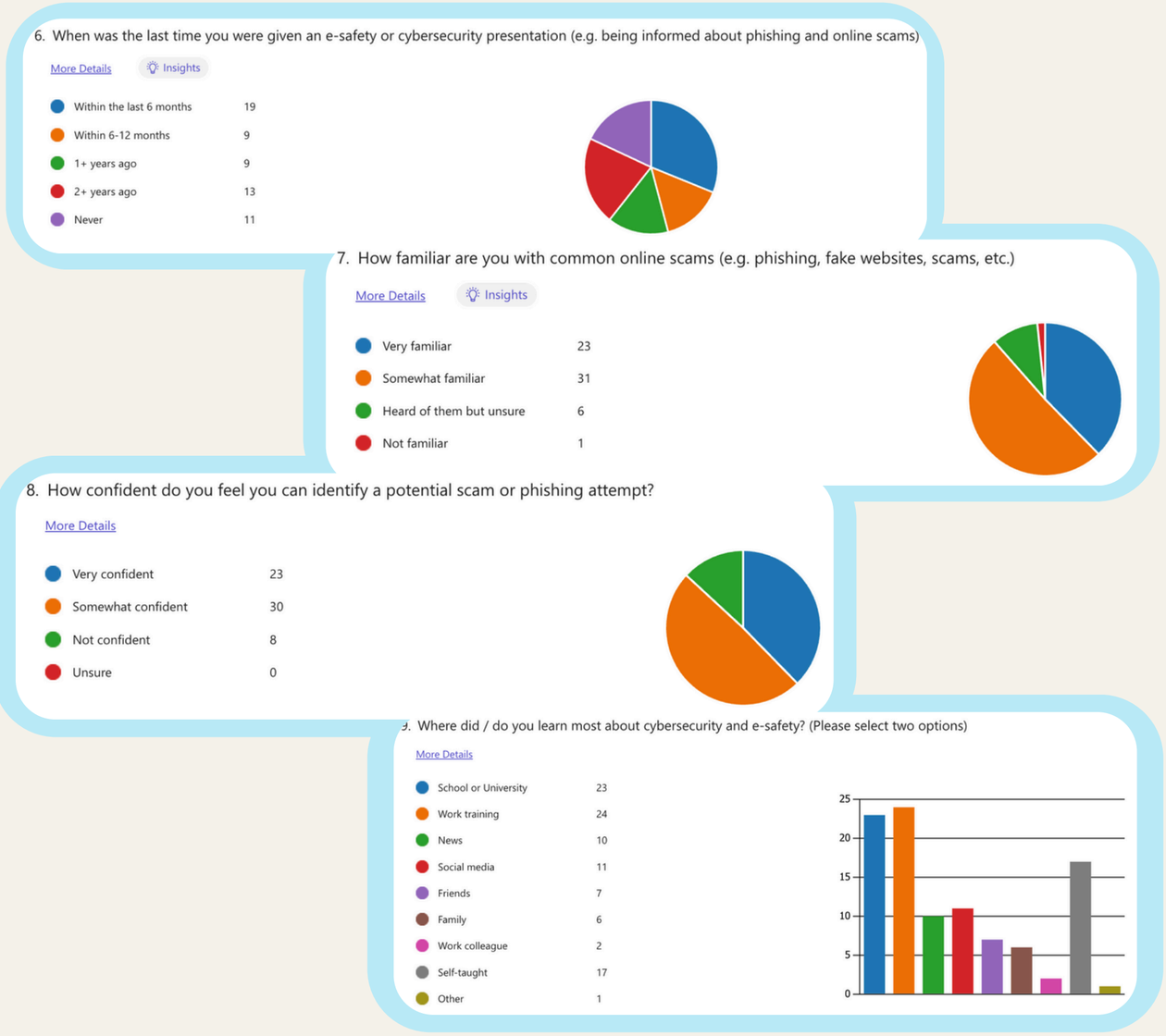
- The questionnaire had **61 respondents** overall, with an average completion time of 9 minutes, demonstrating our intention of **collecting meaningful data** without compromising attention.

- The demographic of respondents showed **skewing towards 18-24 year-olds** (28/61 respondents), along with a significant representation of older adults between 55-74 years old (23/61 respondents).

- Sharing the questionnaire primarily via links resulted in an **unconscious respondent bias**, as many participants are individuals with relatively high digital confidence.
- **95%** of respondents use technology **multiple times a day**, meaning they are all in a position to be exposed to online scams.



Questionnaire: Awareness of Cybersecurity



- Respondents displayed a mixed but overall **moderate familiarity with online scams**, with most participants selecting “somewhat familiar” (31/61 respondents) and “very familiar” (23/61 respondents).

- 53 respondents felt confident in identifying scams, but **often rely on outdated indicators** such as spelling errors or suspicious URLs, highlighting **gaps in confidence vs. knowledge**.

- **More than 50% of respondents have not received cybersecurity training** in at least 1-2 years (33/61 respondents), representing the need for accessible and ongoing support.

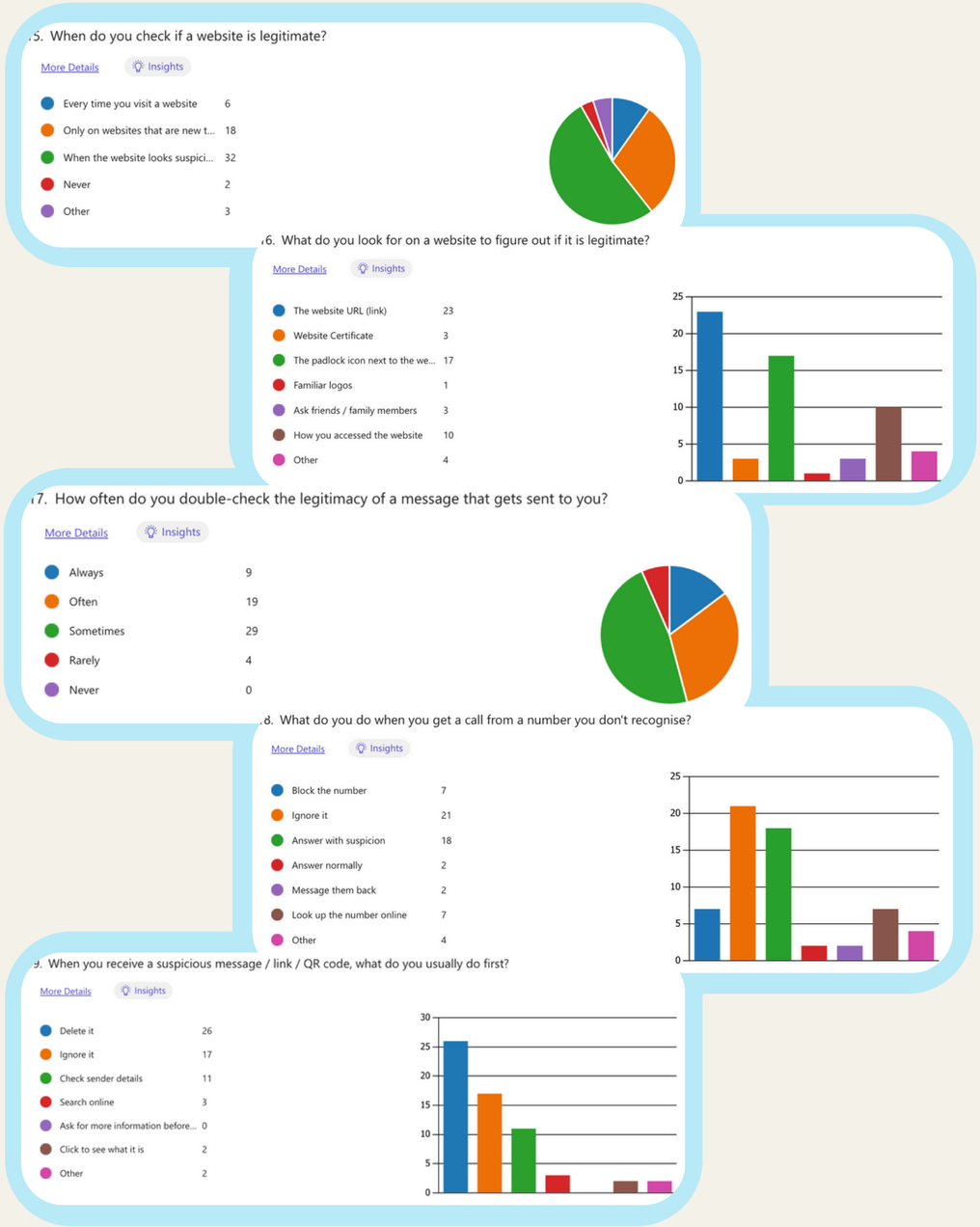
Questionnaire: Personal Experiences

- **67% of respondents have been targeted by an online scam**, with the most common occurrences being phishing emails (22/61 respondents) and text scams (13/61 respondents).

- 48% of respondents realised they were being scammed beforehand, whereas 20% of respondents either realised too late or not at all, indicating a **lack of consistency in being able to recognise potential scams**.

- Respondents **primarily rely on basic cues** such as spelling and grammar, link legitimacy or sender address to recognise scams, displaying the **importance of improved scam detection** as AI and other technologies advance further.

Questionnaire: Behaviour Using Technology



- 52% of respondents only check a website's legitimacy if it seems suspicious, and 30% of respondents only check the legitimacy of websites that they haven't visited before. **Only 9% of respondents check website legitimacy every time**, suggesting reactive cybersecurity behaviours rather than preventative.

- Upon receiving a suspicious message, 70% of respondents either delete or ignore it, but 18% of respondents still interact depending on **whether they gauge the sender's details to be safe**.

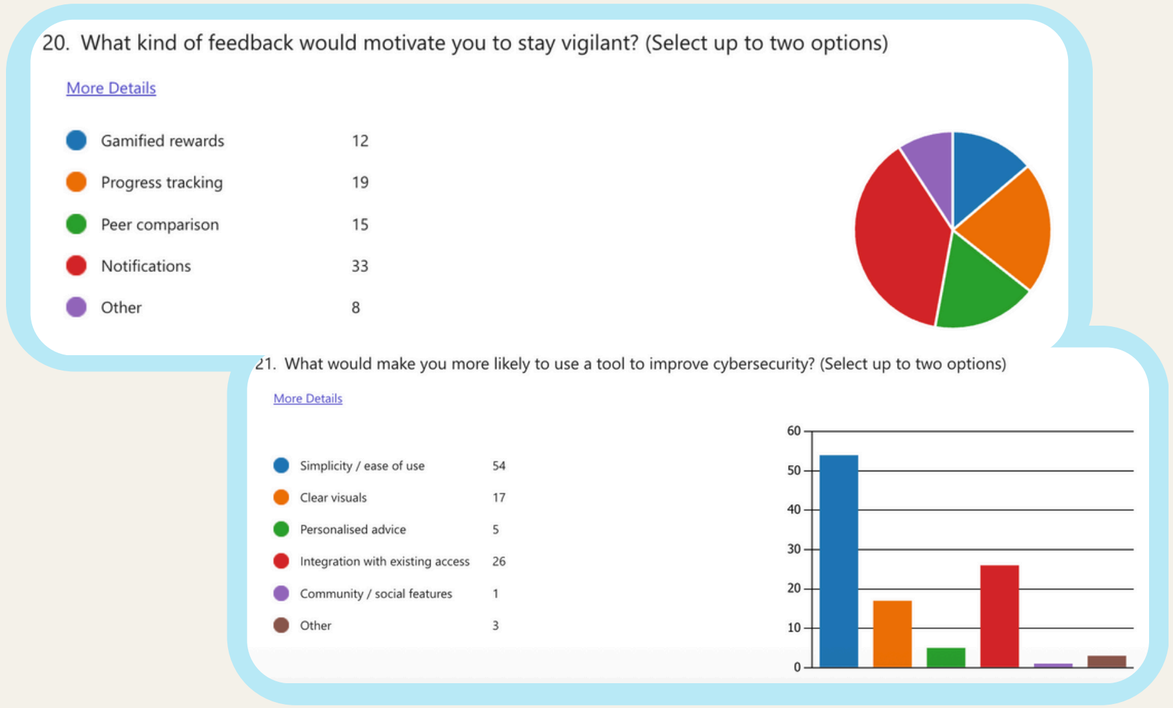
- Scam calls had the greatest variety of responses, with 34% of respondents ignoring calls from unknown numbers, 30% of respondents answering with suspicion and **only 11% of respondents blocking the number immediately**, demonstrating the **demand for tools that can assist** in scam call detection.

Questionnaire: Desired Support

- The most desired element of support is ease of use, with 89% of respondents selecting this. This is indicative of the **collective need for simple security tools** that don't require significant amounts of effort.

- 43% of respondents prefer **integrated tools**, and 28% of respondents prefer **clear visuals**, displaying the desire for non-technical and intuitive support.

- Responses indicate that **the most significant contributor to staying vigilant is notifications**, with 54% of respondents selecting this answer.



Questionnaire Evaluation

Purpose

To gather a large number of **quantifiable insights** into users' everyday online behaviour, experiences and support needs. It aims to uncover areas to address and **how material** those issues are. It follows UCD practices for **gathering data early** from a **broad and diverse group** (Adams & Cox, 2008).

Contribution

The questionnaire revealed that **cybersecurity knowledge is outdated** and the most common things people look for are **poor grammar and spelling** mistakes. People commonly encounter phishing via emails and texts. **Some still engage with scams** under pressure.

Learning

Users are confident with devices but not security. Their **reactions are often emotional** or rushed, highlighting the **need for reactive measures**. Solutions should be **easy to use**, as **notification-based support** and proactive risk detection is widely valued.

Effectiveness

The questionnaire effectively identified digital confidence, habits, patterns and pain points. However, it faced **sampling bias**, with most respondents aged 18-24 or 55-64. Including scenario-based questions could have provided **deeper behavioral insights**, and the use of mainly quantitative self-report measures (questionnaires), didn't reveal why users engage with scams.

Influence on next stage

The data ruled out solutions that rely on users **manually detecting** scams, since respondents mostly verified only when already suspicious. This directed our next steps towards qualitative methods like **interviews** and **personas** to explore emotional drivers why people engage with scams in greater depth.

Adams, A., & Cox, A. L. (2008). Questionnaires, in-depth interviews and focus groups. In Cairns, P., & Cox, A. L. (Eds.), Research Methods for Human-Computer Interaction (pp. 17–34). Cambridge University Press.

Interview Purpose

We conducted 8 **separate interviews** across a **variety of people** with the aim to gain **diverse perspectives and insights** into certain topic areas.

We also conducted an interview with a **professional in the field** with the aim to create **an anti-persona**; finding differences in knowledge so we don't exclude people from using our product that don't have a vast experience.

Interviewees


Software Engineer


Business Student


Data Analyst


Cleaner


University Student


Cybersecurity Professional


Stay at home parent


Retired person

Question Topics

1. Demographic information
2. Technical and cybersecurity knowledge
3. Previous exposure to scams
4. Fears or challenges in relation to scams

Interview Design

We decided to go with a **semi-structured interview script** to allow us to fully explore people's experiences, regardless of their level of expertise with respect to scams and cybersecurity.

We followed the interview structure: **Background**, **letting off steam and** addressing issues (Adams & Cox, 2008).

Letting off steam was important to **establish the interviewees' understanding** of cybersecurity and main issues around the topic.

We have a **'forked' format** for our interview, where once we understand if someone is an expert/not been scammed, we ask them **tailored questions** different to those that are not experts or have been scammed.

This allowed us to drill into the **most powerful or interesting questions** for the interviewee, or change track if they are not giving **engaging responses.**

1 **Background information**

Getting information about the interviewee such as demographic data, technical knowledge and past experiences so shape the interview.

2 **Letting off steam**

Asking interviewee about their overall thoughts on scams and cyber security, so they can let out any strong opinions first.

Understanding the User - Choose depending on experience

Those who have been scammed

Asking questions about what happened, when they realised it was a scam and repercussions.

Those who haven't been scammed

Asking about their opinions of cybersecurity and scams, and their concerns of loved ones.

Cybersecurity experts, both scammed and not

Asking about trends/insights from the industry, their experiences and main concerns.

(Adams & Cox, 2008). - Adams, A. and Cox, A.L., 2008. Questionnaires, in-depth interviews and focus groups (pp. 17-34). Cambridge University Press

Interview: Cybersecurity Professional (Age 38)

Insight	Design Implication
Scams are becoming more **personalised and AI-driven**	Training needs to **adapt** in **real time** to reflect new tactics
Users **overestimate their ability** to spot scams	Promote awareness by making users **stop and think** rather than react automatically
Phishing training decays over time	Include **gamified refreshers** to maintain knowledge
Security systems that are **too complex** lead to **avoidance**	Design needs to balance safety with simplicity
Older users often **avoid digital tools** altogether due to **fear**	Include supportive, confidence-building interfaces for all abilities

Key Interview Quotes

- “People think scams are obvious - but AI makes them look frighteningly real now”
- “Organisations have playbooks; individuals need one too”
- “The biggest barrier isn’t knowledge; it’s fatigue. If security feels like a hassle, users give up”
- “The challenge is to balance security and usability”

Interview: Retired (IT) Person (Age 70s)

Insight	Design Implication
Reluctant to learn new software	Digital products must be **intuitive, without excessive technical information**
Isn't up to date with new scams because new ones are out every few months	Should not require the user to update themself on new scams, instead telling them things that are **relevant to them when needed**
Doesn't want to have to keep updating the software for the latest features	Product should make tech more interesting **and not be bogged down**. Should work light on the system and update itself without the user doing anything
Knows privacy risks, but effort feels overwhelming	**Should not hinder what they are already doing** and take up unnecessary time. It should inform without directly teaching
Find junk emails annoying	Should help users by **hiding any spam and junk** to clear the clutter

Key Interview Quotes

- "Its easier to do it on the phone than to boot up the computer"
- "[technology] has gone so berserk over the last few years that I am now so far behind... I just don't want to spend the time to catch up... I am retired"
- "[privacy] is another thing to keep up with"
- "When I am at work, I hear all around people talking about different scams"

Interview: Parent (Aged 60s)

Insight	Design Implication
Scams are becoming more personalised and scary	Intervention needs to adapt in real time to **reflect new threats** posed to users, or allow for updates
Work phones offer different risks, as you're more likely to need to pick up from unknown numbers	Consider mechanisms to **ensure trust in callers** where there is a need to pick up
Worries about older relatives, they may have less confidence and be more susceptible	Allow **connectivity between users** so people can check in with friends or relatives
Scam messages are often scary, can seem personally relevant and rely on time pressure to cause irrational actions	Design needs to be able to **intervene in a timely manner** and reassure user that not responding to the scam wont be detrimental (to prevent impulse actions)
Finds comfort in identity verification notifications from mobile banking apps	Look into how existing anti scam interfaces work, **implement alerts and notifications**

Key Interview Quotes

- "I have a good general feel of where I need to be aware online, but people are getting smarter"
- "Daytime TV often educates especially elderly people about new and emerging scams"
- "Im sure most people are aware they shouldn't click random links, but a text of the right context with some unfortunate timing can seem more real such as parcel scams"
- "You want to be able to trust people but lots have malicious intent"

Interview: Junior Software Developer (Age 21)

Insight	Design Implication
People working in senior positions are falling for phishing emails	Must appeal to **all levels of abilities** so they are protected and alert
Many incorrectly believe that they can easily identify a scam	Use intervention methods to **encourage users to think before they act**
People are only aware of scams once they see them around them	Product needs to be **informative and repeatable** to refresh awareness
People blindly trust their IT team at work	Product needs to reduce dependence and **improve individual alertness**
Introducing new software at work can be met with resistance	**Prioritise ease of use** and simple installation so people don't choose comfort over safety

Key Interview Quotes

- "I just stay alert to everything I'm clicking on and just don't be naïve"
- "We sent a fake phising emails to our accounts at work... senior purchase ledger filled out all the information"
- "The more scams, the more things that happen to my peers, the more vigilant and hyperaware I become of these scams incase it happens to me"
- "...because of the rapidly growing capabilities of technology that no one will ever be fully informed"

Interview: Student (Non-Tech) (Age 19)

Insight	Design Implication
People with less exposure to cybersecurity awareness have less interest in learning it	Needs to be accessible for people who **can't protect themselves** against malicious actors
Interface familiarity is valued as a factor for users' trust	Should have a simplistic interface that **establishes trust with its users,** building familiarity over time
It is easier for users to understand the consequences of bad cybersecurity practices than it is for them to understand preventative measures that can be taken	Should be automatic and **not reliant on users' own knowledge of cybersecurity**
Users are more likely to realise the need for internet safety when they or someone they know has been impacted by a security incident	Intervention strategies should **remind users of their relevance in any cyber-safe practices** at all times
People rely on information to be given to them	Product should consider that common cyber-safe practices in the cyber/tech industry may be **completely foreign to the average user**

Key Interview Quotes

- "I definitely could learn a few more things, but I just stick to what I know"
- "Maybe there is more information that I'm just not accessing myself"
- "I try not to use websites that might give me a virus or take my information"
- "I was lucky enough to not learn the hard way and just learn what I needed to be careful of"

Interview Evaluation

Purpose

To understand people's real experiences with scams, their **emotional reactions**, daily behaviours and barriers they face with cybersecurity.
It allowed us to get **qualitative research** that questionnaires could not provide and get **deeper psychological insight.**

Contribution

Interviews revealed behaviours such as reliance on **surface cues**, **outdated knowledge**, frustration with **complexity** and **fear** among older users. These insights have highlighted the **need for simplicity**, reassurance, automation and **real-time support** across varying confidence levels.

Learning

We learned most users do not want to learn new security behaviours, they want tools that work in the background, required **little effort** and offered timely, nudges that interrupted impulsive decisions without feeling intrusive. Stress, urgency and emotional triggers make users overlook scam cues.

Effectiveness

Interview provided helpful insights but our 8 interviewees did not represent all demographics. Some demonstrated **social-desirability bias**, wanting to seem more competent than they were, struggling to describe technical details. Next time, we would include **ethnographic methods** as interviews limited our ability to see **in-the-moment behaviours.**

Influence on next stage

These findings will guide us to create **personas** that reflect different confidence levels and vulnerabilities. **Scenarios** will be developed reflecting potential scam moments which will in turn help us redefine our problem statement, ensuring we implement **clarity**, real-time support and **ease of use**.

Empathise Reflections

Aims

- Identify a **problematic behaviour** that is a widespread, relevant issue with the backing of data
- Conduct background research to gain a **diverse range of sources** of information and to understand solutions that exist within the space
- To gather a large number of responses for the questionnaire and interviews to give **reliable and representative data**
- Use a combination of these insights to **find areas for development** moving forward with our design process

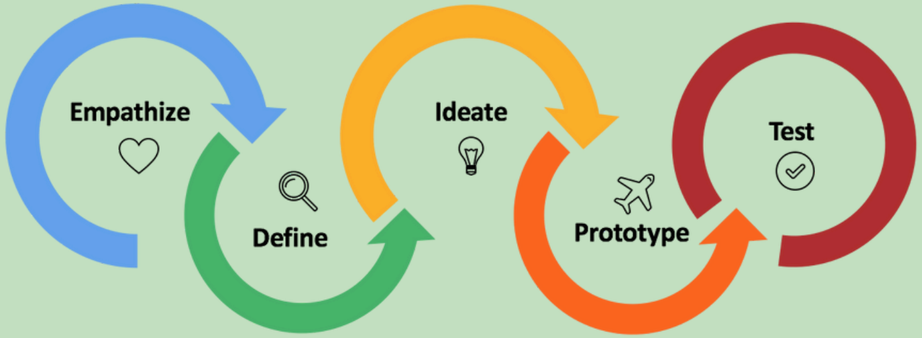
Learnings

- **Younger people are exposed to a higher volume of scams** whilst older people are more likely to fall victim
- Interviewing was an enjoyable process where we genuinely engaged and were receptive to the tailored and **embraced the semi-structured interview format** to extract **interesting narrative** and **overarching themes**
- **Human behaviours** are often the reason scams work, due to **emotive or rash responses**
- The problem we are aiming to solve is a **wicked problem**

Next Steps

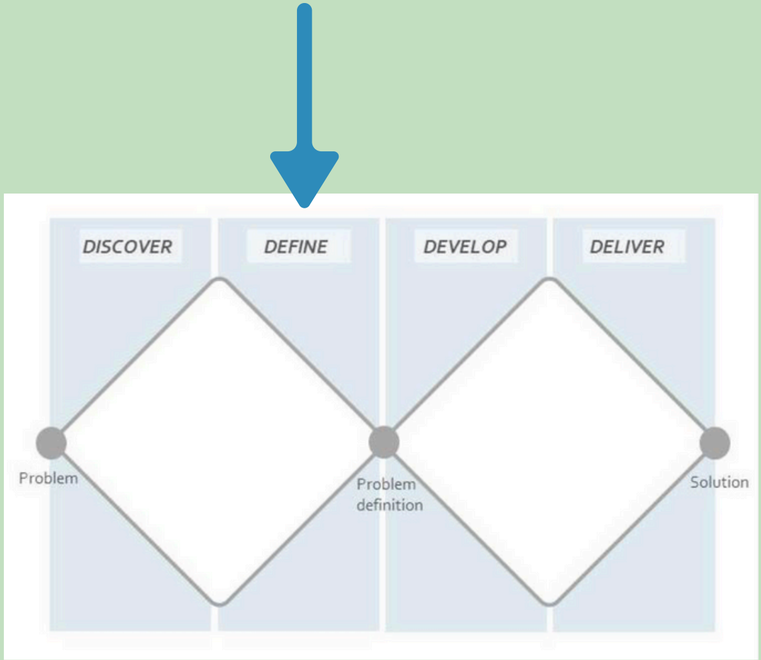
- Consolidate our **primary and secondary research** into personas in order to personify and segment needs of our potential users, since we are designing for adults in different stages of life
- Consider an **anti-persona** that can be used to ensure features aren't being catered towards experts in cybersecurity
- Create **scenarios and storyboards** that facilitate an understanding of **how and when** our design needs to **intervene**

To **Define**



Define

Understanding the problem



Persona Reasoning

We decided to create personas based on findings from our questionnaire and interviews, as well as the research we conducted. We felt that **4 personas were sufficient** in representing differences in our target demographic and **providing a clear reference point** when considering potential solutions, ensuring they **met our users' goals and needs**.

Georgie

Georgie represents a large number of results from our questionnaire. Having qualities reflective of our audience research and interview responses makes him a key part of our target demographic.

Jess

Jess represents the large number of young adults discovered in the questionnaire results, that use tech a lot but that don't have an in depth knowledge of staying safe online and emerging scams.

Linda

Linda addresses those with lower technical ability and less awareness, leading to a higher chance of being scammed, giving another perspective to our other personas. She also represents users that use technology at work.

Marcus

Marcus is our anti-persona. Being a cyber security analyst, he allows us to define our demographic's boundary and ensure focus of design remains on more naive users, ensuring the design stays simple and not catered to professionals.

Persona 1

Georgie Crosby

Age: 74

Job: Retired

Technical Ability

Frequency using Tech

Cybersecurity Awareness

Income

Georgie is retired, previously working for a local supermarket for much of their life. Georgie hasn't had much cybersecurity training through their life, having minimal training in the early 2010s and their technical skills has gradually declined during their retirement in the last decade as "there was little need to keep up to date". They primarily use a smartphone in their day to day but has needed to step back on their use in recent years due to declining vision. Their phone is used for social media, banking and messaging people but beyond that, they let their grandchildren help with the rest.

Goals

- Stay in touch with friends and family without fear of being scammed
- Continue to use apps as he usually does
- Doesn't want to spend much money or any if possible for cybersecurity
- Wants a simple and accessible way to know cybersecurity without any technical jargon

Pain Points

- Hasn't received any recent cybersecurity training or education
- Unfamiliar with modern threats. Knows about phishing scams and fake websites
- Has decreased movement due to weakend muscles
- Tends to trust messages that appear to be known companies
- Gets overwhelmed by technical prompts
- Avoids doing something new to avoid breaking things

Persona 2

Jess Hayle

Age: 20

Job: Student

Technical Ability

Frequency using Tech

Cybersecurity Awareness

Income

Jess is a 2nd year student at the University of Birmingham, studying BA Philosophy. She uses technology daily. For university, she uses her laptop to take lecture notes and complete assignments. She uses her phone a lot throughout the day watching TikToks, chatting with friends and for many other purposes like shopping for clothes. She hasn't been scammed online before but has lately become increasingly anxious of emerging scams, such as one where an official university email requested urgent transfer of tuition fee money. She was nearly scammed once trying to buy a ticket off Facebook for a club event as she was desperate to go, but a friend managed to step in in time. This experience has made her more tentative with engaging with unknown accounts.

Goals

- Continue to use phone as normal
- Get clear, quick indicators if something she is engaging in could be a scam
- Be more up to date on new or emerging scams
- Know tell-tale signs of fraudulent links or websites
- Feel more confident when trying to buy things second hand online
- Wants to know parents and grandparents are safe online

Pain Points

- Previously ignored important calls because it came from an unrecognised number
- Time consuming nature of ensuring latest software updates are installed and pass
- Sees conflicting information about cybersecurity on TikTok
- Feels less confident about being secure on her laptop as she uses less than her phone
- Often acts impulsively

Persona 3

Linda Davies

Age: 46

Job: Solicitor

Technical Ability

Frequency using Tech

Cybersecurity Awareness

Income

Linda is works for a reputable law firm as a senior solicitor in Manchester, where she has been for over 10 years. She uses technology day to day for work, but prefers to switch off when she gets home. While she is confident in her professional ability, she is less sure about her cybersecurity skills, having fallen for scams previously. Linda learns about common scams that have been around for years from her work's cybersecurity course, and is cautious about clicking suspicious links and answering calls from unknown phone numbers. However, she is less familiar with newer scams, including AI deepfake scams, one she fell for recently.

Goals

- Stay protected online to avoid being scammed again
- Maintain her professional reputation and trust between her clients and coworkers
- Be informed and educated about new and upcoming scams - such as deepfake scams
- Wants a simple and straightforward way to stay safe online as doesn't have time for complicated installations

Pain Points

- Has some troubles trusting technology now that she has already been scammed before
- Doesn't have lots of time to read about new scams due to her heavy workload
- Is only well versed in technological products used for her job
- Is able to detect older, more obvious scams but is oblivious to newer scams

Anti-Persona

Marcus Borwick

Age: 36

Job: Cybersecurity Analysist

Technical Ability

Frequency using Tech

Cybersecurity Awareness

Income

Marcus is a cybersecurity analyst at a major financial firm in London with a computer science degree and over a decade of experience in phishing simulations, network monitoring and penetration testing. He is technically confident and at home, he manages his own router with advanced custom security settings.

Marcus does not use consumer detection tools, he views them as redundant and profit-driven. He already possesses the skills to identify scams and believes his personal network security is stronger than antivirus software, which he believes are fear-based marketing for most users and adopts safe practices online.

Goals

- Maintain a high level of security control through his own knowledge and systems
- Customise or automate security solutions rather than relying on guided tools
- Avoid interruptions or redundant alerts that slow him down.
- Stay informed about emerging attack trends through his work and OSINT

Pain Points

- Finds consumer apps patronising and over-simplified
- Dislikes constant pop-ups which are distracting
- Would not engage in gamified features
- Feels most tools are not targeted well and are fear-driven, he does not believe in profiting from fear-mongering
- Skeptical about the data these tools collect and the privacy trade-offs for something he does not need

Personas Evaluation

Purpose

To turn research data into **relatable user representations**. As Grudin states, “Personas focus attention on a **specific target audience**” and prevent vague “design for everyone” thinking.

Contribution

Personas explicitly highlighted differences in **confidence, stress triggers, and digital literacy**. They helped us recognise and understand contrasting behaviours. We picked out **three primary personas** to represent three generations of user and their differences.

Learning

We found that individuals **respond** to scams in **diverse ways**, with motivations influenced by age, experience, and confidence. It was **difficult to encapsulate** these differences within personas, as each profile is unique. By developing 3 personas, we ensured that the **needs** of them were **not neglected**.

Effectiveness

We built the personas on insights uncovered by our background research, questionnaires and interviews, which was an effective way to encompass this data. However, because the questionnaires formed most of the evidence and the sample was limited, the personas **may not fully capture the diversity of behaviours** present in the wider population.

Influence on next stage

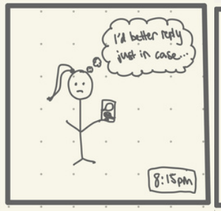
Throughout the UCD design process, we will use these personas to aid design, ensuring **each feature responds to their needs** and pain points. Through scenarios, we can bring their challenges to life which will give us a deeper understanding.

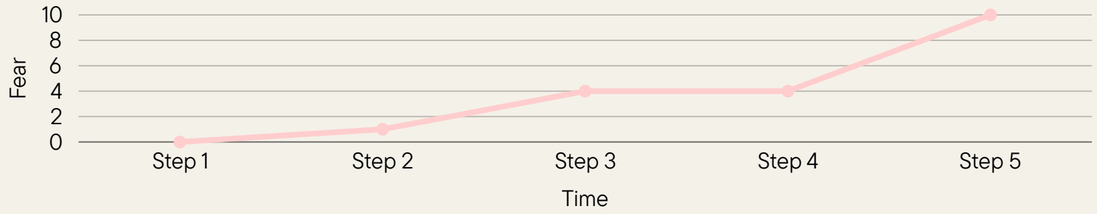
Grudin, J. & Pruitt, J. (2002) Personas, participatory design and product development: An infrastructure for engagement.

Scenario 1

“Jess regularly shops online for clothes, beauty products and other items. She is scrolling on her phone when she gets a text stating there is an issue with one of her deliveries and that she is required to send some extra details to receive her item. As she has purchased something a couple days ago, she clicks the link and fills out the information as she couldn’t remember the delivery provider in the moment. Later, she finds out it was a scam text when she tells friends about it”

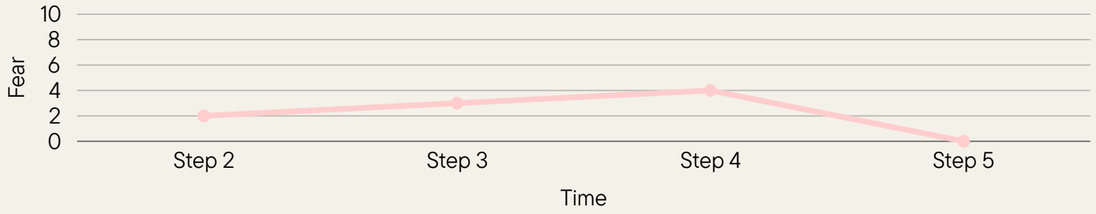
Before

Scrolling on phone	Gets delivery message	Recalls ordering online recently	Responds due to context	Identity is stolen
				
Jess is enjoying her daily scrolling time on her phone, and is happy to finally get time to switch off her brain after a long day of lectures, She does not suspect that she is about to become a target.	After business hours, Jess gets a message from an unknown number, claiming that she is required to input her delivery details for a package she has ordered.	Jess thinks about this, and know she has recently ordered some clothes that she needs for a party this weekend.	Jess needs the clothes to be delivered and the text is threatening they will be returned is she doesn't so act she clicks the link and fills out her details.	The malicious actors targeting Jess now have access to Personally Identifiable Information, allowing them to impersonate her and eventually steal her identity.



After

Gets delivery message	Recalls ordering online recently	Notified of suspicious text	Identity is not stolen
			
After business hours, Jess gets a message from an unknown number, claiming that she is required to input her delivery details for a package she has ordered.	Jess thinks about this, and know she has recently ordered some clothes that she needs for a party this weekend.	Jess is interrupted by her phone screen glowing with a warning, alerting her that the message she has just received is extremely suspicious and that it is a common delivery scam.	Jess is prompted to check directly with the site, finds her parcel is in fact on the way and deletes the message. She now has learned about this scam type.



Scenario 2

“Georgie is on social media and uploads a photo of him and his daughter. He gets a video call from who he believes to be his daughter, which is of bad quality but looks and sounds like her. She asks for money to be sent to her landlord as her account has been frozen and she can’t pay her rent. He sends the money without thought, but it turns out it was an impersonation scam and he lost £5000"

Before

Enjoying relaxing | Upload video to FB | Gets phone call | Child asks for help | Sends Money

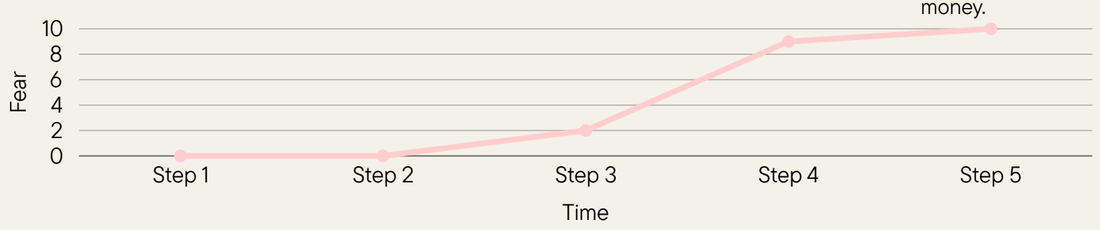
Georgie is enjoying retirement and using his phone, suspecting nothing. He isn’t thinking about much and doesn’t think that he is an upcoming target.

He uploads a video to his Facebook profile of him and his child, when they meet up. They don’t talk much since his child is busy so thought it would be nice to share to his friends.

Late at night, he gets a Facetime phone call. To be safe, he answers it, thinking that at worst he can just hang up and block them.

The person on the other end acts similar to his child, looks like them and sounds like them and although blurry and acting weird, they are saying they are in trouble and need 5k as they are in trouble. They ask to stay on the phone.

They ask Georgie to send money to a new account and tells him in detail the steps to send the money including the new account number and sort code. Due to the stress and fear, Georgie doesn’t think twice and sends the money.



After

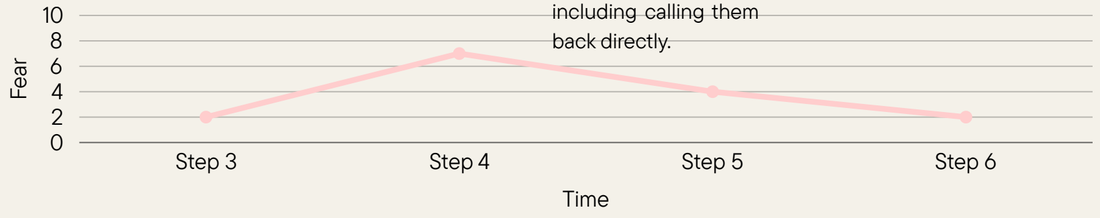
Gets phone call | Child asks for help | Notified of scam | Checks on child

Late at night, he gets a Facetime phone call. To be safe, he answers it, now aware that unknown numbers can be dangerous and to stay aware of everything. He is now more prepared for what is coming.

The person on the other end acts similar to his child, looks like them and sounds like them and asking for 5k because they are in trouble. Georgie is fearful but is more protected, but doesn’t want to dismiss his child.

The phone vibrates and interrupts the call, both audibly informing him and displaying it on the screen, signs of deepfakes and the chances it thinks it is a scam, including details that it detected. It gives him two options, ignore or hang up as well as ways to verify including calling them back directly.

He messages his child back to check if everything is fine. His child replies back informing him that it wasn’t them and to relax.



Scenarios Evaluation

Purpose

To bring to life **realistic scam situations** and reflect the emotional state of the persona alongside. They help us understand when fear is heightened and so when decision making abilities may **decline**.

Contribution

The ‘before’ scenarios revealed **impulsive** reactions, **confusion** and **heightened emotional** states. The ‘after’ scenarios showed how **timely support** can **interrupt unsafe behaviour**.

Learning

We understood the need for **real-time intervention**, reassurance and automatic detection. Through mapping fear, it revealed it’s most important to **intervene before the scam unfolds**, so panic does not occur.

Effectiveness

The scenarios were **built upon assumptions** and not actual observed behaviour. This meant that the full emotion of the scenario may not have been captured. We could have benefited from another scenario to capture other angles, but having 2 allowed us to **streamline the final solution**.

Influence on next stage

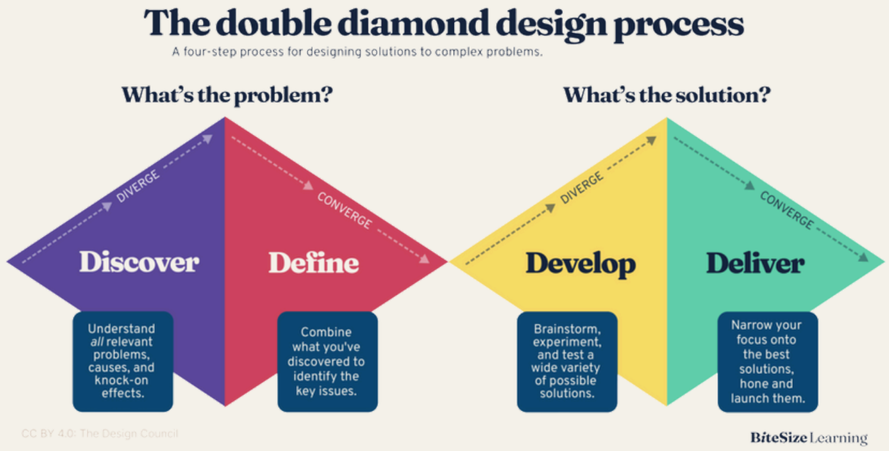
The scenarios have helped visualise when someone will benefit from help, and the **types of scams** we should target. Next, we will brainstorm **further types of scams** to help inform early sketches.

Refined Problem Statement

Problem Statement

People struggle to reliably identify and respond to real-time scam attempts as they often make rushed or emotionally charged decisions. As a result they often miss the warning signs of a scam putting themselves at risk of losing data or money. People who do notice signs often only do so in a reactive way. With increasingly personalised AI driven scams emerging, it is becoming harder to distinguish between scams and genuine communication.

The refined problem statement captures learnings from the emphasise and define stage and gives us a specific problem to solve as we continue through the UCD process. It is a statement built on research, ensuring the design focuses on real user needs.



Evidence Behind This

The questionnaire revealed **people rely on manual checks** and outdated cues. The interview showed stress and urgency triggered **impulsive, unsafe decisions**. Background research highlighted that **emotional pressure reduces rational thinking** and that many users were **overconfident in their ability** to spot scams.

This reflects our personas as:

- **Georgie** has **low confidence -** feels overwhelmed, avoids checking legitimacy, relies on reassurance.
- **Jess** is **impulsive** under pressure. She acts before thinking.
- **Linda** has **outdated knowledge**.

Define Reflections

Aims

- Consolidate our **primary and secondary research** into personas in order to humanise and segment users' needs, since we are designing for adults in different stages of life
- Consider an **anti-persona** that can be used to ensure features aren't being catered towards experts in cyber security
- Create **scenarios and storyboards** that facilitate in understanding **how and when** our design needs to **intervene**

From **Empathise**

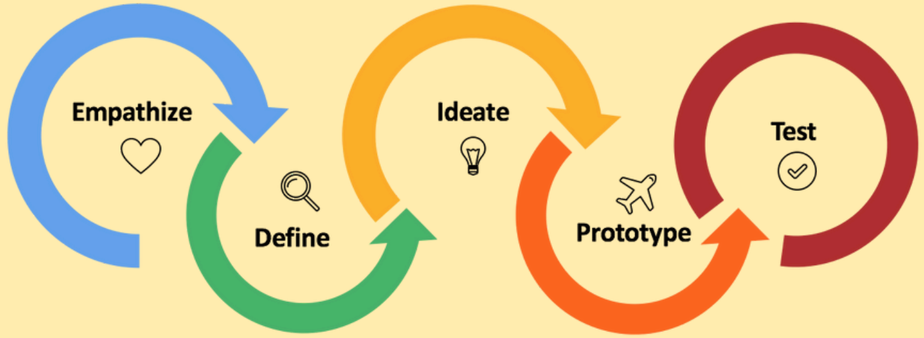
Learnings

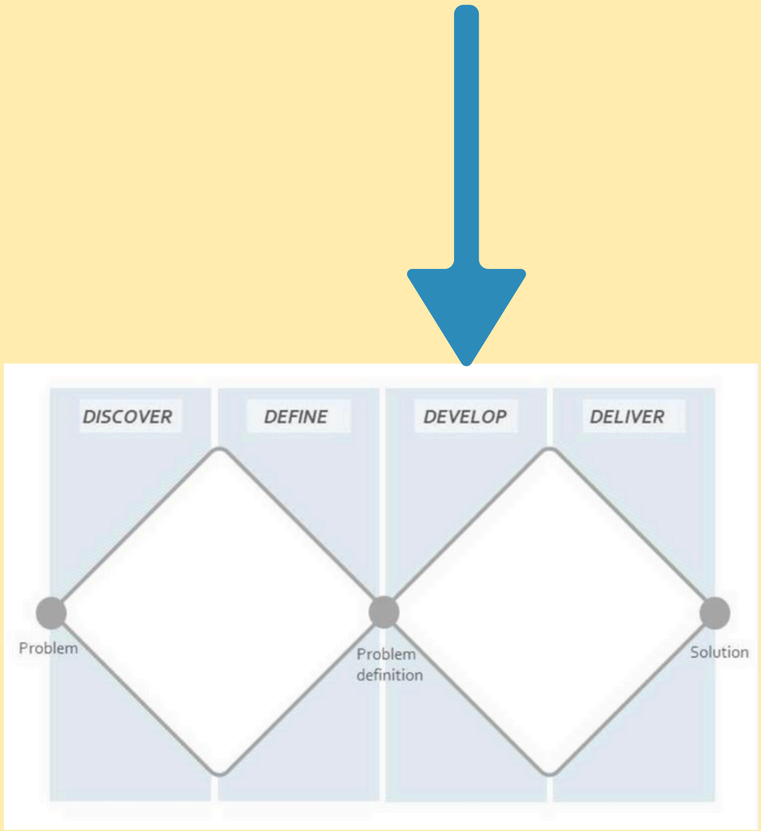
- **Clear and obvious alerts** to the user that **pre-empt or disrupt the scam** to prevent user response
- Team productivity improved as we settled on a routine and hit the **'Norming' stage of working as a team**
- Our **personas** capture our needs for 3 types of users:
  - Younger users that encounter lots of scams **but have some technical knowledge**
  - Users that have been **scammed previously**
  - Older users that need a **simple and accessible way to stay safe from scams**

Next Steps

- Aim to explore different forms of alerts in the **design space**
- Consider **different scam types** individually, then seek how solutions can be adapted or combined to address multiple
- Use different **ideation techniques** to create a broad set of primitive ideas
- Evaluate solution effectiveness **afterwards** using personas and scenarios
- **Seek value** in all ideas proposed and avoid ruling out 'stupid' ideas

To **Ideate**





Ideate

Exploring ideas

Designing for usability - key principles

Communications of the ACM, March 1985, vol 28, no 3. John Gould and Clayton Lewis (IBM - well-know interaction design experts). "Designing for Usability: Key Principles and What Designers Think".

QOC: Creating Criteria for our Designs

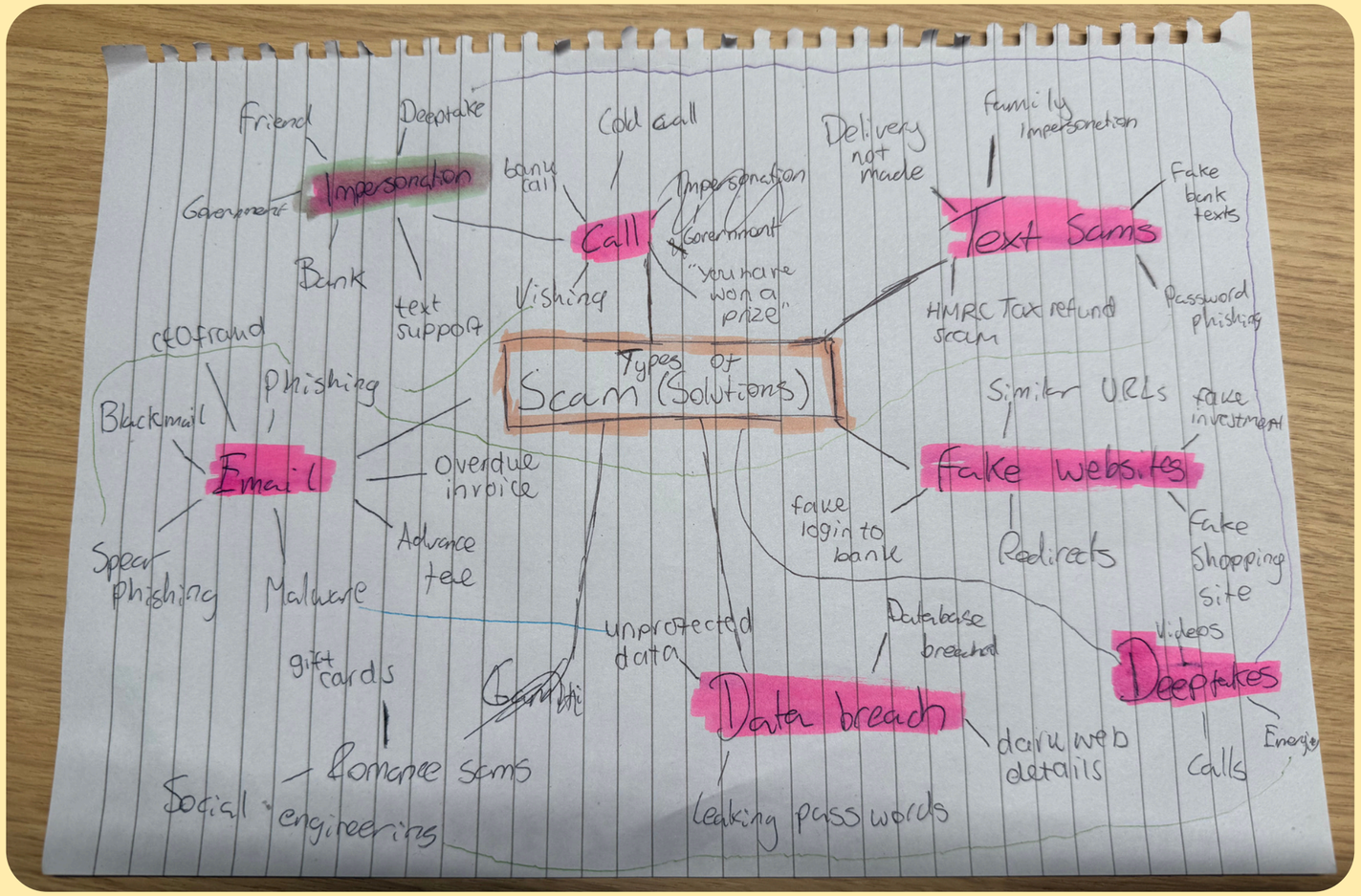
Question: Which design best supports users in reliabily identifying and responding to scam attempts across platforms, especially during rushed, emotional, or high-urgency moments?

Criteria	Meaning	Evidence behind the criteria
Ease of use	Has minimal setup and limited required user inputs	Identified as the top criteria in our Questionnaire, meets needs of **Georgie and Linda.** We also want to ensure it is not catering to **Marcus**
Pre-emptive	Is able to stop scams before the user gets exposed to them	Prevents users being exposed to the threat. Stops scams before **Jess** can act impulsively
Real-Time Intervention	Is able to stop scams as they happen	Critical because users make impulsive decisions under urgency **(1)**
Non-Invasive	Doesn't read excessive personal or identifiable information	**Georgie & Linda** distrust complex or invasive systems
Accessibility	Is able to be used for the average user and works for people with and without the product	**Georgie's** declining vision, desire for simple cues. Accessible for every skill level
Future-Proof	It can be updated to keep up with emerging scams, or be able to be relied upon in the future	Users are worries about emerging scams such as deepfakes from interviews, which **Jess** is worries about
Reflective	Helps the user understand scams or why their actions could cause them to be scammed	**Georgie, Linda and Jess** all want to be informed of scams so they can learn
Reliable	Is able to be trusted for information and has false positive/negative preventions.	**Georgie** wants certainty when messaging family and friends that he is safe
Integration	Can allow users to carry out daily activities without much extra addtional action	**Jess** wants to continue using her phone as normal, and **Linda** wants to avoid complicated installations

https://cybercashworldwide.com/the-dark-psychology-behind-the-tricks-used-by-scammers

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5485927

Brainstorming Solution Types



We brainstormed ideas for different types of scams, we identified 5 main areas to discover:

- **Phone call scams**
- **Text/SMS scams**
- **Fake Websites**
- **Deepfakes**
- **Emails**

We explored ideas within these areas to help inform ideation of our preliminary sketches.

We assigned a different scam type for each person to come up with ideas, but also collaborated on some.

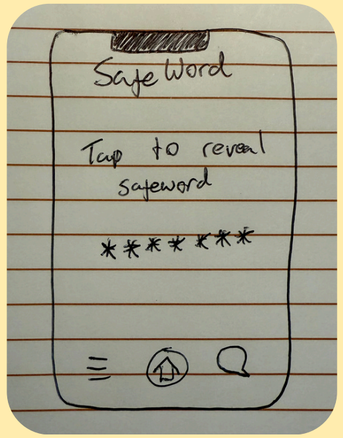
We saw how lots of the themes were interlinked which means we could end up combining features later on. We noticed the sub-theme of **impersonation** linked in with the scam types.

SMS Scam Ideas

SafeWord

Method - Lateral Thinking

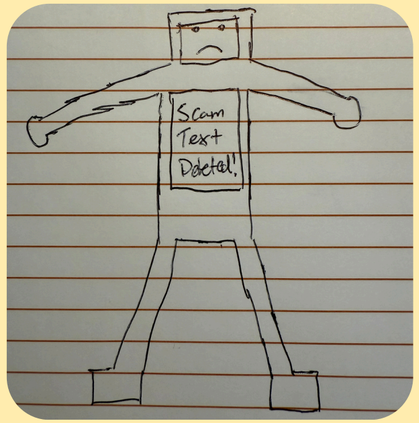
- Created by thinking about how to verify end user
- Random rotating word that only two users know so they can be sure they are talking to who they think they are
- Secured with FaceID
- Easy to use interface for older generation for **Linda** and **Georgie**



TextBuddy

Method - Future Envisaging

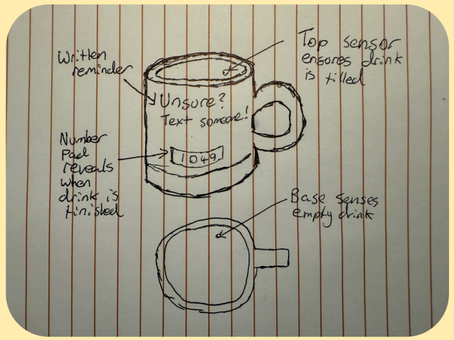
- Ideated by thinking about the future, where we have personal robots to help us avoid scam texts
- Visual indicators if a scam text has come through on a screen
- Useful for **Linda** at work as a quick visual aid, or for **Georgie** at home
- Can advise next steps



Drink & Think Mug

Method - Impossible Combinations (Mug and Text)

- Impossible combination of mug and scam text
- Aims to prevent a heat of the moment response like in **scenario 1 (Jess)**
- User has to pour and finish a tea/coffee before they are able to respond to a text from an unknown number
- Could have an app to go with it
- Encourages user to seek help/advice if they are unsure

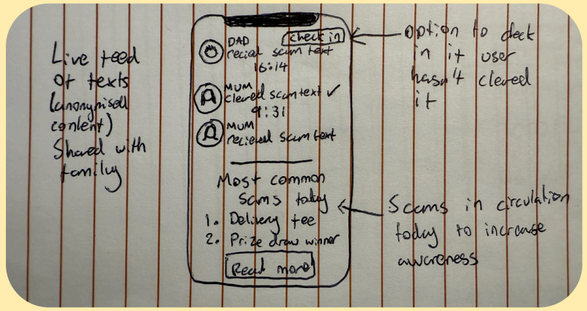


SMS/Deepfake Scam Ideas

CyberConnect

Method - Inspiration

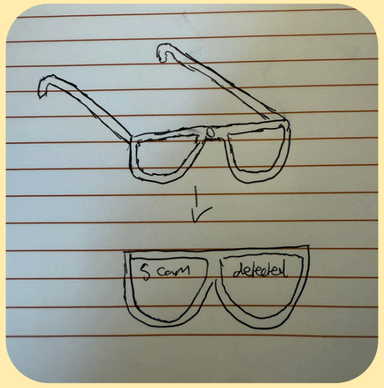
- Took inspiration from social media apps and aimed to apply them to scams
- Notifies a group (e.g a family) when someone within the group has received a scam text
- Users can flag as resolved if they have addressed it
- If they haven't, other users in the group can check in to ensure they haven't reacted
- Intervenes in **scenario 1.**
- Useful for **Jess** to check in with family



Cyber Glasses

Method - Future Envisaging

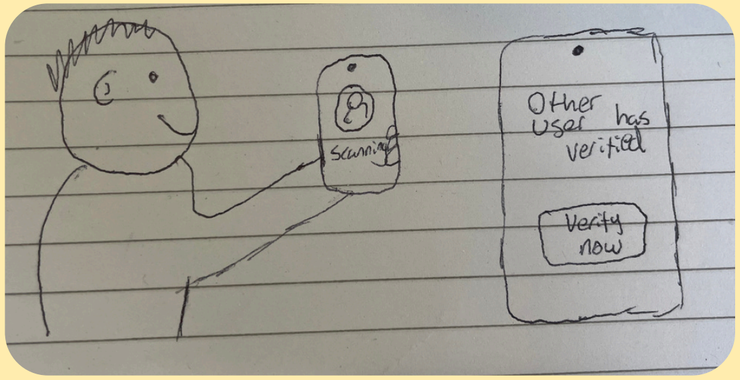
- Imagined humans having super vision that could be connected to IoT
- Reined it in to glasses with a camera that can read the content of a page/text
- Displays text in lenses if scam is detected
- May not be helpful to those who don't wear glasses usually
- Could be blue light glasses for those doing work like **Jess** and **Linda**
- Can intervene such as for **scenario 1**



Verified Identity Communication App

Method - Lateral Thinking

- A phone messaging and video call app that requires users to verify their identity before beginning communication
- Scans messages and blocks malicious texts before users can interact with them
- Notifies users of potential deepfake scams during video calls
- Addresses the need for intervention, depicted in **scenario 2**

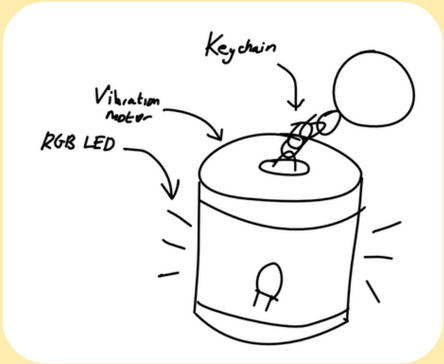


Deepfake Scam Ideas

Security Keychain

Method - Lateral Thinking

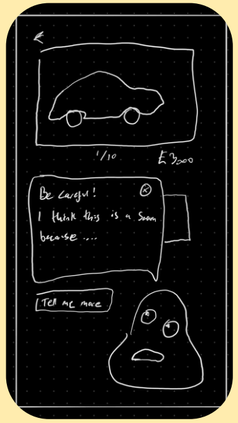
- Device you can put on your keychain, and it lights up depending on the current threat
- Aims to notify the user in real-time by connecting to an application on different devices
- Still kept up to date even if you are not at your device when something happens
- Supports **Georgie and Linda's** needs by being low-maintence



Security Assistant

Method - Future Envisaging

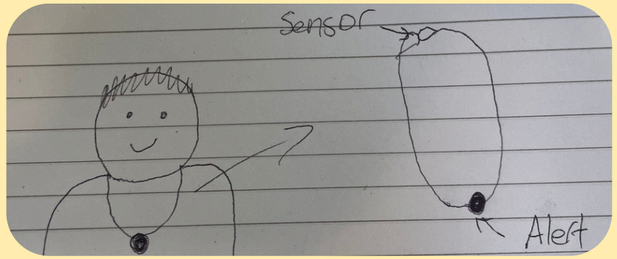
- Imagined a little blob minion that followed you to warn you of danger
- Made it reality by having it as a 'pet' as an overlay
- Pops up when it sees a text, call or video that seems like a threat to warn the user and inform them about it
- Useful for **Jess and Georgie** to tell them about scams



Intelligent Necklace

Method - Future Envisaging

- A necklace that connects to different devices and acts as both a 2FA device and notifies you when your security is at risk or compromised
- Features a vibration motor to subtly alert users of security breaches
- Has an app that rates your security and gives you tips on how to improve it
- Meets **Georgie, Jess** and **Linda's** needs by being easy to use and low maintenance

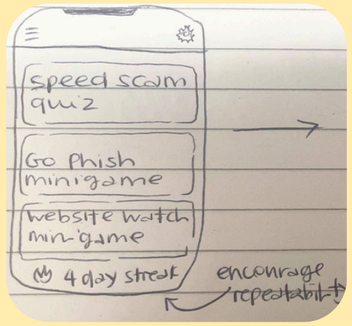


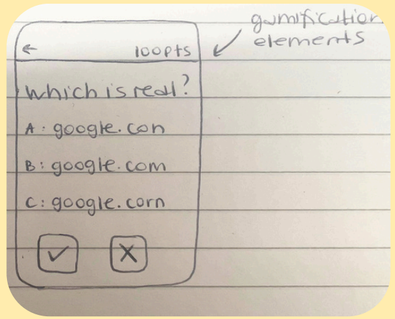
Website Scam Ideas

Gamified Learning App

Method - Gamification

- E-Safety course delivered as a fun, interactive game
- Questionnaire revealed over 70% of people haven't had an e-safety course in 1+ years or ever
- Aims to solve issue of poor education on new scams and naivety on older scams
- Encourages repeatability with "streaks" and "points" with friends

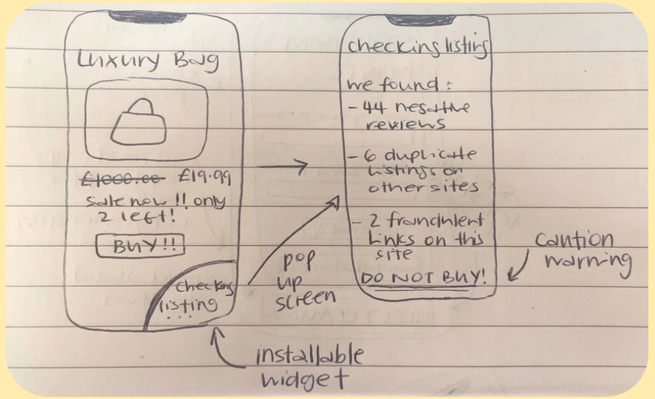




Online Shopping Agent

Method - Lateral Thinking

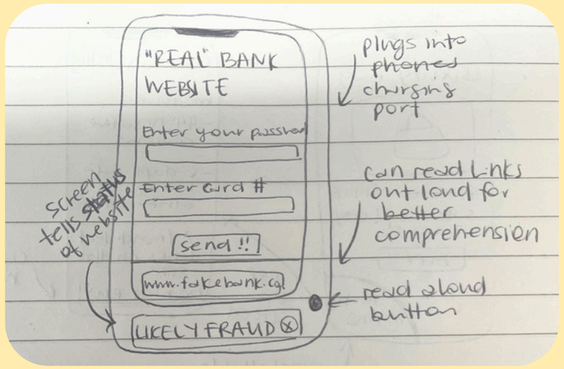
- Online tool that checks online retail sites for negative reviews, duplicate listings and unlicensed links.
- Inspired by **Jess**' pain points and questionnaire responses detailing online shopping scams
- Designed to run in the background, omitting user commitment
- Easy to use for **Georgie and Linda**



Website Link Verification Phonecase

Method - Future Envisaging

- Initial: someone attached to your phone 24/7 to check if a website was fake
- Reduced to a phone case that can check links for you
- Relies on checking website certificates and common signs (e.g. spelling)
- "Read Aloud" option for accessibility
- Solves **Jess'** problem of impulsive actions and **Georgie and Linda's** lack of awareness

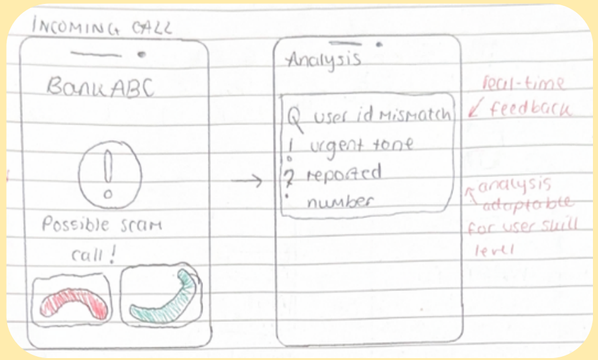


Phone Call Scam Ideas

Scam Call Detector

Method - Lateral Thinking

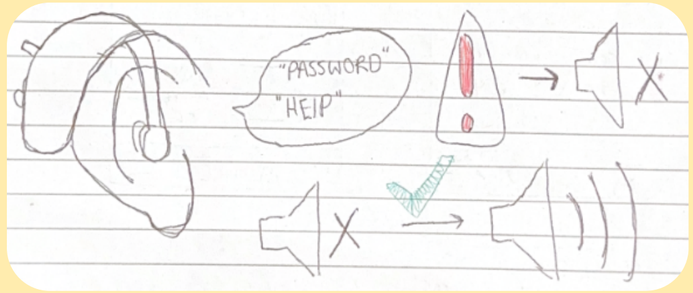
- Warns users during incoming calls
- Flag indicators: ID mismatch, urgency, reported number
- Real-time prompts to pause/reflect
- Post-call summary, adapts to user skill
- Caters to **Linda** and **Georgie**'s needs as is simple to use and set up
- Is informative as provides real time feedback and analysis



Safe-Scam Hearing Aid

Method - Inspiration

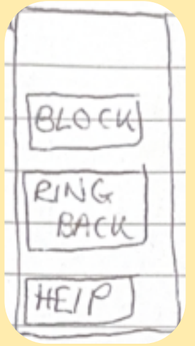
- Detects when user is on the phone via microphone/pressure sensor
- Listens for high-risk scam language e.g "password".
- Mutes hearing temporarily to prevent users giving sensitive information out
- Once call ends hearing returns
- Intervenes in potential scam situations, addressing the alertness concern discovered audiences research
- Useful for older users like **Georgie**

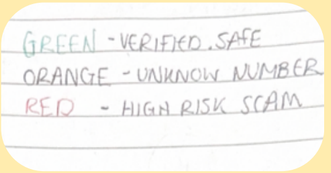


SafeCall Hub

Method - Future Envisaging

- Plug in-hub connected to landline
- glows red/orange/green based on call risk
- Speaker to verbally warn user
- Includes remote with easy options for block number, call back verified number, get help
- Optional mobile app version with similar features
- Helpful for **Linda** and her work calls



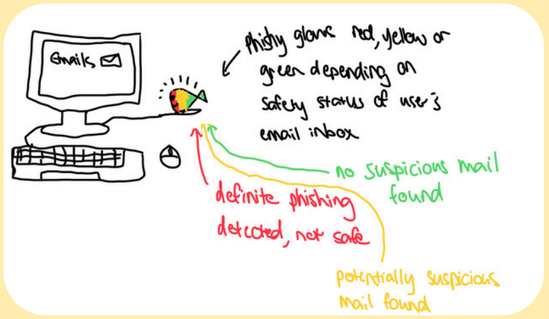


Email Phishing Scam Ideas

GlowPhish Email Scam Detector

Method - Lateral Thinking

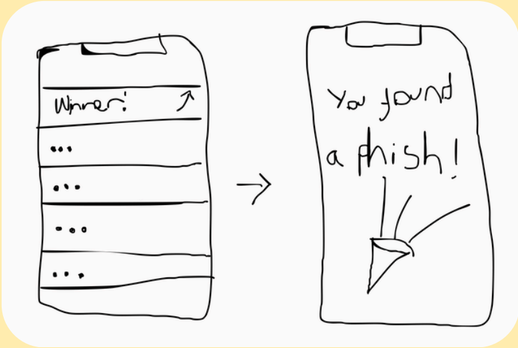
- Physical device “Phishy” that plugs into desktop
- Actively scans user’s mail inbox to detect suspicious activity
- Green glow indicates lack of suspicious emails
- Glows yellow to alert user of possible scams
- Glows red upon detecting suspicious activity that is likely to be a phishing scam
- **Linda** would benefit from this at her desk



GoPhish Email Scam Game

Method - Gamification

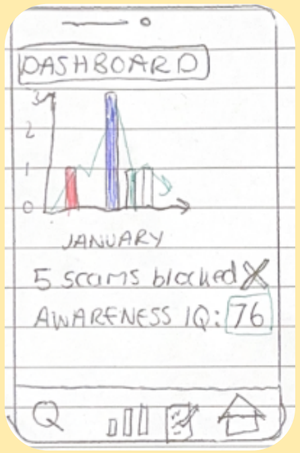
- The app sends random spam messages into the users’ inboxes
- If they click the link, they lose points
- If they forward them to the app, identifying that it is a scam email, they win a point
- Doesn’t inform users explicitly but aims to educate users over time
- Useful for a corporate setting for **Linda**

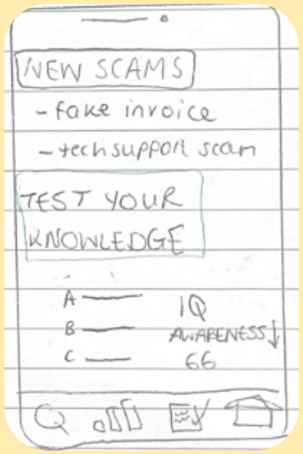


Scam Educator Dashboard

Method - Gamification

- Updates users on new/emerging scams using OSINT data gathering/webscraping, an aim of **Marcus**
- Quick knowledge test to build confidence
- Shows scams blocked and awareness score over time





Six Hats Technique

After reviewing each of our initial ideas as a group, we used the **Six Hats Thinking Technique** to eliminate some ideas and refine others, allowing us to look at our ideas and ensure they meet our essential criteria.

Intuition and Emotions

- “This app just feels like nobody would actually use it ”

- Danny (About gamified learning app)

Processes and Procedures

- “We need to make sure we are talking more about personas needs for this design”
- **Due to not having 6 members, we all shared this role in addition to other hats**

Optimistic Judgement

- “This idea is great because it covers so many scam types”

- Asia (About digital security assistant)

Risk and Caution

- “It needs to be more reliable, false positives can be annoying for **Georgie** causing him to lose trust in the whole product”

- Ilya (About safe call hub)

Creative Ideas

- “We can take this idea even further, like an authorisation device specifically for payments, even integrating with mobile banking apps”

- Ava (About verified identity app)

Logical Thinking

- “Not everyone drinks hot beverages so this wouldn’t work for everyone”

- Steph (About drink and think mug)

*Syed, S. A. (2022) AI-POWERED CYBERCRIME: THE NEW FRONTIER OF DIGITAL THREATS. International Journal of Engineering Technology Research & Management, vol. 6, no. 2. Available at: https://d1wqtxts1xzle7.cloudfront.net/122826710/AI_POWERED_CYBERCRIME_THE_NEW_FRONTIER_OF_DIGITAL_THREATS-libre.pdf

Sketches Evaluation

Purpose

To **explore** the **design space** as much as possible, using different ideation techniques: **Impossible Combinations**, **Future Envisioning** and **Lateral Thinking,** before committing to a direction. Sketching allowed us to **bring-to-life ideas** quickly, visualise possibilities and **“fail fast”** to **avoid prematurely narrowing down the solution**.

Contribution

Each technique generated different ideas: from impossible combinations **pushing the boundaries**, pairing unrelated elements and future envisaging encouraged long-term thinking about **emerging technologies.** Whilst lateral thinking allowed us to **reframe assumptions.** The sketches allowed us to cover a broad range of ideas, revealing a diverse range of **intervention types** - delay, guidance, reassurance, verification.

Learning

We learned the solutions can target different areas and equally meet the problem statement, from **interrupting impulsive decisions** to improving understanding.
As the sketches varied a lot, using the **6 hats** highlighted which ideas aligned with **persona** needs and which were **too impractical** to continue with in a manner which was consistent for each design.

Effectiveness

Sketching was highly effective at **pushing the boundaries** of the problem space **beyond typical solutions**, even if some were less realistic and feasible.
However, as **for most designs each person only explored one scam type**, the range was narrower than it could have been with **everyone ideating across all scam categories.**

Influence on next stage

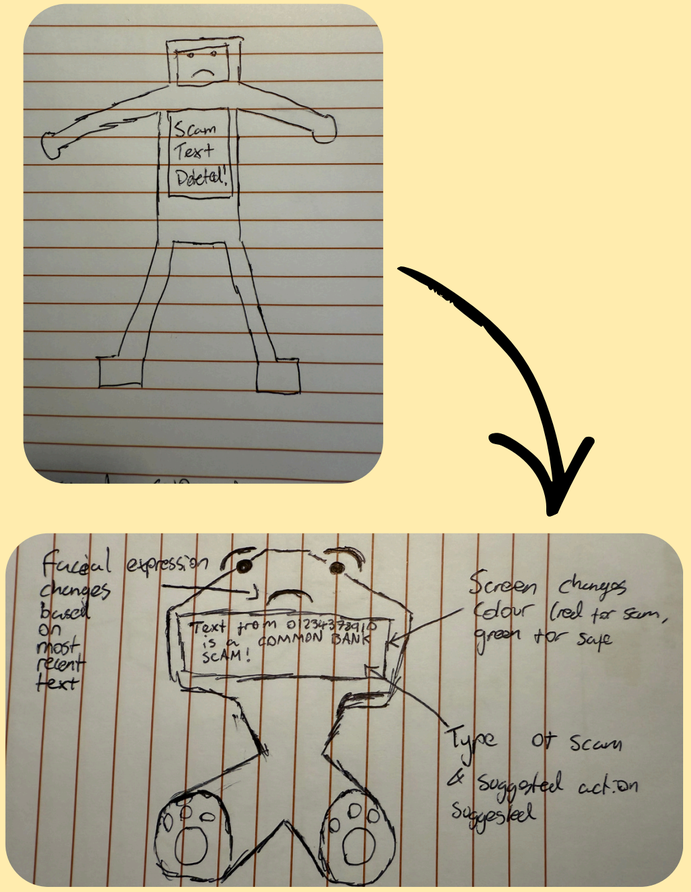
Next steps would be first **refining** the selected ideas then **applying QOC** to see how they meet criteria. The sketches helped us identify **core features** that we should carry forwards into our refined ideas.

TextBuddy Security Assistant

Features Description

- Using **future envisaging**, idea began as the concept of having a robot that intervenes if you have recieved a potential scam, taking your phone and deleting it
- Refined to a **mascot style** product, designed to sit to a desk
- Flashes **red** and looks sad if it believes it's detected a **scam text message**
- Shines **green** and looks happy if the message is **acceptable**, such as a requested verification code
- Will give a **suggested action**, such as delete the message/block the contact, or ask a follow up question. Otherwise the user is allowed to proceeed

Refined Iteration



Reasonings and Potential Limitations

- Fulfils **Linda's** needs, as she mainly uses technology at her work desk and the device gives **straightforward instructions** on how to act when a scam text arises
- Remains **discreet** when not detecting scam messages
- Accessible way for **Georgie** to understand Cyber Security
- Acceptable for **Jess**' as can quickly inform her if something is a threat

Limitations

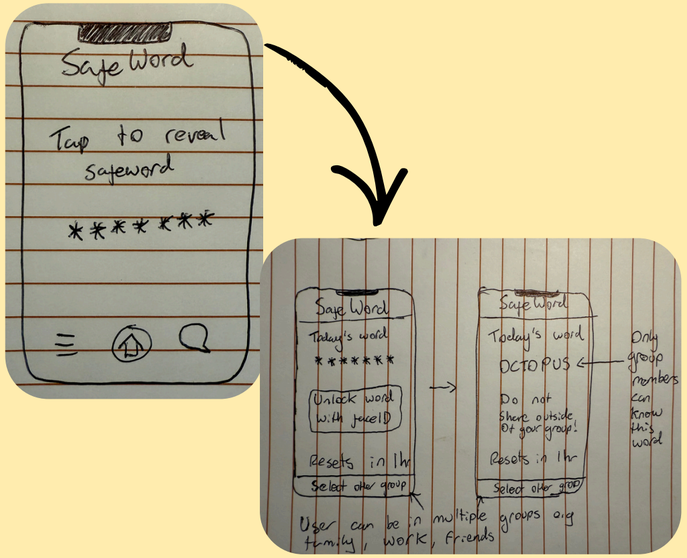
- Can't be carried everywhere, more suited for use at work/for elderly people
- May be annoying if flashing constantly
- Could be seen as slightly **invasive**

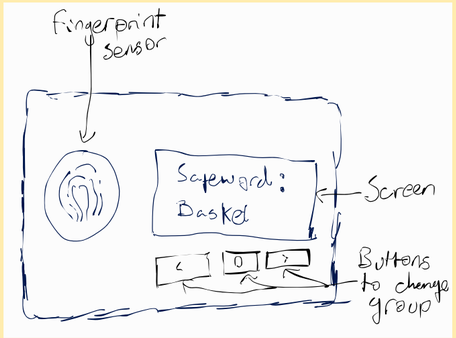
SafeWord Identification Tool

Features Description

- Addresses the '**Family Impersonation Scam**', which one of our interviewees fell victim to
- Developed from the desire to create a way to verify identity between multiple people
- Shared **'safe word'** that refreshes daily that users can use to verify they're speaking to each other
- Iterated to be a credit card sized device, utilising **fingerprint verification** to reveal the 'safe word'
- Can be used beyond text messages, as would work to combat deepfake and phone call scams
- Allows for unique safe words with multiple groups of people
- Safe word can be rotated

Refined Iteration





Reasonings and Potential Limitations

- Helps **Georgie** with his goal of staying in touch with family and friends without the fear or being scammed
- Addresses **Linda**'s pain points of not being able to maintain trust as she can confirm client's identities easily
- **Simple** and **accessible** design provides ease of use to all

Limitations

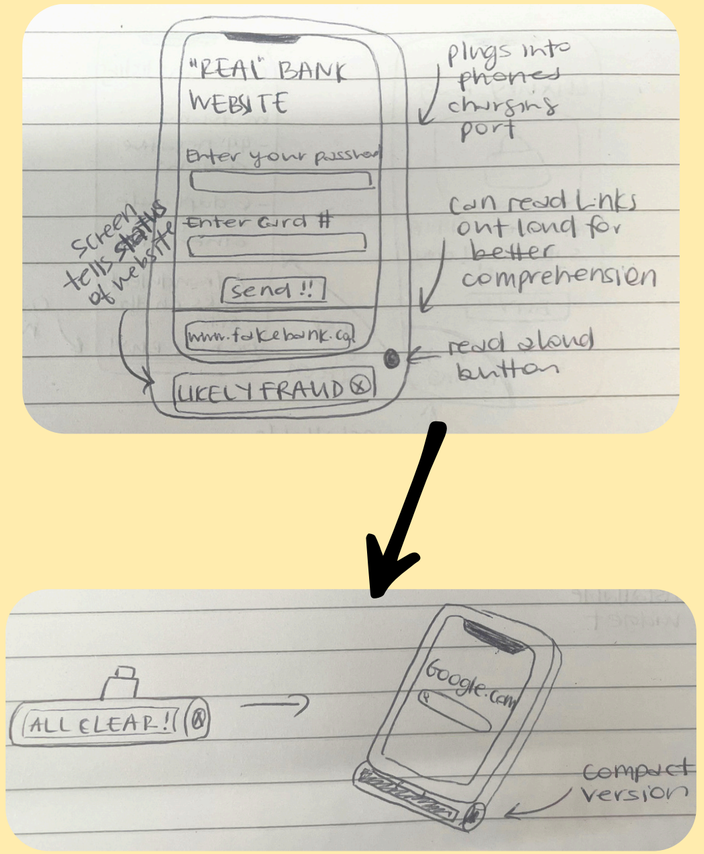
- Doesn't explicitly address Jess' needs
- Only works if the user has a suspicion they aren't speaking to who they think they are

Website Link Verification Phonecase

Features Description

- Original idea was a phone case that could read website links while you were accessing them and tell you if they were legitimate
- Iterated and refined to be a smaller, **more compact** device that has the same functionality
- Increases **ease of use** and make it more accessible, considering our older demographic
- Includes a **charging ability**, giving it a dual purpose and second motivation for use
- Accessibility features include the **'Read Aloud' function**, so users can double check spelling and website links themselves for peace of mind

Refined Iteration



Reasonings and Potential Limitations

- Addresses **Georgie**'s pain points and keeps the older demographic in mind by considering the weight of the device
- Directly solves **Jess**' needs for feeling safe while shopping online

Limitations

- Slightly **invasive** as it reads all the website links you access
- Only works if connected to the phone
- Doesn't directly address **Linda**'s needs or pain points

Digital Security Assistant

Features Description

- Digital and physical assistant that notifies you if your online safety is **compromised**
- **Physical keychain** component allows you to chat with it to learn more about how to stay safe online
- Performs **background screen scans**, and places an informative overlay when it detects something suspicious
- Fun **‘blob’ mascot** gives more information on the scam, what the key signs were to prevent future instances
- Soft **keychain** includes LED light, charging port and **vibration motor for accessibility**
- Connects to mobile app to notify users even when they’re away from their phone
- Uses an LLM to create natural back-and-forth **conversation**

Refined Iteration

Reasonings and Potential Limitations

- Communication with physical device directly addresses issues defined in **Scenario 1**
- Involves a solution to **Jess’** pain points on spontaneous shopping activities with an intervention method

Limitations

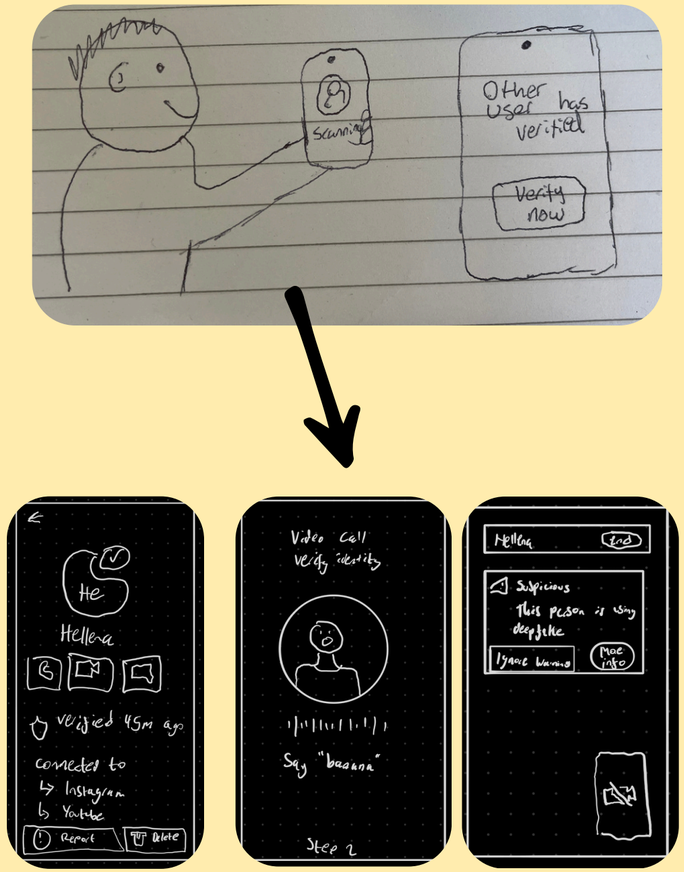
- Could be considered **invasive** or intrusive as it runs in the background and **reads all browsing activity**
- Doesn’t directly address **Georgie’s** needs
- Keychain needs to be carried with the user

Verified Identity Communication App

Features Description

- All in one phone call, messaging and video call app that can **verify your identity** via real-time analysis
- Users are required to scan their face to verify their identity, if they wish to join a video call
- Continuous **call analysis** that detects flags in real-time
- Scans for malicious text messages, hiding them from the user
- Makes sure you are speaking to who you think you are

Refined Iteration



Reasonings and Potential Limitations

- Solves problem of identity verification outlined in **scenario 1**
- Addresses **Georgie**'s pain points, allowing him to ensure he knows who he's talking to when chatting with family and friends
- Meets **Jess and Linda's** needs of immediate intervention

Limitations

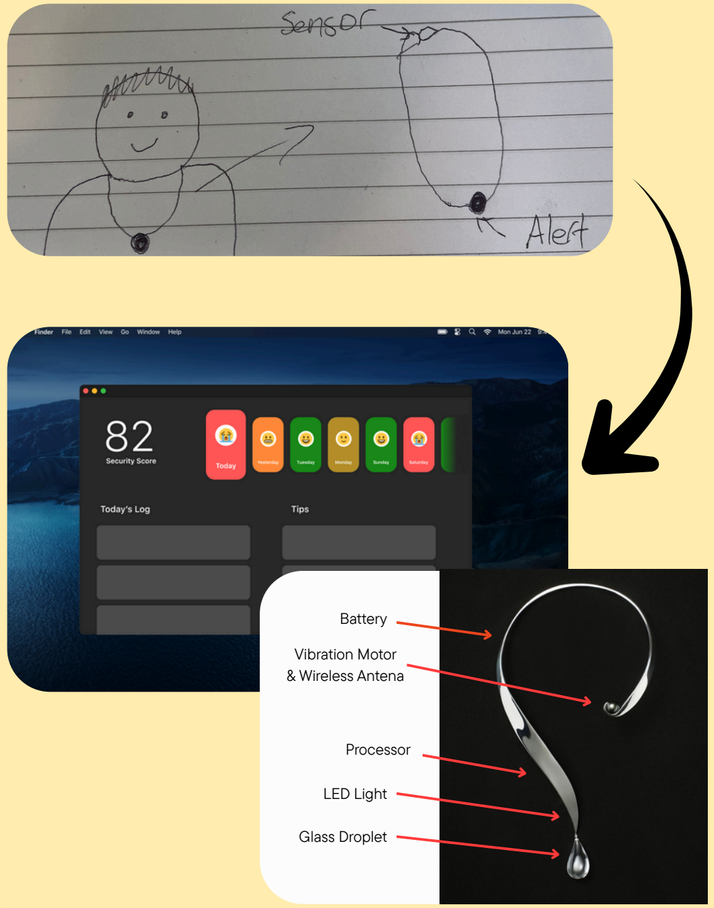
- Only protects communication that occurs through this application (i.e. can still get scammed on WhatsApp)
- Potentially invasive through face scans and voice verification

Intelligent Necklace

Features Description

- Necklace that connects to different devices and acts as a two-factor authentication device
- Sleek solid aluminium design provides a light and modern feel while wearing
- Quick wireless charging for continuous use throughout the day
- Included single glass droplet that glows via LED light to indicate severity of scam
- App runs in the background and provides a ‘security log’ based on scam insights and data footprint
- Provides interactive tips on how to improve security level and help the user to enable specific setting and understand their system

Refined Iteration



Reasonings and Potential Limitations

- Appeals to questionnaire responses, as it’s integrated with existing software
- Dual use with the Two-Factor Authentication capabilities
- Solves problem of ‘Lack of Alertness’ by being interruptive (flashing light)

Limitations

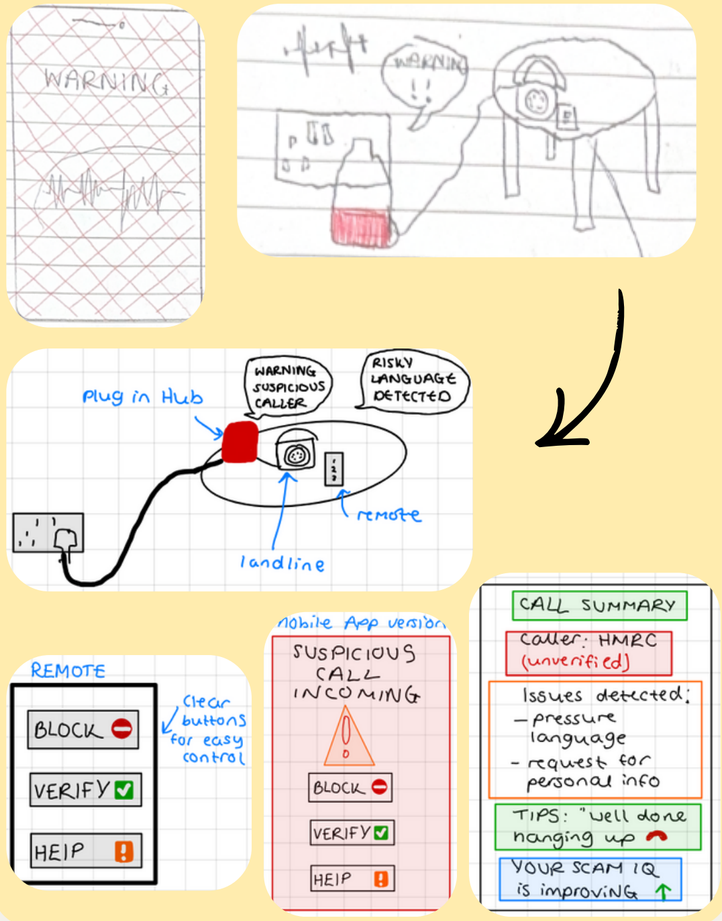
- Only alerts when the user is wearing the necklace
- Does not directly address **Linda** and **Jess’** needs
- Could be considered invasive as must be worn at all times to maximise security

SafeCall Hub Device, Scam Call Detector

Features Description

- A **plug-in smart hub** that connects directly to landline, as well as mobile, that **detects** and **warns** users through scam call situations in **real time**
- Provides incoming **call analysis** and in-call risk monitoring through the mobile app and physical hub
- Uses a colour ring for **risk signalling** for incoming calls
- **Speaker** warning the user “unrecognised number, answer with caution” and provides post-call responses
- Easy to use **remote control** with simple actions: block, verify and help
- **Mobile app** offers scam detection for mobile calls, post-call summary analytics to clearly explain risk signs on the dashboard

Refined Iteration



Reasonings and Potential Limitations

- **Accessible solution** for users who lack confidence with tech
- Addresses the issue in **scenario 2**, by assisting users in real-time
- Meets the needs of **Georgie** by giving **simple, non-technical reassurance**
- Advanced **call analysis** and consise post-call insights help **Linda** stay protected from modern scams without needing technical expertise/knowledge

Limitations

- Device only provides warnings, does not deliberately prevent/protect from cyber attacks
- Can be considered invasive as is listening to calls and reading phone numbers

GlowPhish Inbox Scanner

Features Description

- Physical plug-in device designed to assist with the **detection of malicious email activity**
- Uses scanning **software trained to recognise** scam emails
- Physical indicator depicts email safety - **glows red/yellow/green** depending on the status of the email
- **Direct plug-in** to desktop
- Users are in control of deciding which emails are scanned by GlowPhish
- Option to **enable audio announcements** of email scan results

Refined Iteration

Reasonings and Potential Limitations

- Addresses **Linda's** needs by taking the responsibility off the users to be alert and detect suspicious emails themselves, giving them more time for their work
- Accessible via option for audio announcements and glowing indicators
- Easy to understand design doesn't require technological knowledge

Limitations

- Giving users the ability to choose which emails are scanned creates the possibility to allow some malicious emails through
- Doesn't directly address **Jess** or **Georgie's** needs or pain points
- Only protects when at the desktop

Questions, Options, Criteria

Question:

Which design best supports users in reliabily identifying and responding to scam attempts across platforms, especially during rushed, emotional, or high-urgency moments?

Criteria	Digital Security Assistant	SafeHub Scam Calls	TextBuddy Security Assistant	WebsiteLink Verification Phonecase	GlowPhish Inbox Scanner	SafeWord Identification tool	Verified Identity Communication App	Intelligent Necklace
Ease of use	✅	✅	✅	✅	✅	✅	✅	✅
Pre-emptive Intervention	❌	✅	✅	❌	✅	❌	✅	❌
Real-Time intervention	✅	✅	✅	✅	❌	✅	✅	✅
Non-invasive (privacy)	❌	❌	❌	❌	❌	✅	❌	❌
Accessibility	✅	✅	✅	✅	✅	❌	❌	❌
Future-Proof	✅	❌	❌	❌	❌	✅	✅	✅
Reflective	✅	✅	✅	❌	❌	❌	✅	✅
Reliable	❌	❌	❌	✅	❌	✅	✅	❌
Integration	✅	✅	✅	✅	✅	✅	✅	✅
Targeting Scam Type	Everything	Scam calls	Scam messages	Website verification	Email Phishing	Scam calls & messages	Scam calls & messages	Everything

Using QOC, we established from group discussion that the Verified Identity app met the most criteria and is the best option, giving targeted intervention for both **scenarios 1 and 2,** and meeting at least some needs of **all of our personas**. It also doesn't cater to our **anti persona Marcus** as it is an entry level application.

Where criteria are not met currently, we will aim to adapt and iterate the design to meet these.

QOC Evaluation

Purpose

Our purpose with QOC was to **explore alternate design options**, ways of building our solution, and prevent premature convergence.

Contribution

Our table mapped options against criteria which we justified at the beginning of the ideate phase, which **encapsulated key needs from the personas**. It allowed us to see gaps in solutions where it did not meet all criteria.

Learning

Many solutions did not take all the needs of the personas into account, e.g. simplicity, low confidence. We learnt that **reflective and interruptive features** aligned with our refined problem statement.

Effectiveness

QOC was effective at understanding to what extent our shortlisted designs met the needs of our personas. However, our QOC failed to follow the method rigorously. We only used one high-level question, and loose **subjective** criteria with a black and white scoring method. Generating **multiple questions** and comparing options feature-by-feature would have **expanded the design space** and allowed us to break down the criteria further e.g. “ease of use”. By not doing this, we risked overlooking **strong alternatives** and reinforcing **group bias** rather than achieving the balanced evaluation QOC provides.

Influence on next stage

Whilst the QOC process was flawed, it identified **the Verified Identity Communication App** as the option meeting the most criteria. The next steps will be lofi-prototypes.

Choice Rationale

Verified Identity Communication App

1 - QOC

This option had the highest number of criteria met to answer our question, and the ones it didn't address could be solved through iteration.

2 - Research

According to Experian, impersonation scams are the most common types of fraud, affecting 3x more people than the next most common, causing $2.2bn in losses each year.

3 - Questionnaire

62% of respondents were only somewhat or not confident in suspecting phishing/scam messages.

4 - Interviews

A number of comments arose around the topic of not being able to trust texts, communicating with hacked friends' accounts and the rise of deepfake/AI contact.

5 - User Feedback

On consulting with interviewees on idea, they thought this idea had an easy to use interface that instills trust in communications they may have.

6 - The Human Element

This option best addressed the 'heat of the moment' reaction that causes scams to work, by allowing both pre-emptive and live scam prevention.

7 - Personas

- **Jess** - The product works with existing apps and gives clear indicators of potential scams.
- **Georgie** - Simple way to stay safe online, expecially when speaking to friends and family.
- **Linda** - Can speak to her clients with confidence.

8 - Scenarios

- **Scenario 1 -** Intervenes by alerting user of a potentially malicious message and hiding it initially.
- **Scenario 2** - Identifies deepfake content on video call and advises user to hang up immediately.

Feature-Persona Weighting Analysis

We evaluated the features our chosen idea required by using a Feature-Persona Weighted Priority Matrix*, to ensure we were viewing this from the perspective of our stakeholders, which are represented by our personas. As we had multiple personas with different pain points and goals, we decided to give each persona numerical weightings to balance each of their needs. We then multiplied each feature by its given -1 to 2 appreciation score* to get each features importance level.

	Georgie	Jess	Linda	Marcus	Weighted Sum
Weightings	*40*	*40*	*20*	*0*	*100*
Identity Verification	2	1	2	1	160
Real-Time Scam Detection	**2**	**2**	**2**	**1**	**200**
Cybersecurity Information	1	1	1	0	100
Ease of Use	2	1	2	1	160
Real Time Intervention (Notifications)	**2**	**2**	**2**	**0**	**200**
Local functionality	1	1	1	1	100

Georgie: **40 weighting**, because he is one of our two majority personas, representing the older, digitally illiterate generation.

Jess: **40 weighting**, as she is the other majority persona, representing naive users with problematic behaviours.

Linda: **20 weighting**, has less influence as her needs are already encapsulated by Georgie and Jess, but still needs representation as an affected persona.

Marcus: **0 weighting**, as he is our anti-persona so we're not marketing to him.

-1: Harms the Persona
0: Does not affect Persona (positively or negatively)
1: Benefits the Persona
2: Essential feature for the Persona

*Stoll, J., McColgin, D. & Gregory, M. (2008) Persona-weighted feature matrix excerpt. In: Stoll, J., McColgin, D., Gregory, M., ... & Edwards, W.K. 'Adapting Personas for Use in Security Visualization Design', Chapter in conference paper. Available at: https://www.researchgate.net/figure/Persona-weighted-feature-matrix-excerpt_fig2_226135040

Ideate Reflections

Aims

- Aim to explore different forms of alerts in the **design space**
- Consider **different scam types** individually, then seek how solutions can be adapted or combined
- Use different **ideation techniques** to create a broad set of primitive ideas
- Evaluate solution effectiveness **afterwards** using personas and scenarios
- **Seek value** in all ideas proposed and avoid ruling out ‘stupid’ ideas

From **Define**

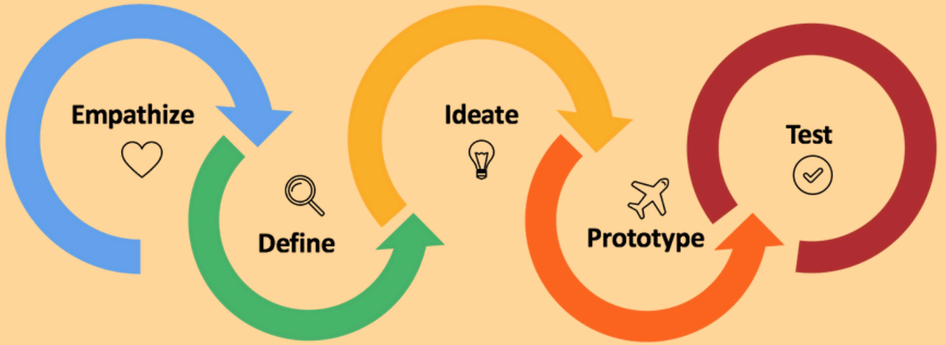
Learnings

- Creative thinking techniques like **future envisaging** vastly helped in formulating novel and unique ideas; a mix of physical and app-based designs
- **QOC** worked well, but **could have been split into more questions**
- Lots of solutions ended up catering to multiple scam types
- Some of the more **‘out there’** ideas helped us **consider features** to implement, such as the ‘drink and think’ mug with a time-based physical intervention
- **Collaboration sparked creativity** in creating and iterating

Next Steps

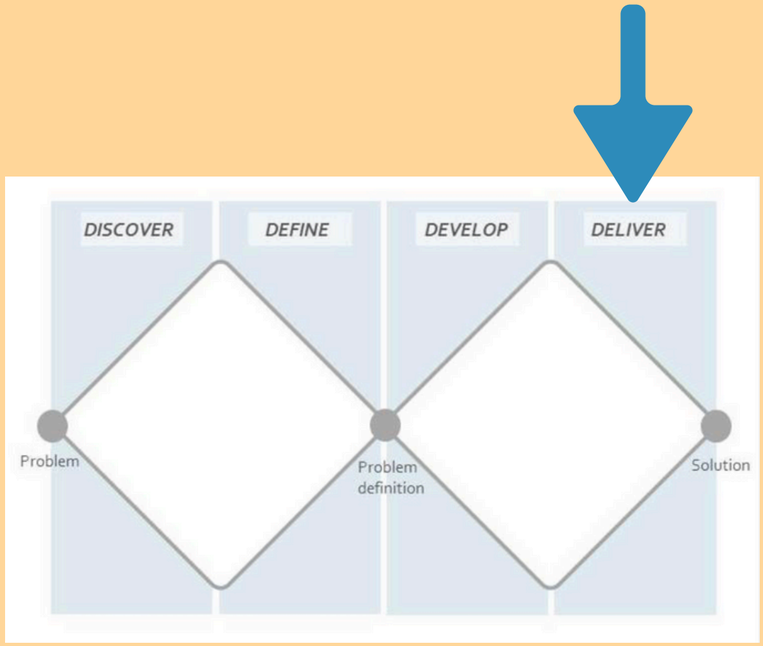
- Continue to **iterate and refine** the product in ways which meet user requirements
- Keep good design principles, such as the ‘**10 design principles**’ in mind when prototyping
- Keep **accessibility and ease of use** as a priority for **core functionality**, with scope for more advanced features for more capable users. This means less tech savvy users such as **Georgie** are understood

To **Prototype**



Prototype

Bringing ideas to life

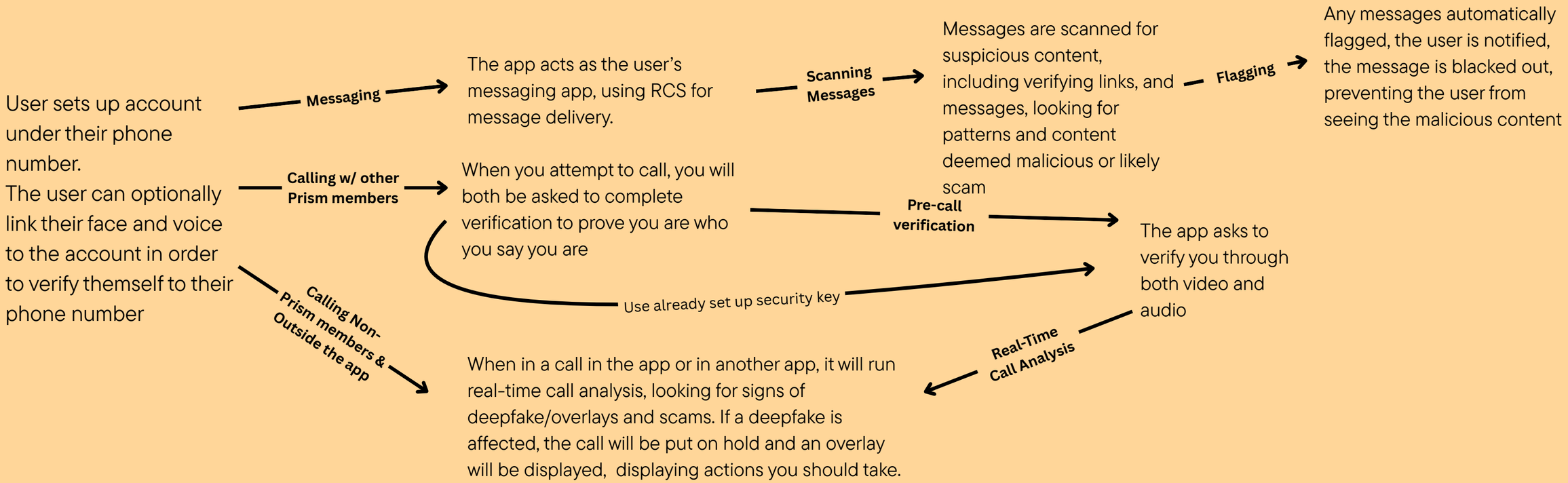


Chosen Idea

Prism Connect

Formerly “Verified Identity Communication App”

Initial App Feature Breakdown

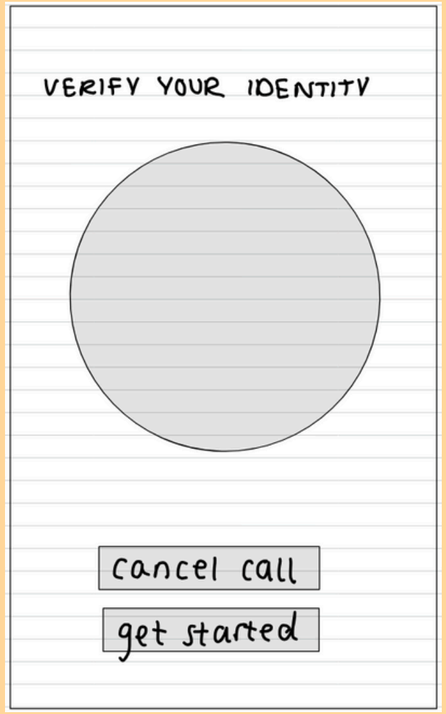


Lo-Fi Prototype

Spatial Structure of the App

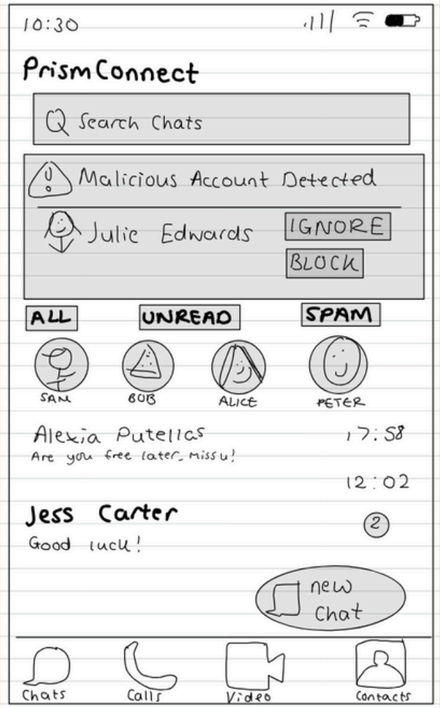
Card-based prototypes showing various screens of the app. Mapping the **spatial structure** of the app allows easy walkthroughs.

Identity Verification Screen



Lets the users **verify** their identity using face and voice biometrics, preventing impersonation and deepfake scams, **reducing reliance on user's judgement** during high-pressure calls.

Main Chat Page

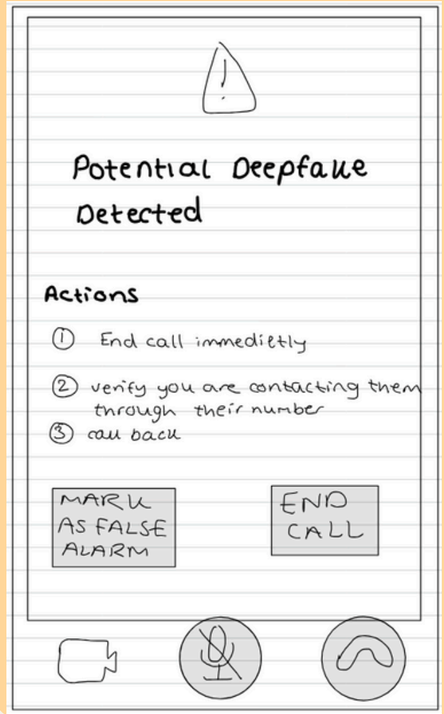


The main page flags **suspicious contacts** and messages early, **reducing reliance** on users noticing risks only when **"something feels off."**

Malicious Message Detected Screen

The message page which hides **malicious** content. It **reduces reliance** on **outdated cues** and prevents rushed responses.

In-call detection



During a call, if suspicious indicators are detected, the call will be paused. This forces the user to **reflect** and interrupts **impulsive decisions.**

Incoming call detection

The app will warn users when a number is unverified, telling them to decline the call or answer with caution. This interface is very **clear** and allows users to make **safer decisions**.

Key Prototypes

We decided to design a physical tag to pair with our app to streamline verification with the aim to improve **ease of use**, one of the most important criteria for **Georgie, Linda and Jess**. It is used to verify identity the user's in seconds after initial setup.

To setup the tag, a user completes an **identity scan in the app** which assigns the tag to the user. Their identity can be confirmed at any time by simply **using their fingerprint** and **scanning the tag** with the user's mobile device. Below shows how this design meets the **10 design principles**.

Lo-Fi

Hi-Fi

Innovative
A new and faster way to verify identity and instil trust in communications.

Makes the product useful
The tag links with the app harmoniously to deliver the verification service

Aesthetic/Unobtrusive
Simple elegant key fob that adds minimal volume to a user's keychain. Comes in multiple colours to meet users' tastes

Helps understand a product
Fingerprint icon guides gives a visual prompt on how to use it

Honest
Biometric fingerprint data stored locally on key, not shared elsewhere.

Durable/concerned with environment
Casing made from recycled plastic. Its small size and simple design limits number of components. Uses an ultrasonic fingerprint sensor so it can withstand water and dirt.

Consequent to last detail
All aspects considered form usability, form factor and integration with the app

Little design as possible
Has only the minimum required features to assist the user when using the app

Lo-Fi Prototype Evaluation

Purpose

For **testing** out ideas, to evaluate and get feedback. To **answer questions** such as the spatial structure, how the **app flows** and what the product **looks like**.

Contribution

Card based screens helped us **visualise** the app's spatial structure and simulate high-pressure moments. The physical key-tag prototype provides a simple way of verifying your identity using biometrics, and how **external verification can improve ease of use**.

Learning

The security key provided reassurance for our persona **Georgie**, providing verification in a simple way. We understood the importance of considering materials to be environmentally friendly and to follow **little design as possible** to aid ease of use. We also recognised the need to **prevent heat-of-the-moment responses** through **time delays** on risky messages and to better explain **in-call-scam-detections**.

Effectiveness

Our lofi-prototypes captured the app's flow partially and the security key's feel, though the key did not take into account a battery, so **dimensions** may **not** have been **accurate**. Our card-based prototype remained digital, limiting the **tactile interaction and card rearrangement.** Limited screens restricted the combinations of flows to occur.

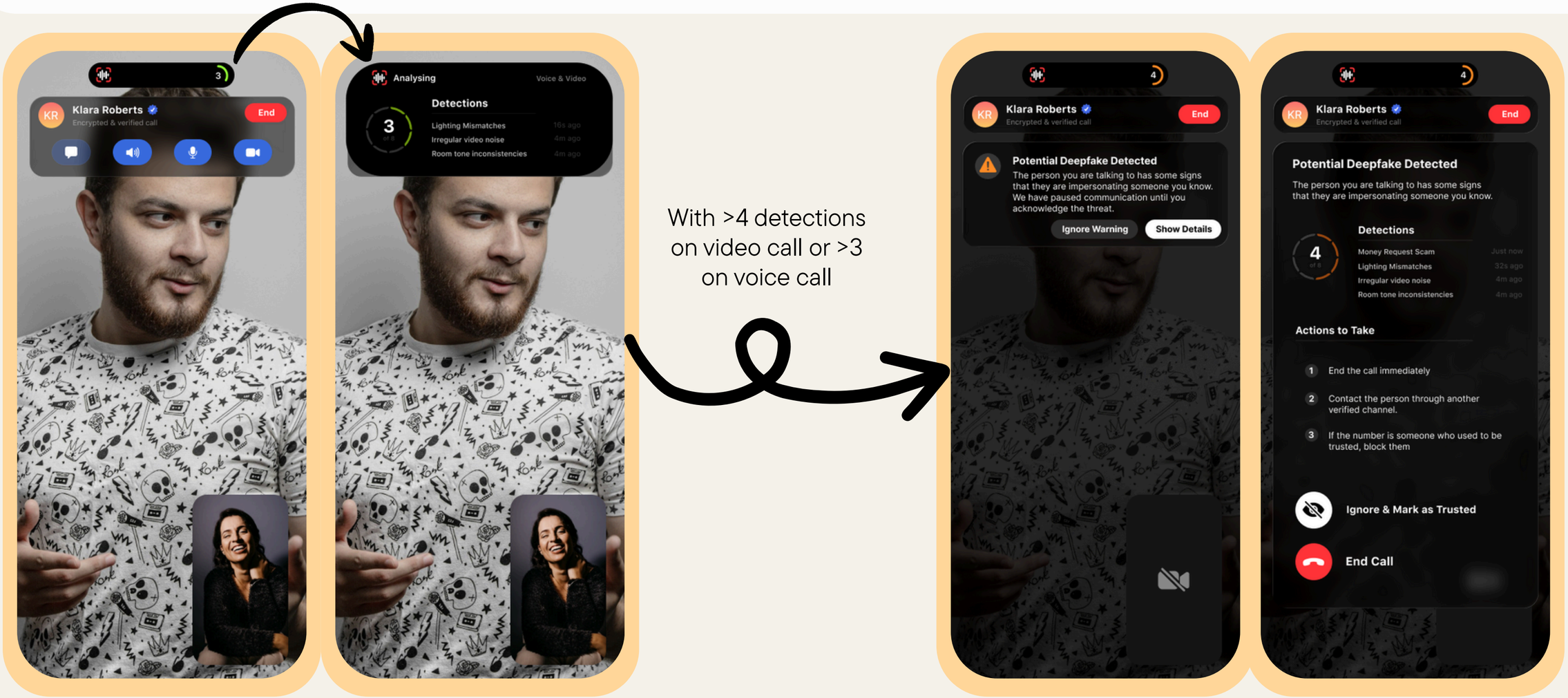
Influence on next stage

The Lo-Fi prototypes directly shape the Hi-Fi design. The next steps are to refine these into full wireframes, add explanations to the **in-call detections** to improve user understanding for **Linda**, and strengthen **message-blocking** to stop impulsive replies from **Jess**.

Wireframes

Refined Design

Iterating from our lo-fi prototype, we decided to display which signs of deepfake or scam so users know why the interaction may be malicious. This helps them learn the signs and stay up to date, as needed by **Jess and Linda**. The breakdown makes it as easy as possible for **Georgie** to comprehend this more advanced topic.



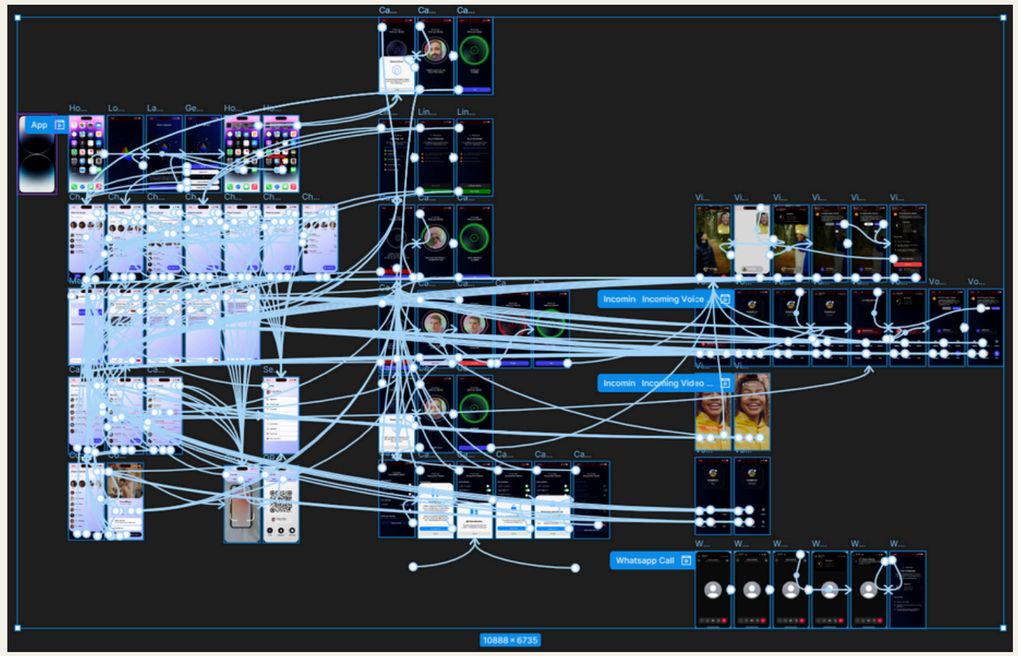
Wireframes

Final Version

The wireframes were further built out in full and interactions were added to make the the hi-fi prototype, allowing for usability testing and feedback.

Try the Demo

Screen Flow Diagram



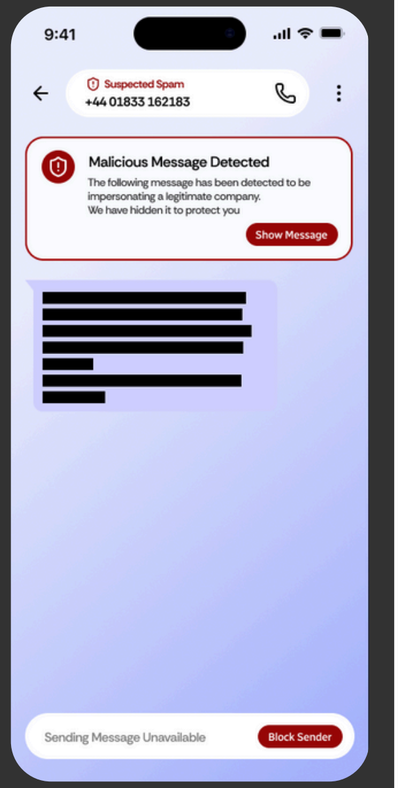
Features Summary

Universal features work with existing apps, whereas native features work within PrismConnect

PrismConnect

Universal Features

SMS Detection

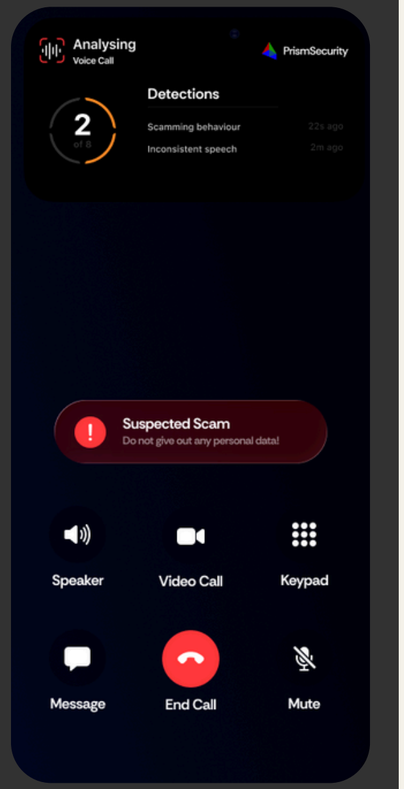


- Hides suspicious texts for 10s and warns users of threats within to prevent hasty actions

Mutual Verification with our Security Key

- Set up the security key with a single face scan and fingerprint to assign your identity to the key
- Send family or friends a notification to request verification of their identity and verify yours
- Trust the end user on any app or phone call

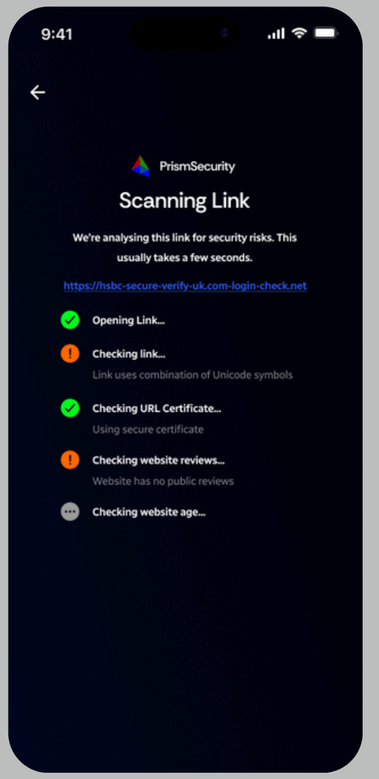
Live Call Analysis



- Listens to live incoming video and voice calls and alerts if the caller has malicious intent

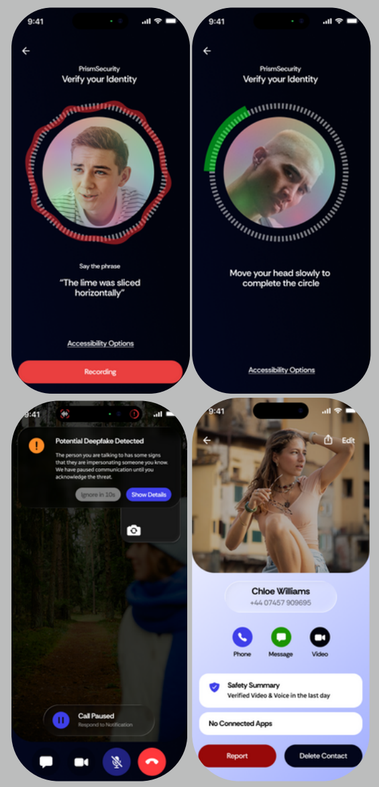
Native App Features

Web Link Verification



- Clicked links have their authenticity checked in a safe, virtual instance

Video & Voice Scanning



- Users who don't wish to use the key can scan before each call with voice and face scans

Functional Details

The defining factor for PrismSecurity is that it is **able to detect threats, either before or during an attack**. These methods include pre-call verification, real-time call analysis and message analysis. This aims to fix **Scenario 2**, by preventing the scammer from using the child's voice.

Verification Methods

Voice

Dynamic spoken pass-phrase

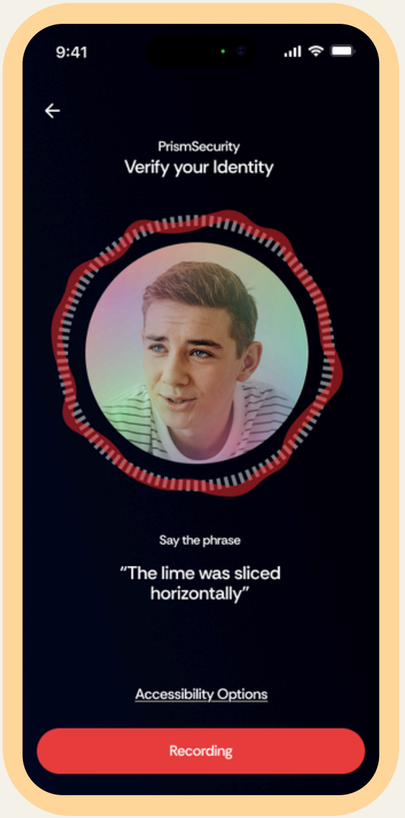
- User is prompted to speak a phrase while the camera records their face.
- Phrase is unique per verification attempt, generated by a randomisation algorithm similar to time-based 2FA codes.
- Random and unusual phrasing reduces risk of: Brute-force attempts (limited to three attempts), and replay attacks, because attackers cannot predict or reuse past phrases.

Detection of synthetic voices

- Synthetic voices typically pass through a neural vocoder, which leaves identifiable artefacts in a mel-spectrogram. (Voice modulation artifact)
- Spectral inconsistencies between natural and generated audio are analysed to distinguish genuine speech from model-generated speech (inferred from typical anti-spoofing techniques).

Facial-speech synchronisation check

- Camera captures lip and jaw movement during the spoken phrase.
- System compares audio timing with facial movement to detect mismatches typical of deepfake overlays and situations where a real voice is used with a fabricated or altered face.



Verification models are designed for incremental server-side updates, allowing users receive improved detection methods automatically without requiring configuration, helping **Jess and Linda** stay up to date with scams with little additional effort.

Sun, C., Jia, S., Hou, S. and Lyu, S. (2023). AI-Synthesized Voice Detection Using Neural Vocoder Artifacts. [online] openaccess.thecvf.com. Available at: https://openaccess.thecvf.com/content/CVPR2023W/WMF/html/Sun_AI-Synthesized_Voice_Detection_Using_Neural_Vocoder_Artifacts_CVPRW_2023_paper.html.

Agarwal, S., Farid, H., Fried, O. and Agrawala, M. (n.d.). Detecting Deep-Fake Videos from Phoneme-Viseme Mismatches. [online] Available at: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w39/Agarwal_Detecting_Deep-Fake_Videos_From_Phoneme-Viseme_Mismatches_CVPRW_2020_paper.pdf.

CatalyzeX (2016). Data Generation Using Pass-phrase-dependent Deep Auto-encoders for Text-Dependent Speaker Verification. [online] CatalyzeX. Available at: https://www.catalyzex.com/paper/data-generation-using-pass-phrase-dependent [Accessed 28 Nov. 2025].

Functional Details

Video

Video calls starts with video verification, followed by voice verification, ensuring multi-factor validation before a call is approved, and significantly reducing the chance of deepfake. This feature addresses **scenario 2** directly, preventing the scammer from using the child's picture.

Head-rotation challenge

- User rotates their head in a circular motion.
- Protects against: Static photographs and Low-quality 3D models unable to relight or deform naturally.

SAFE method (for devices without LIDAR)

- Users must maintain eye contact while moving their head.
- Detects crude 3D impostors by comparing gaze stability with head motion — something low-detail models cannot emulate.
- Developed in response to UK regulations requiring robust face-verification methods; early commercial systems had weaknesses this method aims to mitigate.

Playback-attack detection

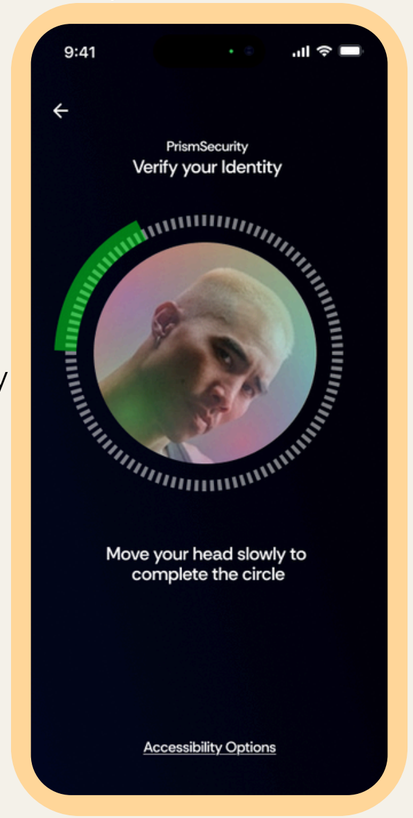
- System uses a CNN-based RGB matrix analysis
- Identifies signs of a pre-recorded video played back to a camera (e.g., screen refresh artefacts, moiré patterns).

Warping and noise analysis

- Deepfake models often generate a face that must warp artificially when head pose changes.
- CNN checks for edge warping, texture flickering, and noise patterns inconsistent with real skin under motion.

Blood-flow (aliveness) detection

- Subtle changes in specific red channels are measured to verify real physiological signals.
- Prevents attackers from passing verification using photos, videos, or static masks.



Becattini, F., Bisogni, C., Loia, V., Pero, C. and Hao, F. (2023). Head Pose Estimation Patterns as Deepfake Detectors. ACM Transactions on Multimedia Computing, Communications, and Applications. doi:https://doi.org/10.1145/3612928.

Boehm, A., Chen, D., Frank, M., Huang, L., Kuo, C., Tihomir Lolic, Martinovic, I. and Song, D. (2013). SAFE: Secure authentication with Face and Eyes. doi:https://doi.org/10.1109/prisms.2013.6927175.

Li, Y. and Lyu, S. (n.d.). Exposing DeepFake Videos By Detecting Face Warping Artifacts. [online] Available at: https://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Li_Exposing_DeepFake_Videos_By_Detecting_Face_Warping_Artifacts_CVPRW_2019_paper.pdf.

Morescalchi, D. (2022). Intel Introduces Real-Time Deepfake Detector. [online] Newsroom. Available at: https://newsroom.intel.com/artificial-intelligence/intel-introduces-real-time-deepfake-detector.

Functional Details

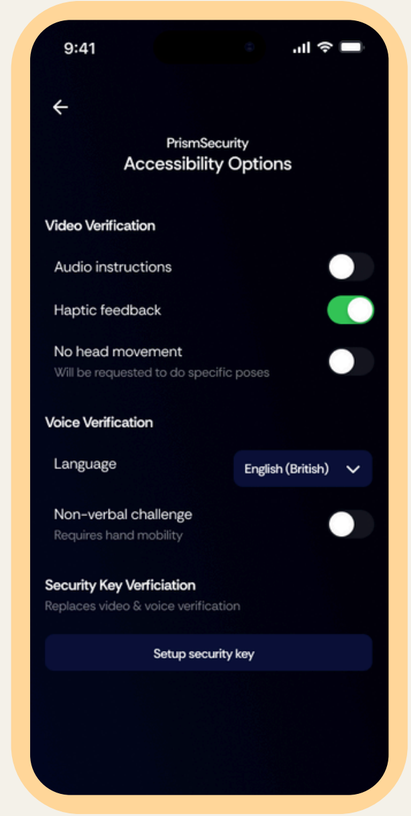
Accessibility Options

Audio guidance

- Real-time spoken instructions help users align themselves with the camera.
- Particularly important for users with limited or impaired vision.

Haptic feedback

- Device vibrations assist users in moving through scanning positions.
- Intensity increases when approaching unscanned regions.
- Can be disabled to accommodate users with pain or instability (e.g., arthritis).



No-head-movement mode

- For users with limited mobility such as **Georgie**
- Replaces rotation tasks with a sequence of facial expressions (e.g., smile, shock).
- Tests for warping artefacts, aliveness signals and 3D shape consistency derived from 2D deformation during expressions.

Multiple language support

- Instructions and spoken components are provided in languages selected by the user.
- Ensures accurate phrasing for voice-based challenges.

Non-verbal challenge mode (for non-speaking users)

- User performs hand gestures or simple hand-to-face interactions.
- Hands are difficult for generative models to reproduce accurately due to complex articulation and occlusion.
- CNN analyses gesture shape, motion features and physics-based interactions between hand and face.

Functional Details

Prototype

Security Key Verification

The security key must be FIDO-certified to ensure the hardware follows secure cryptographic protocols and cannot easily be spoofed. This would help **Georgie** as his eyes get worse and will struggle more to read details. It also helps **Jess and Linda** with quick authentication with minimal disruption to their usual online activities.

Initial identity binding

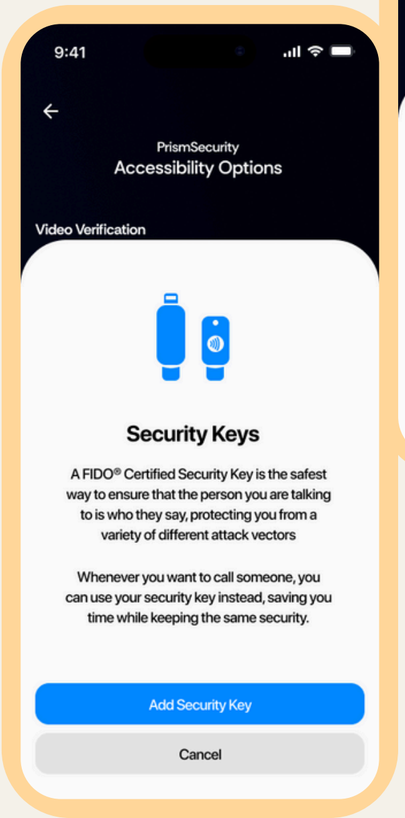
- User completes full video + voice verification.
- System stores mathematical encodings (embeddings) of the user's face and voice locally.
- These embeddings are linked cryptographically to the security key.

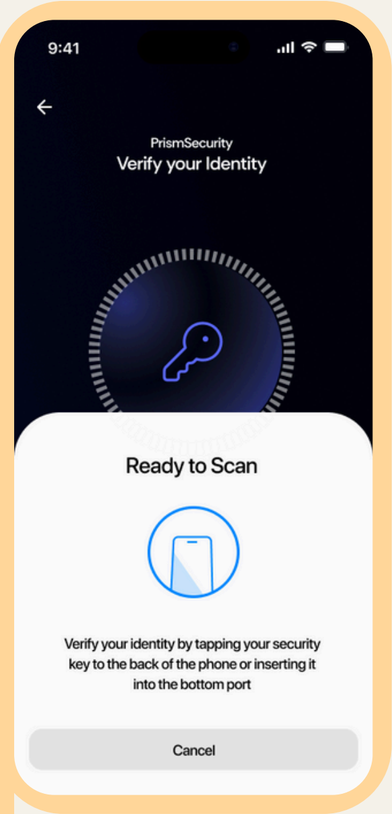
Ongoing biometric checks on key use

- A short burst of video frames is captured.
- System verifies:
  - Match against the stored identity.
  - Aliveness via micro-variations in blood-flow channels.
  - RGB matrix artefacts to detect video playback rather than a live face.

MAC-based security key validation

- On each authentication attempt, the device checks:
  - The integrity of the key using its Message Authentication Code.
  - Whether the key is linked to one or more stored identities.
- Allows a user to associate multiple keys with their profile.





Functional Details

Real-Time Analysis

Intermittent clip analysis

- Periodically captures short audio–video segments during the call.
- Runs the same deepfake, spoofing, and liveness checks used in pre-call verification, to check one hasn't been activated mid-call or conditions have changed.

On-screen detection indicator

- A small but noticeable banner appears at the top of the screen.
- Integrates with normal phone calls for ease of use for **Jess, Georgie and Linda**.
- Shows that real-time analysis is active, the current count of detected suspicious indicators.
- The indicator becomes progressively more visible as more warning signs accumulate.

Escalation threshold

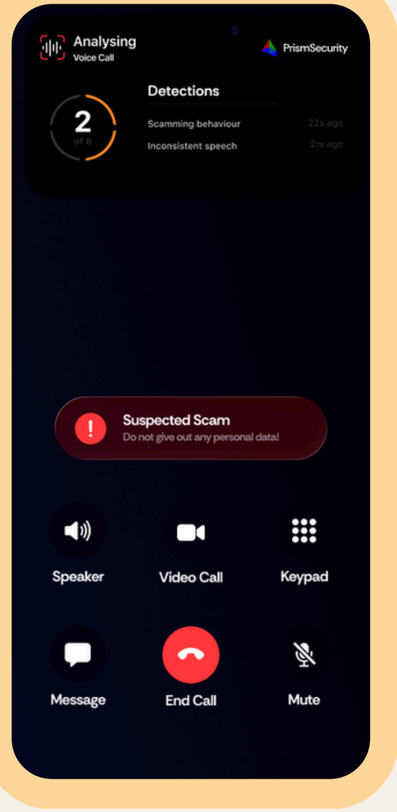
- When 3–4 signs of potential scam or spoofing are detected, the call is automatically paused, a notification with a customised alert sound is played, in the app, the popup cannot be dismissed without user interaction - based on behavioural research, it has been shown that passive dismissal leads to missed warnings.

Audio and video behavioural analysis

- Audio and video streams are analysed independently for scam-like behaviour: Examples include unnatural pauses, scripted scam patterns, or visual signs of identity spoofing.
- Model is trained on a large curated dataset of known scam interactions.

User-submitted data

- Users can choose to contribute recordings of scam attempts - These samples expand the dataset and improve detection robustness over time.



Wang, J., Shi, J., Wen, X., Xu, L., Zhao, K., Tao, F., Zhao, W. and Qian, X. (2022). The effect of signal icon and persuasion strategy on warning design in online fraud. Computers & Security, 121, p.102839. doi:https://doi.org/10.1016/j.cose.2022.102839.

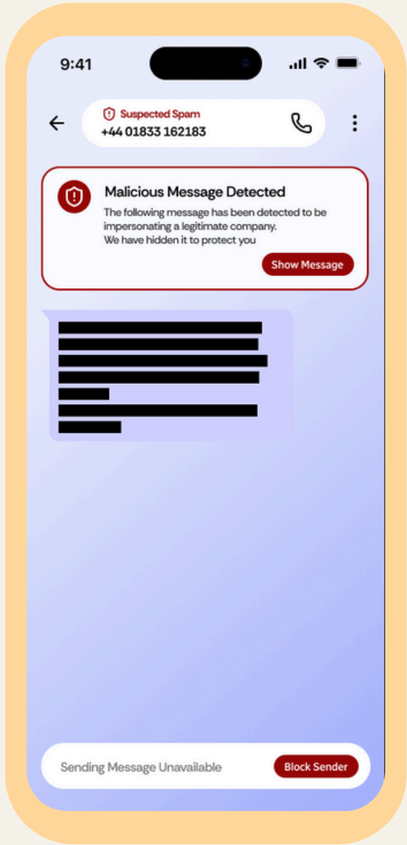
Vassilev, V., Donchev, D., Tonchev, D. and Василев, В. (2021). Impact of false positives and false negatives on security risks in transactions under threat. [online] repository.londonmet.ac.uk. Available at: https://repository.londonmet.ac.uk/6776/.

Zhou, S., Liu, X.F., Nah, F.F.-H., Harrison, S., Zhang, X., Zhen, S., Yeung, D., Hsiao, J.H., LC, R., Chan, A.B., Wang, X., Jiang, C.L., Lin, F., Li, J., Wong, A.W.-K., Chan, L.L.-H., George, B. and Li, P. (2024). Understanding and Fighting Scams: Media, Language, Appeals and Effects. Lecture notes in computer science, pp.392–408. doi:https://doi.org/10.1007/978-3-031-76821-7_27.

Functional Details

Message Scanning

After our Lo-Fi card prototype, we recognised the need to **further reduce impulsive behaviour** so we added delays to responding for 10s, **inspired** from the mechanism from the **think and drink mug**. It directly addresses **Scenario 1** and common scam types by preventing heat-of-the-moment reactions for impulsive users like **Jess** and explaining why messages are harmful, supporting users like **Georgie**.



Transformer-based message classifier

- Incoming messages are evaluated before the user sees them.
- Model is trained on a constantly updated database of verified scam messages.

Spam filtering

- If a message exhibits strong spam-like characteristics:
  - Sender is immediately routed to the spam folder.
  - The conversation is muted.
- Limits user exposure to low-effort scam attempts.

Confidence-based hiding

- Messages with medium or high scam confidence are hidden from view and replaced with a warning prompt.
- User must acknowledge the warning before continuing, preventing passive reading.

User awareness

- Warning prompt explains which features caused the detection.
- Helps the user recognise similar patterns in future interactions.

Transmission blocking

- User cannot send further messages until the warning is acknowledged.
- Reduces the chance of continuing engagement with a scammer.

Wash, R. (2020). How Experts Detect Phishing Scam Emails. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), pp.1–28. doi:https://doi.org/10.1145/3415231.

Jensen, M.L., Dinger, M., Wright, R.T. and Thatcher, J.B. (2017). Training to Mitigate Phishing Attacks Using Mindfulness Techniques. Journal of Management Information Systems, 34(2), pp.597–626. doi:https://doi.org/10.1080/07421222.2017.1334499.

Functional Details

App Privacy

- Messages sent within the application (not SMS/RCS) are **stored using the Signal protocol**, a secure end-to-end system where **only the recipient and sender are able to see the message**.

- When a user adds a security key, their account is linked to that key and the user is asked to do **face and voice verification**. The app sends the video and voice to the server, after which, the server generates a **mathematical representation of both the face and the voice** (hashed) to be referenced in the future. This result is then stored in the database.

- For voice and video verification, the server runs checks and **never stores the result**, keeping it in RAM and **deleting it within 10 minutes**.

- Real-time analysis uses **both local and cloud based detection**. The phone itself, runs checks every 10s while the call is running. **The server runs constant checks**, working in the middle to run heavier models to detect for scams.

Cohn-Gordon, K., Cremers, C., Dowling, B., Garratt, L. and Stebila, D. (2017). A Formal Security Analysis of the Signal Messaging Protocol. [online] IEEE Xplore. doi:https://doi.org/10.1109/EuroSP.2017.27.

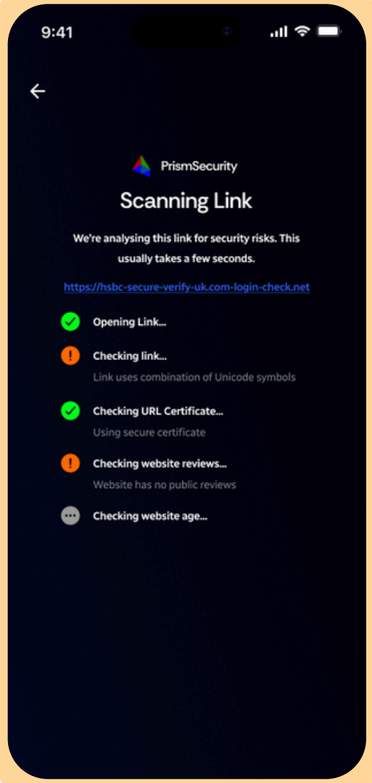
Your paragraph Wikipedia Contributors (2019). Face ID. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Face_ID.

Functional Details

Website Scanning

Virtual browser container

- When a user taps a link in the application, the system loads the site inside a controlled sandbox.
- This environment is isolated from the main system to prevent:
  - Malware execution.
  - Access to cookies or local storage.
  - Device fingerprinting.



Real-time status updates

- The user is shown progress indicators explaining:
  - What is being checked.
  - Why certain checks are necessary.
- Transparency helps reduce user impatience during security examination.

Detection results and mandatory delay

- If red-flag indicators are found (examples inferred: phishing patterns, fake login portals, malicious scripts):
  - Warnings are displayed with a summary of the detected risks.
  - A 10-second mandatory wait is enforced before allowing the user to continue.
- This pause helps disrupt impulsive clicking behaviour often exploited by scam links, for **Jess'** online shopping and impulsiveness.

www.scamadviser.com. (n.d.). Scamadviser.com | check a website for risk | check if fraudulent | website trust reviews |check website is fake or a scam. [online] Available at: https://www.scamadviser.com/.

Scam Prevention Summary

Feature	SMS Scams	Deep fakes	Phone Call Scam	Website Scams	Impersonation	Phishing
Voice Verification	✗	✓	✓	✗	✓	✗
Video Verification	✗	✓	✗	✗	✓	✗
Message Scanning	✓	✗	✗	✗	✓	✓
Website Scanning	✓	✗	✗	✓	✓	✓
Secure Key Tag (Physical Product)	✓	✓	✓	✗	✓	✗

The above matrix shows **how the combination of our features address the scam types in different ways** of concern from our research. All work to **prevent impersonation**, the main purpose of the app. The physical key can, once set up, be used to prevent scams from the start. The other scanning and verification features allow for **ongoing safety** in an interaction, and real time alerts on texts and call threats to educate the user, meaning it caters to a variety of technical proficiency from **Jess to Georgie.** Phishing is not explicitly covered in the current version covered as it was deprioritised due to being a lesser threat.

Prototype Reflections

Aims

- Continue to **iterate and refine** the product in ways which meet user requirements
- Keep good design principles, such as the '**10 design principles**' in mind when prototyping
- Keep **accessibility and ease of use** as a priority for **core functionality**, with scope for more advanced features for more capable users. This means less tech savvy users such as **Georgie** are understood

From **Ideate**

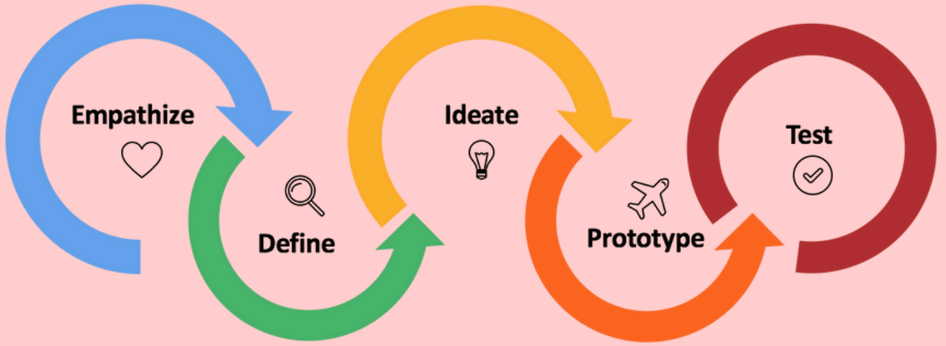
Learnings

- Addition of extra **in person collaboration** sessions helped us align goals and bounce ideas off each other
- Having **physical prototypes** in the form of a figma app and physical security key allowed us to **simulate usage to show a proof that the product works in practice**
- Rooting design in **good practice and principles** guided and made decisions when prototyping

Next Steps

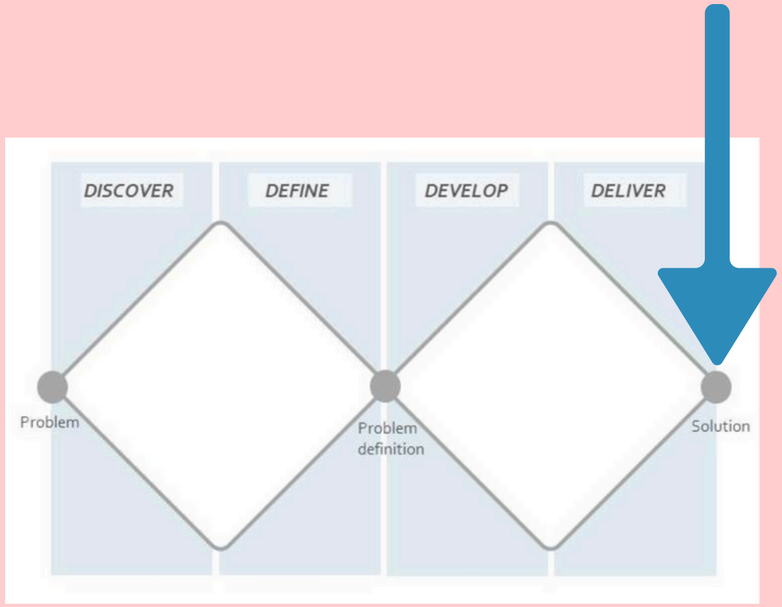
- Test the product **in-situ against scenarios** to see how well our product prevents the undesirable behaviour of falling victim to scams
- Gather **feedback from users** that we interacted with in empathise
- Evaluate the **product against personas needs** to see where we have improved their situation and where there are limitations
- Identify areas where there would be **scope for development** in a **future iteration** of the product

To **Test**



Test

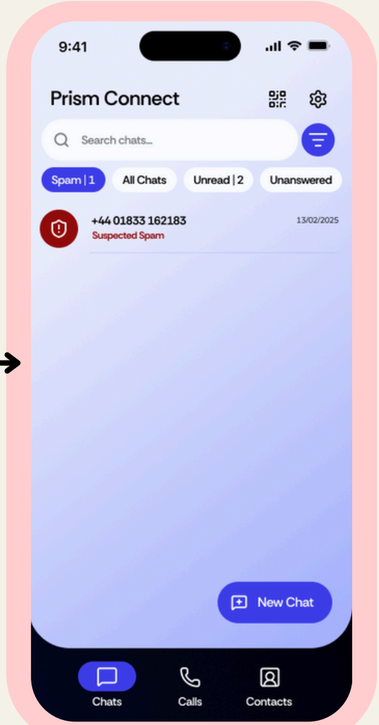
Bringing ideas to life



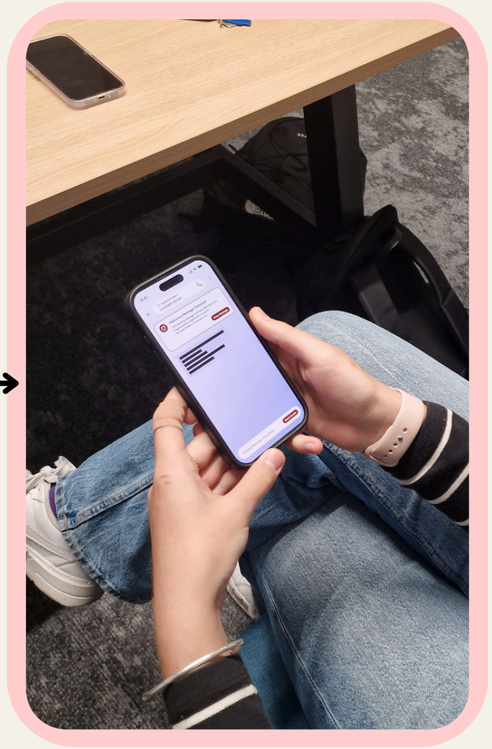
Scenario 1 Walkthrough

Jess is a student, scrolling on her phone. She has recently been ordering from various online shopping sites.

She receives a message from an unknown number, but it's automatically put in spam. They claim to be a well-known postage company requesting her details to deliver a parcel but Jess is unable to read it before reading the **warning**.



Before Jess can respond, she is notified about the potential scam, informed about the scam indicators detected and **restricted from replying** to the message until she acknowledges them.

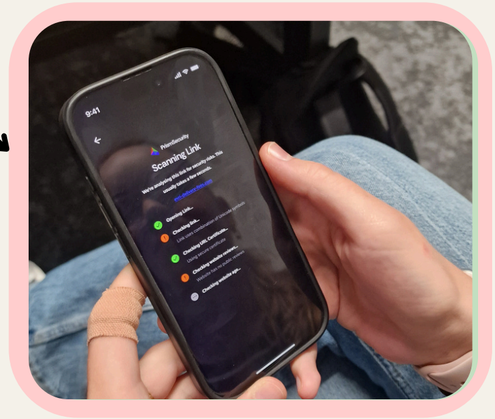


Ignores

Jess can quickly block the sender and is **protected** from the information theft scam.

Clicks Link

Jess clicks on the link and the app scans it and informs her that it is a scam site, **forcing her to wait 10s** to reconsider or go back to the app.



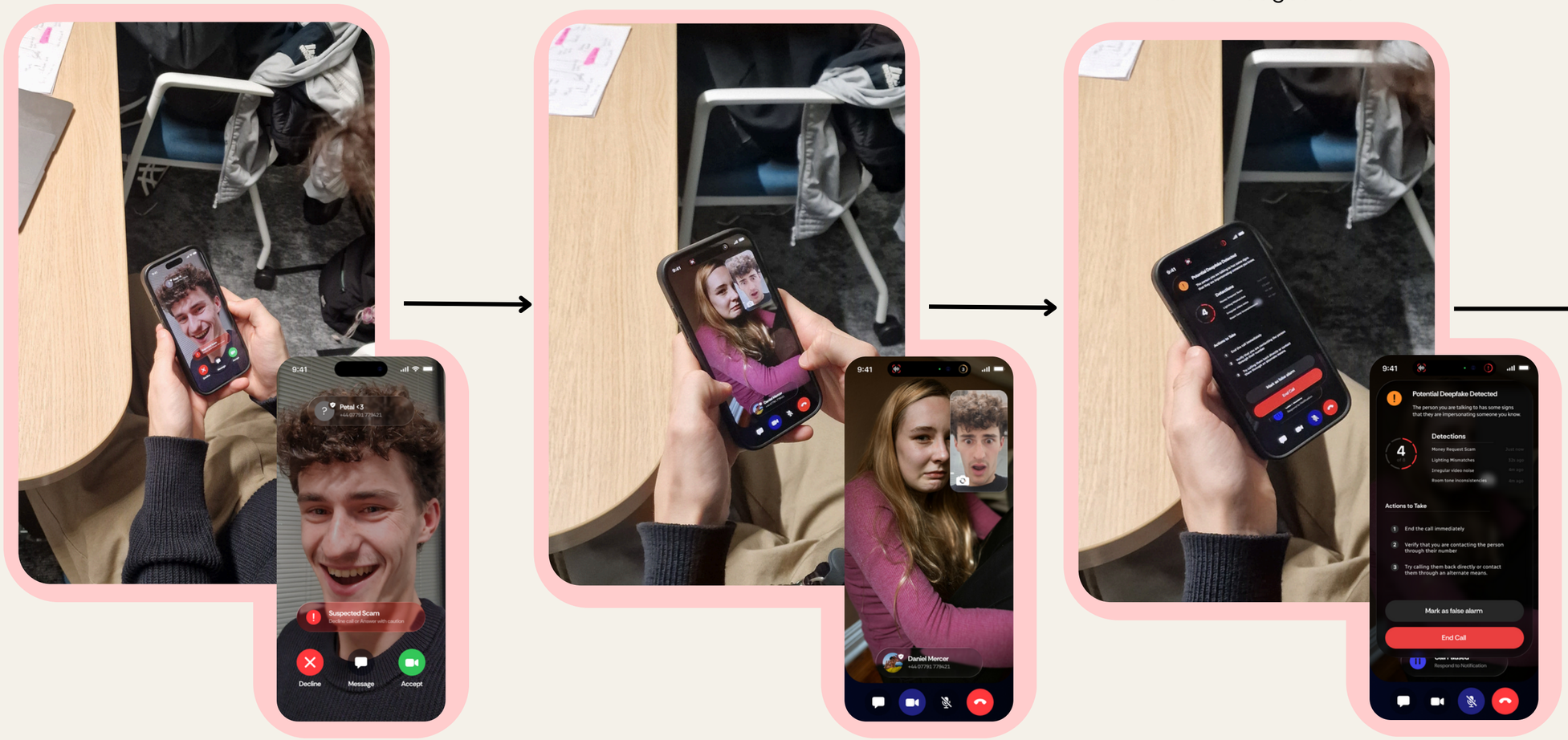
Scenario 2 Walkthrough

Georgie gets a video call on WhatsApp from a number claiming to be his child.

The video call begins, and Georgie's child says that they are in trouble and need money urgently.

The app **pauses the call** as it has recognised multiple **inconsistencies** that point to a deepfake scam, Georgie **cannot resume connection** until he acknowledges this.

Georgie ends the call and messages his real child to check if they're alright, confirming the call was a scam and everyone is safe.



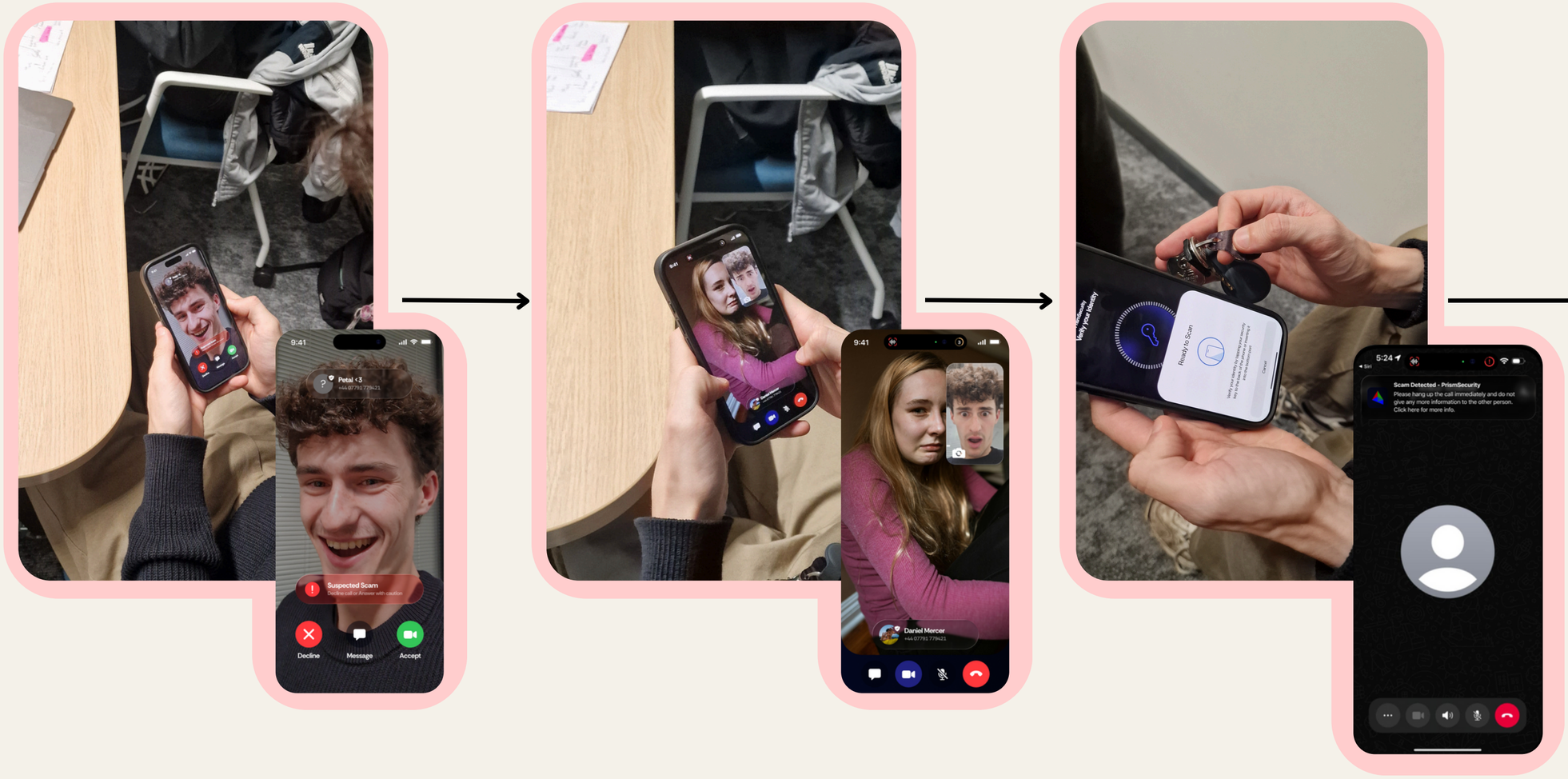
Scenario 2 Walkthrough 2 (Key)

Georgie gets a video call on WhatsApp from a number claiming to be his child.

The video call begins, and Georgie's child says that they are in trouble and need money urgently.

Georgie is informed by the app of the suspicious nature of the call. He **verifies his identity** using his pre-configured **security key**, scanning his **fingerprint** to activate. This allows his to send an **identity request notification** to his daughter to confirm if it is really her.

The impersonator of Georgie's 'Daughter' is unable to verify themself in the verification check, so Georgie hangs up and doesn't answer. He texts her daughter through the in app secure messaging to double check she is ok.



User Feedback of Product

Positive

"The overlays of warnings are good on the call and text feature as they infor m and alert me but still allow me to exercise my own judgement"

"I like all the guardrails in place to stop scams from unfolding, it seems like a comprehensive solution"

"The hiding of messages is a cool feature as it warns you to take caution before you see it, going into reading it with a different mindset"

"The design [of the app] is very professional looking, it feels to the standard of my normal apps like WhatsApp"

Constructive

"It's a bit hard to see how this would work with my normal texts in this [Figma] prototype, it would be good to be able to see other examples of texts coming in and out."

As it was not feasible to make functional custom text inputs in Figma, this may have prevented fully 'bringing the product to life' for some users with contrived examples.

"I struggled to find a few of the features without assistance and some of the buttons didnt do anything"

This showed some features may not be signposted well enough, but also is another limitation of Figma. Within our time bounds we couldn't get every single setting and detail created and linked in our prototype.

"The key fob is nice and small, I could see myself having this on my keyring, but almost forgettably small"

An unexpected by-product of making the fob small was that it wasn't noticeable. Having it on a keyring has the advantage of being always with you and discrete, but also now we know this may mean it isn't always front of mind for some users.

Think Aloud Testing

Purpose

We used think aloud testing on a couple of our users to see if their **mental model matches the expected functionality of the app**. Through this testing, we found there were **2 interesting quotes** from users that informed quirks in functionality that we weren't aware of before this exercise.

"I have picked up a call from the suspicious number, I am speaking to the scammer, I can't see the screen to see the alerts as my phone is on my ear"

Learnings

This showed us that users that don't pick up phone calls with headphones or on speaker mode would not be able to see the alerts from the app in call mode. Whilst not strictly to do with their mental model, it showed an important **oversight** in our design.

Next Steps

There should be **another cue** such as vibrations, flashing phone torch or speak-aloud function alongside warning popups.

"I am looking for the way to pair my security key, I am looking through the screens, I can't quite see where I am meant to go"

Learnings

The user was unable to locate the option to pair the security key, suggesting that what appeared clear in the prototype may **not align with user expectations** (e.g., finding it in settings). While this may partly reflect **limitations of Figma**, where not all buttons function across pages, it remains an important consideration for design clarity.

Next Steps

On app startup for the first time we should have added a popup with the option to add in a key or information of where to find it in the future.

Test Reflections

Aims

- Test the product **in-situ against scenarios** to see how well our product prevents the undesirable behaviour of falling victim to scams
- Gather **feedback from users** that we interacted with in empathise
- Evaluate the **product against personas needs** to see where we have improved their situation and where there are limitations
- Identify areas where there would be **scope for development** in a **future iteration** of the product

From **Prototype**

Learnings

- The design successfully intervened in both scenarios when testing
- Users were impressed by the design and completeness of scam prevention
- The figma prototype had some limitations where test users may have been expecting a fully working app
- There was a difference between one user's mental model and our app
- The solution prevents different scams in multiple different ways

Next Steps

- Consolidate the final design into a poster, showing at a glance the steps we took to get to it and show the features
- Conduct evalutaion and measure our achievements against our objectives
- Have project debrief with team to discuss how we felt the project went; what went well and what could have gone better.
- Apply learnings from the whole exercise to any design work in the future, putting users first

To **Evaluate and Beyond**

Evaluation

Reviewing and Critiquing our Designs and Methods

Personas Design Evaluation

As Georgie and Jess represented a large section of our target demographic, it was important that majority of their goals were met. The pitfalls we encountered came from specific functionalities we didn't directly include (i.e. online shopping scams), which could be implemented in a future expansion of the product.

Goal Met?

Not | Partially | Fully

Georgie

Stay in touch with friends and family without fear of being scammed

→ The app provides a way to confirm identity before beginning communication and alerts to suspicious texts

Continue to use apps as he usually does

→ The app includes integration with existing applications, allowing him to continue using apps as usual

Doesn't want to spend much money or any if possible for cybersecurity

→ This need is met as the application is free to download, and the optional security key would cost around £20, which in his price range

Wants a simple and accessible way to know cybersecurity without any technical jargon

→ Provides explanations for why interactions are flagged as scam, but does not directly teach cybersecurity

Jess

Continue to use phone as normal

→ The app does allows regular use of phone as before

Get clear, quick indicators if something she is engaging in could be a scam

→ The app provides clear and concise reasonings behind given detected scam

Be more up to date on new or emerging scams

→ As the app receives updates and learns from new scams, it informs users of these emerging scams

Know tell-tale signs of fraudulent links or websites

→ The app provides multiple reasonings behind the scam indicators it detects

Feel more confident when buying things second hand online

→ Partially met as it can verify website shopping links and seller's identity, but not their true intentions

Wants to know parents and grandparents are safe online

→ Can guarantee their safety if they're using the app

Personas Design Evaluation

Goal Met?

Not Partially Fully

Successfully meeting Linda's needs and not Marcus' needs shows that we have successfully created features that benerfit our target audience, instead of trying to appeal to the wider market. We understood that we could not create a one-size-fits-all product, and stayed within our chosen demographic's boundaries.

Linda

Stay protected online to avoid being scammed again

→ The app will keep her protected by blocking malicious messages and calls to prevent being scammed again

Maintain her professional reputation and trust between her clients and coworkers

→ The security key identity authentication allows her to maintain trust in who they're speaking and sending files to, helping in a corporate setting

Be informed and educated about new and upcoming scams - such as deepfake scams

→ Partially met as it does not directly inform her about scams but does explain tell-tale signs of deepfake scams if encountered

Wants a simple and straightforward way to stay safe online as doesn't have time for complicated installations.

→ The app installation is simple and pairing the optional security key doesn't require convoluted setup

Marcus - Anti-Persona

Maintain a high level of security control through his own knowledge and systems

→ Partially met as the app maintains high level security, but not directly through his own systems

Customise or automate security solutions rather than relying on guided tools

→ Not met as the app is fixed as guided, he cannot customise or automate his own solutions

Avoid interruptions or redundant alerts that slow him down

→ Not met as the app provides alerts and notifications if it detects a scam or threat

Stay informed about emerging attack trends through his work and OSINT

→ Partially met as does inform about emerging scams as they unfold but not more technical attacks he would find in his line of work

After referring back to our problem statement, we believe that for these personas, we've successfully addressed each of our individual needs. However, there are clear areas for improvement, specifically in the explicit education of scams and how to spot them.

Evaluation of Methods

Positives

Breadth of Methods

- **Multi-method** approach
- Brainstorming, research, SWOT, interviews, personas, scenarios, sketching, 6 Hats, QOC, Lo-Fi prototyping
- Built a **strong backbone** to **design process**
- Enabled **confident decision-making**
- Each method added a new **lens**

Brainstorming

- Exploration of design brief using brainstorming **was crucial** to uncover problems
- **Prevented jumping** to **typical** problems and solutions that fit the brief or problem statement

Personas & Sketches

- Translate **raw data** from research, interviews and questionnaires into **relatable**, **clear** user representations
- Personas **guided decisions** throughout design process
- Anchors design in **real behaviours** & needs
- Sketching **pushed creativity** and expanded solution space beyond the obvious

Prototyping

- Brought **ideas to life**
- Allowed us to **test flow** and **feel** of features
- **Guided refinement** towards the Hi-Fi design
- Understood what worked and what did not

Evaluation of Methods

Critiques

Research & User Input

- Improve sampling: questionnaire skewed to 18-24 which created a **sampling bias**
- Have more detailed scam-specific questions.
- Use **anti-persona** more throughout as he was not used in deliberation nearly as much as primary personas

Scenarios & User Behaviour

- We could have grounded our scenarios in real scam stories gathered from interviews and questionnaires.
- This would have made the scenarios **evidence-based** and **more realistic**, reflecting **real user experiences** situations and where intervention would have meaningfully helped

Ideation & Methodology

- **Broaden ideation**: each team member could have sketched ideas for multiple scam types to increase diversity of ideas
- Improve **QOC methodology,** use multiple, more specific questions and criteria to get a more **granular level of criteria** (e.g. breaking down ease of use) **reduce group bias** and **overlooking alternatives**

Prototyping

- Incorporate **technical constraints earlier** (e.g. battery size in security key)
- Move **beyond digital-only** prototypes to test different arrangements in the flow
- Iterate **higher-fidelity versions sooner** to reveal practical usability issues
- Consider other protoyping options that could be more complete as per user feedback

Evaluating Our Solution Against Identified Gaps

How our Final Design Addresses Key Areas for Development

We revisited the gaps identified from the background research and used this as evaluation criteria for our final prototype

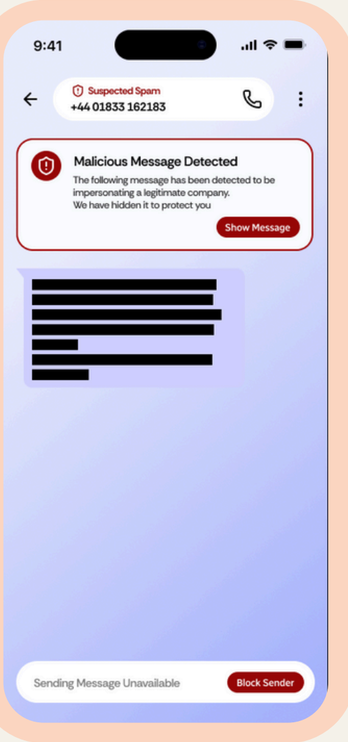
Areas for Development	How Our Final Prototype Addresses This
Reactive, not preventative	Possible scam text content is hidden from the user and they are warned why before they can see. Calls from suspicous sources are flagged
Limited Personalisation	Users can use as many or as little features as they feel comfortable, so more vulnerable users like **Georgie** could use the full suite of app
Poor Accessibility/ Readability	App uses plain language and bright colours like red to signal warnings for scam messages. Accessibility settings are baked in such as low movement face scans for **Georgie**
No reflection or follow-up	Malicious text content is explained and time block is applied to help impulsive users like **Jess** understand how responding could cause harm. Users are informed about why their calls or videos are likely to be scams.
Static examples and outdated training	The app receives updates so that new threats can be addressed, helpful for **Jess and Linda's** needs. The learning model also gets trained on interactions.
1-dimensional, single device	The presense of a physical security key means the app can be used across multiple mobile devices with a singular identity.

Usability Heuristics Evaluation

Heuristic Evaluation of the SMS Scanning Feature Using Nielsen's 10 Usability Heuristics

Interface	Issue	Heuristic(s)	Frequency	Impact	Persistence	Severity
Chat Filter Screen	Suspected Scam" label is red but small and could be overlooked	1. Visibility of System Status	2*	1	3	**2**
Chat Filter Screen	Filter "Unanswered" is unclear, texts may not need an acknowledgement	2. Match Between System and Real World	3	1	1	**1.6**
Scam Chat	"Ignore" vs "Hide" is ambiguous, the difference may be unclear	2. Match to Real World / 4. Consistency and Standards	3*	3	2	**2.6**
Scam Chat	Countdown timer ("Ignore in 10s") is confusing, why does it auto-ignore? What does it do?	3. User Control and Freedom	3*	3	4	**3.3**
Scam Chat	Sending message disabled, no explanation as to why	10. Help and Documentation	3*	3	3	**3**

Frequency 0 (rare) to 4 (common)
Impact 0 (easy) to 4 (difficult)
Persistence 0 (once) to 4 (repeated)
Severity = Total of F+I+P/3



*Frequency dependent on detection of scam messages/content

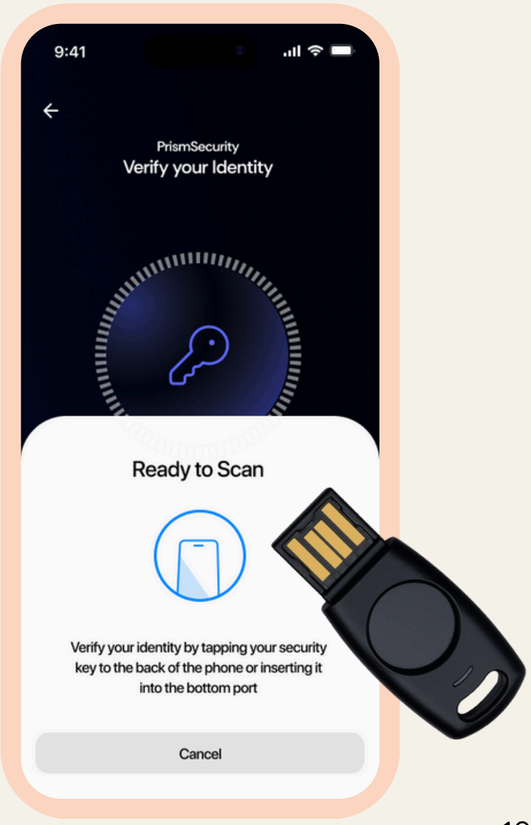
Nielsen, J. (1994) 10 Usability Heuristics for User Interface Design. Nielsen Norman Group.

Usability Heuristics Evaluation

Heuristic Evaluation of the Biometric Verification with our Security Key Feature Using Nielsen's 10 Usability Heuristics

Interface	Issue	Heuristic(s)	Frequency	Impact	Persistence	Severity
Add Security Key	The option to verify identity using the security key is hard to find within the app	3. User Control and Freedom	3	4	3	**3.3**
Security Keys Explanation	Popup alone does not clearly explain the purpose of the key: "you can use key instead of calling someone", the wording is confusing,	2. Match to Real World	2	2	1	**1.6**
Verify Identity	Unclear initially what verification involves for the user	1. Visibility of System Status	2	3	2	**2.3**
Security Key ready to Scan	Lacks error states e.g. key not detected, timeout	5. Error Prevention	3	4	3	**3.3**
Verification Complete	"Next" button could be more descriptive e.g. "Return to Settings"	2. Match Between System and Real World	2	2	1	**1.6**
Security Key Active	"Remove" is bold but there is no confirmation message to reduce accidental removal risk.	5. Error Prevention	2	3	2	**2.3**

Frequency 0 (rare) to 4 (common)
Impact 0 (easy) to 4 (difficult)
Persistence 0 (once) to 4 (repeated)
Severity = Total of F+I+P/3



Nielsen, J. (1994) 10 Usability Heuristics for User Interface Design. Nielsen Norman Group.

Usability Heuristics Evaluation

Purpose

Identify **potential usability problems** of PrismSecurity. Ensure features align with UCD principles.

How we conducted it

We independently evaluated the main 2 features using **Nielsen's 10 heuristics**. We then compared findings verbally and combined into one table. We **rated issues** by **frequency**, **impact** and **persistence**.

Our Focus

SMS Scam Filtering and **Security Key Verification** are the two main features due to persona needs and research. Focusing features ensured **depth** over **breadth**.

Evaluation of the method

Using Nielsen's 10 heuristics gave us a **clear structure** to **assess usability** and quickly **highlight problems**. The method was fast and easy to apply, but scoring can be **subjective**, and stronger expert input would have improved **reliability**.

Key Findings

- Risk information is ambiguous
- Missing system feedback about how the key works
- Poor visibility of critical system states e.g. no error states

Design Impact

- Add clearer explanation of risk and verification steps
- Simplify wording and improve consistency
- Add recovery states to confirm actions

Nielsen, J. (1994) 10 Usability Heuristics for User Interface Design. Nielsen Norman Group.

Future Expansion

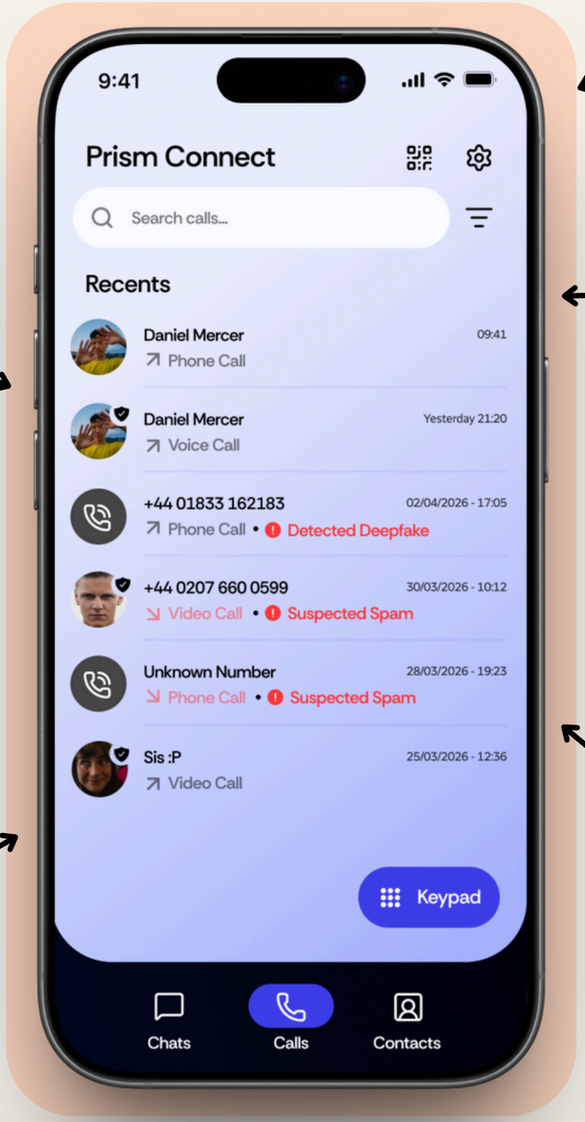
Additions informed form user feedback, testing, persona and heuristic evaluation

Phishing

- The app protects against **smishing** and **vishing** but not email scams.
- Future expansion could involve **automatic email forwarding** from unknown emails, so users are **protected from phishing**
- Any suspicious emails would appear in the app as spam and get deleted from the users inbox

Latest Scams Page / News

- A "**latest scams"** page would help users stay informed and aware
- **Notifications for major threats** could highlight **active scams** keeping the user **informed**
- Fully meets **Jess'** need of **staying up to date and informed**



Vibration when scam detected

- Based on user feedback, adding a vibration to get the user's attention could help notify

Marketplace Protection

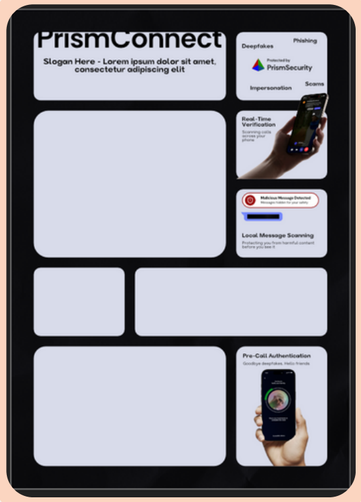
- Current implementation does not provide protection against **fake shopping sites**
- A feature to add **buyer/seller protection**
- This would address **Jess'** needs

Advanced Website Protection

- The website checking feature analyses surface-level features and reputation
- Technical users can **bypass detection** by exploiting safe websites or routing through trusted sites
- The checker dos not review website code or navigate through multiple pages
- Adding a **website crawler with code detection** could strengthen protection and close these gaps

Poster

- Our goal was to show how PrismConnect protects users in multiple ways.
- We studied feature showcases like Apple's WWDC layouts for inspiration.
- This led us to use simple blocks, strong visuals and short scenario highlights to communicate key ideas.



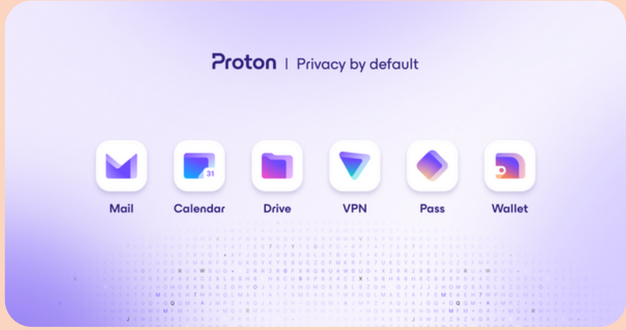
In Progress

- Final poster uses colour to separate features visually.
- Prioritises innovative features at the top of the reading order.

- Iterative design showed we needed more detail to meet criteria.

Final Poster

- Designed PrismConnect so it can be expanded in the future.
- Added "Protected by PrismSecurity" to create a recognisable sign to build trust
- Took visual inspiration from Proton's consistent privacy-focused design language.



- Featured the statistic "86% have received impersonation scams" prominently to capture attention.
- Structured flow to naturally guide the viewer through features, app UI and QR code.
- Included an Easter egg—"scammer" behind the phone—to subtly show scam detection.

Team Management

Working together to solve a problem

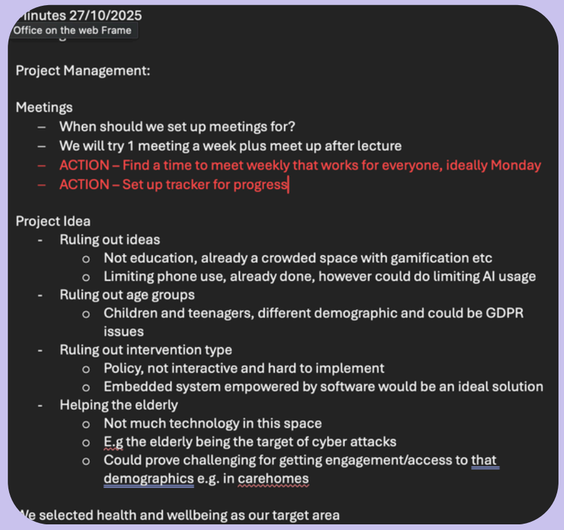
Facilitating Open Communication

Team Meetings

- Our team arranged to **meet twice a week**, holding one online session and one in-person session throughout the semester.
- We used these meetings as an opportunity to discuss workload, **share opinions** and ensure that we were all on the same page regarding our assessment.
- We had one person assigned to get **meeting minutes**, so that the actions could be seen at a later date.
- These acted as checkpoints to ensure we were hitting our aim of a stage per week, allowing us to adjust timelines if necessary

Communication Channels

- In addition to regular meetings, we held a team vote, deciding **which platform was most suitable** for regular communications.
- We decided upon WhatsApp, utilising this as a tool for arranging meetings, asking questions and staying on track.
- Additionally, we made use of **Canva comments** to provide feedback for individual slides, allowing for **easy group collaboration**.



inutes 27/10/2025
Office on the web Frame

Project Management:

Meetings
- When should we set up meetings for?
- We will try 1 meeting a week plus meet up after lecture
- ACTION – Find a time to meet weekly that works for everyone, ideally Monday
- ACTION – Set up tracker for progress

Project Idea
- Ruling out ideas
  - Not education, already a crowded space with gamification etc
  - Limiting phone use, already done, however could do limiting AI usage
- Ruling out age groups
  - Children and teenagers, different demographic and could be GDPR issues
- Ruling out intervention type
  - Policy, not interactive and hard to implement
  - Embedded system empowered by software would be an ideal solution
- Helping the elderly
  - Not much technology in this space
  - E.g the elderly being the target of cyber attacks
  - Could prove challenging for getting engagement/access to that demographics e.g. in carehomes

We selected health and wellbeing as our target area

Extract of Meeting 1 minutes

Improvements

- Our meetings could have run more smoothly if we had **consistently made use of an agenda**, rather than relying on each individual to bring up a discussion point.
- Could have assigned slides to review for each member each week to **reduce checking** later on.

Managing Tasks and Resources

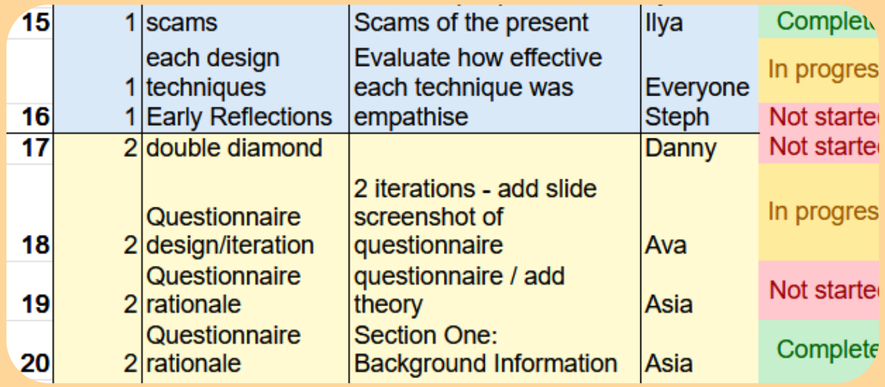
Sharing Resources

- Making sure that we all had **access to the relevant resources** was a priority for our group, and we catered to this need by creating a **shared OneDrive folder**, home to all of our group's meeting minutes, ideas spaces, research items and prototype sketches.

Meeting Minutes	October 27	Ilya Sullivan (MEng	5 items	Shared
Prototype SS	Tuesday at 11:3...	Ilya Sullivan (MEng	67 items	Shared
071125 - Meeting with Russell.docx	November 7	Ava Boomla (BSc C	15.6 KB	Shared
calendar.xlsx	A few seconds a...	Stephanie Franks (L	42.8 KB	Shared
Cybersecurity Questionnaire.xlsx	November 20	Stephanie Franks (L	36.6 KB	Shared
Exploring Problem Space.docx	November 6	Ilya Sullivan (MEng	59.1 KB	Shared
Problem-Breakdown.docx	October 30	Ilya Sullivan (MEng	29.7 KB	Shared

Managing Tasks

- To ensure **equal responsibility and accountability** for tasks allocated to team members, we created a live tracker using Excel.
- This tracker includes slide numbers, titles, brief descriptions, the name of the member responsible and a **dropdown progress measure** with options of "Not Started", "In Progress" and "Complete".



15	1	scams	Scams of the present	Ilya	Complet
	1	each design techniques	Evaluate how effective each technique was	Everyone	In progres
16	1	Early Reflections	empathise	Steph	Not starte
17	2	double diamond		Danny	Not starte
18	2	Questionnaire design/iteration	2 iterations - add slide screenshot of questionnaire	Ava	In progres
19	2	Questionnaire rationale	questionnaire / add theory	Asia	Not starte
20	2	Questionnaire rationale	Section One: Background Information	Asia	Complete

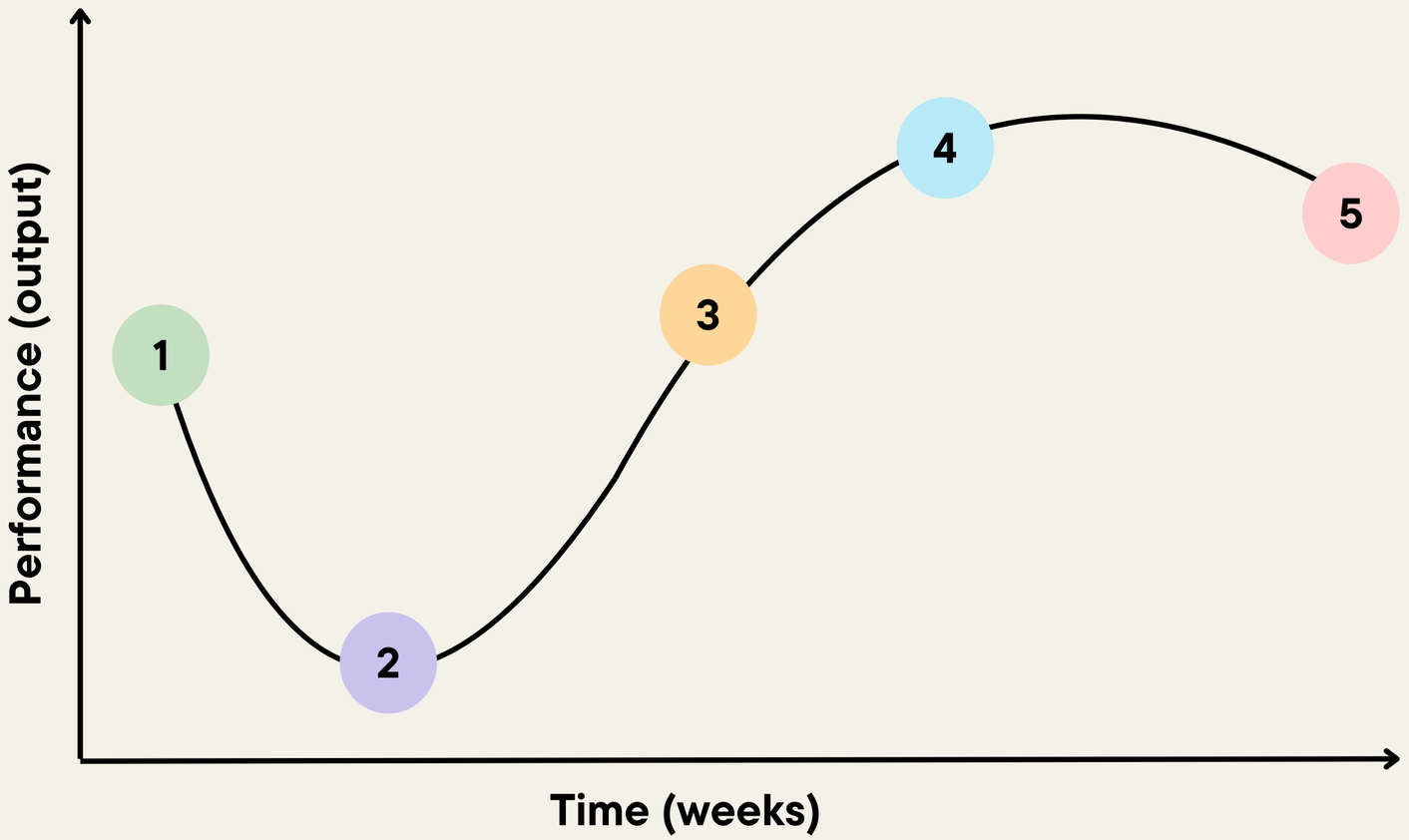
Improvements

- Our organisation of resources should have been **labelled more clearly**, to decrease the chance of confusion.
- We could have managed tasks more effectively by implementing a **RACI matrix** which was similar to our approach. By **clarifying roles** into responsible, accountable, consultation and informed, it would have **improved communication, collaboration** and **organisation** in a group project where no task is handled by a single person.

Group Dynamics

Analysing Performance Against **The Tuckman Model**

We noted on reflection that our project journey followed the curve in the model



1 - Forming (Day 1)

The initial enthusiasm for the project gave a positive team environment where we had lots of ideas for the empathise stage.

2 - Storming (Week 1 end)

We had some issues with lack of clarity on what we were actually trying to achieve, with some members feeling ‘out of the loop’ going into one of our meetings. The large number of tasks and assement criteira caused some overwhelming. We chose to have a reset and discuss where we had got to, and then made trackers and plans to ensure we were all on the same page going forward.

3 - Norming (Week 2)

At this point we had settled into a routine of meetings and deadlines, had landed on our undesirable behaviour, and had conducted research. This lead to more shared aligned vision which started to help us work on completing group and individual tasks.

4 - Performing (Week 3)

At this point we felt output of slides and design progress peaked as we were having more efficient meetings and getting clear action points. We also started giving content constructive criticism allowing us to improve and ensure we were having users and design priciples at the core of our project. We were also working together in person outside of meetings more as we sawe the benefits of bouncing ideas off each other.

5 - Adjourning (End)

As the deadline closed in our output dropped slightly due to stress of submission and deadlines in other modules.

Design Techniques Index

Emphasise	
Technique	**Slide numbers**
Brainstorming	3-5, 7-8
SWOT Analysis	6
Background Research	9-19
Early Problem Statement	20
Questionnaire	21-28
Interviews	29-36

Define	
Technique	**Slide numbers**
Personas	39-44
Scenarios	45-47
Storyboards	45-47
Refined Problem Statement	48

Ideate	
Technique	**Slide numbers**
QOC	51, 69-70
Brainstorming	52
Initial Sketches	53-58, 60
6 Hats	59
Refined Ideas	61-68
Choice Rationale	71
Feature Persona Matrix	72

Prototype	
Technique	**Slide numbers**
Lo-Fi Prototype	76-78
Wireframes	79-80
Features summary	81
Functional Details	82-89
Scam Feature Matrix	90

Test & Evaluate	
Technique	**Slide numbers**
Scenario Walkthroughs	93-95
User Feedback	96
Think Aloud Testing	97
Persona Design Evaluation	100-101
Evaluation of Methods	102-103
Evaluation of our Solution	104
Usability Heuristics Evaluation	105-107
Future Expansion	108