

USENIX

THE ADVANCED COMPUTING
SYSTEMS ASSOCIATION

# Anxious and Aware: Examining the Effects of Social Anxiety on Social Engineering Resilience and Vulnerability

Martin Dechant, Victoria Woo, Florence Tomlin, Bettina Moltrecht, and Mark Warner, *University College London*

https://www.usenix.org/conference/soups2026/presentation/dechant

This paper is included in the Proceedings of the Twenty-Second Symposium on Usable Privacy and Security.

August 24–26, 2026 • Hannover, Germany

ISBN 978-1-939133-59-5

Open access to the Proceedings of the Twenty-Second Symposium on Usable Privacy and Security is sponsored by USENIX.

# Anxious and Aware: Examining the Effects of Social Anxiety on Social Engineering Resilience and Vulnerability

Martin Dechant
*University College London*

Victoria Woo
*University College London*

Florence Tomlin
*University College London*

Bettina Moltrecht
*University College London*

Mark Warner
*University College London*

## Abstract

Social engineering attacks manipulate individuals into acting against their own interests, by exploiting users' cognitive capabilities. Less is known about how characteristics of social anxiety, a trait that exists on a spectrum ranging from mild situational anxiety to a clinical threshold affecting daily life, may influence susceptibility to social engineering attacks. The present study starts to address this gap. We conducted an online survey ($N = 42$) to identify real-world social engineering tactics, which informed role-playing interviews (N = 20) that compared responses in individuals on the higher and lower end of the SoAnx spectrum. Our findings show that characteristics of SoAnx can create vulnerabilities and benefits depending on the type of attack. Participants suggested technology solutions, ranging from educational resources and automation to social community support. We contribute new qualitative insights into how SoAnx may affect both vulnerability and resilience against social engineering attacks.

## 1 Introduction

Social engineering attacks are defined as the use of psychological manipulation to influence individuals into actions against their best interests [49]. They have been on the rise with the advancement of technology and increased online connectivity, leaving many victims distressed [1, 12, 13, 24, 29, 66, 68]. Social engineering attacks have been identified as a primary issue in cybersecurity, where human vulnerabilities as opposed to technological weaknesses have become the target of criminals [1, 4]. Criminals increasingly exploit individuals' need for social connection through schemes such as romance and pig butchering scams, where offenders pose as ideal partners to groom victims over time for financial gain [64, 70]. Ultimately, such manipulative tactics can coerce victims into surrendering personal information and financial assets, with some cases escalating to forms of sexual extortion [3, 38, 64, 67, 70, 72].

To prevent victimization, researchers have investigated whether certain individual characteristics, such as loneliness or low self-esteem, affect susceptibility [59, 61, 62, 64, 72]. This research has shown that certain psychological factors increase vulnerability. However, specific work on the impact of certain mental health conditions, like social anxiety (SoAnx), where many of these identified risk factors are present at once, is lacking [30, 59]. SoAnx is characterized as a continuum, ranging from experiencing anxiety in certain social situations, all the way to a clinical threshold where anxiety symptoms affect the whole life of the individual [15, 59]. Socially anxious individuals are more likely to engage with others online than in-person, putting them at increased risk of social engineering attacks. Additionally, they may be more susceptible to manipulation, as their tendency to avoid conflict or social pressure can make them more likely to comply with requests [16], a dynamic often exploited in social engineering attacks [49]. On the other hand, socially anxious individuals' heightened vigilance—stemming from their tendency to closely monitor social cues online [15, 59] and offline [45] —may help them detect deceptive or suspicious cues. Considering that SoAnx is a highly prevalent condition we argue that there is an urgent need for researchers to better understand the association of SoAnx and social engineering attack risks.

To address this gap and inform the design of more effective support strategies, we conducted a two-part study consisting of a scenario development survey (study 1) with people who recently experienced a social engineering attack, followed by role-playing interviews (study 2) to better understand possible risk and protective factors relating to the scenarios and varying levels of SoAnx. In study 1, we identified three types

Copyright is held by the author/owner. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee.
*USENIX Symposium on Usable Privacy and Security (SOUPS)* 2026.
August 23–26, 2026, Hannover, Germany.

of social engineering attacks experienced by participants: (a) exploiting trust and desires, (b) extorting with compromising material, and (c) exploiting empathy and urgency. The identified attacks informed the development of three scenarios used in study 2, in which 20 participants (10 with high and 10 with low trait SoAnx) were guided through the scenarios as part of semi-structured interviews. Our interview findings suggested shared vulnerabilities were present across both groups, such as barriers to reporting scams. Additionally, we found SoAnx was able to foster resilience against some social engineering attacks while amplifying susceptibility to others. In other words, our work demonstrates that SoAnx creates context-dependent vulnerability: rather than uniformly increasing risk. In terms of designing security systems, participants proposed automation, risk indicators, and community resources, which we position as preliminary directions requiring further research rather than prescriptive solutions.

## 2 Related Works

We situate our research within three overlapping areas of work which we explore in this section. Firstly, we present prior literature on SoAnx to understand the characteristics of this trait. Secondly, we explore mechanisms used within social engineering attacks and related vulnerabilities to these attacks. Thirdly, we explore previously developed prevention methods, designed to help protect people from these forms of attack. Finally, we conclude with our research goals and motivations.

### 2.1 Characteristics of Social Anxiety

Social anxiety (SoAnx) has been characterised as *"a marked, or intense, fear or anxiety of social situations in which the individual may be scrutinised by others"* [5]. SoAnx involves a combination of trait and state anxiety, where state SoAnx refers to temporary experiences of anxiety in response to specific social situations (e.g., job interviews, public speaking) [15, 33, 39, 59]. Trait SoAnx refers to more consistent personality tendencies in a person that come with persistent cognitive and behavioural biases (e.g. increased focus on social threat stimuli) and often present as an increased fear of negative evaluation by others as well as certain safety and avoidance behaviours (e.g. avoiding public events) to reduce these negative experiences [15, 59]. These characteristics for SoAnx are important to consider in order to better understand possible vulnerabilities to social engineering attacks. While there has been prior work focusing on psychological factors that increase vulnerability to social engineering attacks [17, 19, 40, 46, 47, 55, 64, 71, 72, 74], there is still little research in the intersection of SoAnx and social engineering attacks.

### 2.2 Social Engineering Mechanisms and Vulnerabilities

The landscape of social engineering attack vulnerability is complex given how many different factors they aim to exploit including psychological and situational vulnerabilities: Bullée et al. found 88% of attack patterns employ persuasion principles targeting users directly [12, 63], while perpetrators strategically target individuals experiencing economic distress and desperation [21, 36, 42]. The success of social engineering hinges on the exploitation of fundamental human needs [61, 68] such as distraction, social compliance, the herd principle, dishonesty, kindness, need and greed, and time pressure [61]. Vardhan et al. found that 84% of participants succumbed to at least one social media scammer-like profile, with "trust" and the "fulfilment of need" being cited as primary drivers [66]. While these principles apply broadly to the general population, Whitty and others show that susceptibility is not uniform and can be influenced by demographic, personality, and psychological factors [71].

Prior work has explored the role of *user demographics* in the vulnerability to social engineering attacks. Findings suggest that females [28, 58], and especially older women are often more susceptible to phishing emails [40]. Age has also been explored as a general vulnerability factor, with mixed findings indicating that younger age groups (age 18-25 years) [58] as well as older age groups were more vulnerable to phishing attacks than other age groups [40]. Older adults, face heightened online security vulnerabilities, particularly those with mild cognitive impairment [46]. Furthermore, Zhai et al. found that seniors were more at risk in DeepFake scams because they may lack knowledge about scam activities and the skills needed to navigate evolving technology spaces [74], thus supporting findings that show vulnerability is impacted by user confidence and risk perception [51].

Research has also explored the role of different *personality characteristics* with personality traits such as agreeableness, neuroticism and openness [25, 28], being associated with increased vulnerability to certain type of attacks. Other personality traits such as curiosity, can override a user's recognition of danger and knowledge about cyberattacks [6]. As, Stajano and Wilson pointed out, people do not fall for scams because they are *"gullible, but because of human nature"* [61]. In contrast, those with high levels of conscientiousness [20] and extroversion [20] have found to be less susceptible to these forms of attack.

*Psychological factors* that have been shown to increase vulnerability include mild cognitive impairment in older adults [46], and dysfunctional impulsivity [51]. Scammers have also been shown to systematically exploit low self-esteem and security concerns to establish trust before inflicting financial and emotional harm [72]. The exploitation of vulnerabilities can also extend to family dynamics. Deng et al. observed that filial piety, normally a source of support, could become a

"double-edged sword" in online social engineering attacks, as seniors sometimes pressure family members into complying with scammers' requests [17].

Those affected by social engineering attacks experience significant emotional distress, including feelings of self-blame, and shame [3, 13, 17, 55, 65, 72] with consequences extending to families and close communities [13, 17]. These emotional and social consequences can contribute to underreporting of scams [3, 17, 67]. Budimir et al. suggests that individual who experience intense emotional negativity are less likely to engage in adaptive help-seeking behaviours, while those with minimal emotional response may lack the motivation to seek support [11].

Taken together, prior work shows that universal psychological principles create baseline susceptibility across populations, while demographic factors (age, gender) and psychological states (emotional distress, cognitive impairment, low self-esteem) create differential vulnerability.

## 2.3 Social Engineering Attack Prevention

Proposed prevention strategies include policy changes, design interventions such as nudges, and social support systems [2, 17, 34, 44, 51, 52, 74, 75]. Education about social engineering has proven effective in helping individuals to protect themselves [2, 44], while design strategies can support vulnerable users without limiting autonomy [51, 74, 75]. Social support from family and friends can provide additional external perspectives to protect against manipulative tactics used in social engineering [17, 52]. In summary, social engineering attacks are a multi-layered threat characterised by three key aspects: an expanding digital attack surface that reduces barriers to deception, systematic exploitation of psychological principles that transcend technological defences, and profound consequences that extend beyond individual victims to create systemic barriers to reporting and recovery. Given these systematic patterns of exploitation, understanding how specific psychological vulnerabilities might create differential susceptibility becomes critical.

## 2.4 The Scope of the Study

Given the absence of prior research in this area, we conducted an exploratory investigation to help us understand how SoAnx characteristics might influence responses to social engineering attacks. Rather than assuming uniform vulnerability, we sought to identify whether the distinctive cognitive and behavioural features of SoAnx create context-dependent patterns across different attack types—as suggested by the literature but not yet empirically examined. This preliminary study employed a participatory approach to ensure ecological validity while maintaining ethical boundaries around exposing participants to actual deceptive content. We formulated the following three exploratory research questions to identify patterns and generate initial insights:

**RQ1:** What are common social engineering attacks experienced by participants with varying levels of trait SoAnx?

**RQ2:** How do these social engineered attacks affect individuals with trait SoAnx?

**RQ3:** What technological designs do participants suggest that could support them during socially engineered attacks?

## 2.5 Ethics Statement

Both studies were approved by the ethics board of the University College London (Ethics ID: LMS REC - 2025-0470-961). All participants provided written informed consent after reviewing study information, could withdraw anytime, and received mental health support resources. Data were de-identified, securely stored on encrypted systems, and reported using pseudonyms.

# 3 Study 1: Pre-screening and Scenario Development Survey

## 3.1 Methods

In study 1, we aimed to understand participants' real-world experiences of social engineering attacks **(RQ1)**, pre-screen participants for the second phase, and use the descriptions of their experiences to inform the development of a set of role-playing scenarios. Further we aimed to identify specific contextual details, emotional aspects, and progression patterns to shape our scenario design. We deployed a survey to collect information on the types of social engineering attacks participants with both high and low SoAnx, submitted in response to an open-ended question. While existing social engineering taxonomies and frameworks exist that could form the basis of the scenarios [50, 73], we wanted to develop scenarios (to be used in study 2) grounded in the lived experiences of those with varying levels of trait SoAnx, to help surface attacks that may be more prevalent amongst this population. Finally, our scenarios were refined based on prior work to strengthen their credibility [3, 14, 19, 37, 50, 64, 65, 67, 69].

### 3.1.1 Participants and Procedure

A total of 67 responses were collected through the participant recruitment platform of the University College London in the United Kingdom and advertised on Instagram and LinkedIn of the profiles of the researchers. Study adverts described the aim of the study, information about ethics and data management, and a link to participate in the screening task. Inclusion

criteria required participants to be: (1) at least 18 years of age, (2) capable of providing informed written consent, and (3) have experienced at least one social engineering attack (self-reported). The study was conducted over a two-week period from the 3rd of June till 17th of June 2025.

We compensated participants' time with a 5 GBP shopping voucher. The survey began with the study information, consent form, and mental health resources, followed by questions about whether participants had experienced a social engineering attack. If they indicated that they had experienced a social engineering attack, they were invited to provide details of their experience in an open-ended text box. Participants then completed the Liebowitz Social Anxiety Scale (LSAS) [39, 57], the Patient Health Questionnaire (PHQ-9) [35], and the Generalised Anxiety Disorder Scale (GAD-7) [60]. Using LSAS, PHQ-9, and GAD-7 together allowed us to better understand the current mental well-being, especially their levels of SoAnx, and to exclude participants who showed extreme levels of anxiety or depression using GAD-7 and PHQ-9 [56]. Participants who were excluded due to severe mental health risks were referred to mental health experts and offered mental health resources.

#### 3.1.2 Measurements

**Liebowitz Social Anxiety Scale (LSAS).** The LSAS [26, 57] was used to assess participants' levels of trait SoAnx. The scale consists of 24 items that evaluate fear and avoidance in a range of social and performance situations. Each situation is rated separately for fear (0 = none to 3 = severe) and avoidance (0 = never to 3 = usually), producing a total score ranging from 0 to 144. Higher scores indicate greater SoAnx.

**Patient Health Questionnaire (PHQ-9).** Depressive symptoms were measured using the PHQ-9, a 9-item self-report measure [35]. Each item is rated on a 4-point Likert scale (0 = not at all to 3 = nearly every day), resulting in a total score from 0 to 27. Higher scores reflect greater severity of depressive symptoms.

**Generalised Anxiety Disorder Scale (GAD-7).** General anxiety symptoms were assessed using the GAD-7, a 7-item measure [60]. Items are rated on a 4-point scale (0 = not at all to 3 = nearly every day), yielding a total score between 0 and 21. Higher scores indicate more severe general anxiety symptoms.

#### 3.1.3 Analysis

Of the 67 responses, we removed 18 participants, of which seven had not experienced social engineering attacks and 11 entries were considered fraudulent data [54]. Suspected fraudulent responses were reviewed by two researchers independently. Metadata was reviewed, together with the submitted responses. We identified fraudulent responses where IP addresses were duplicated across multiple submissions, suggesting the same user was submitting multiple entries. We also identified VPN usage from the same set of providers being used in submissions that were submitted close in time, with written responses that were similar. We excluded six additional responses lacking sufficient detail and combined one duplicate, resulting in 42 participants. All 42 participants' attack descriptions were analysed together, regardless of LSAS scores, to identify common social engineering tactics experienced across the SoAnx spectrum.

We analysed all 42 participants' free-text responses describing their scam experiences. Two researchers independently coded responses using an inductive reflexive coding approach at the semantic level [7, 8]. Both coders met frequently to discuss each others developing codes, to support shared understanding and the collaborative development of themes. Codes were grouped into broader themes representing distinct social engineering attack types. During initial open coding for study 1, approximately four codes were generated across participants' responses, which were then refined into the three final themes based on prevalence and distinctiveness of manipulation mechanisms (i.e., the initial hook used to engage victims). This approach allowed themes to be developed from participants' lived experiences, ensuring analysis remained grounded in real-world social engineering encounters to inform scenario development.

#### 3.1.4 Sampling Rationale and Characteristics

Study 1 identified common social engineering patterns from participants' lived experiences rather than imposing predefined taxonomies, ensuring study 2 scenarios reflected real-world attacks. We refined scenarios using established frameworks [50, 73] to enhance ecological validity.

Participants' LSAS scores ranged from 0 to 133 (M = 69.86, SD = 29.70, Mdn = 74), showing substantial variability. The distribution exhibited moderate negative skew (skewness = -0.57, kurtosis = 0.10), with the sample mean substantially exceeding typical clinical cut-offs for SoAnx disorder (above 30), indicating strong representation of clinically significant SoAnx (see Figure 1).

Participants reported mild-to-moderate levels of depression (PHQ-9: M = 9.58, SD = 6.15) and anxiety (GAD-7: M = 8.81, SD = 6.21), with both distributions approximately symmetric and near-normal (skewness = 0.35 and 0.13, respectively).

### 3.2 Findings Study 1: Commonly experienced social engineering attacks (RQ1)

In our survey, we aimed to understand the basic elements and tactics within social engineer attacks experienced by our participants with varying levels of SoAnx. Although many of the descriptions could have been placed in different themes, we limited each social engineering attack to a single theme.

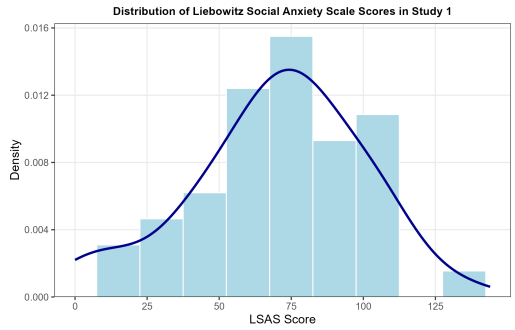



Figure 1: Distribution of LSAS scores (N = 42). Blue line = observed distribution. The distribution exhibits moderate negative skew (skewness = -0.45), with strong representation of higher SoAnx (M = 69.86, SD = 29.70).

This allowed us to better understand the main lure of the scam that gained participants' engagement. Some participants shared more than one experience during the screening and Table 1 counts the inputs per story. From this we developed three themes focusing on common experiences of our participants: (1) Exploiting Trust and/or Desires, (2) Extorting with Compromising Materials, and (3) Exploiting Empathy and Urgency.

#### 3.2.1 Exploiting Trust and/or Desires

This theme describes scams where perpetrators leverage a victim's trust or personal desires to manipulate them ($N = 27$). The typical mechanism involves an attacker making initial friendly contact to build a false sense of rapport or using trust symbols of known entities which may bypass the victim's defence. Once trust is established, they present a deceptive opportunity that appeals to the victim's desires, such as a promising job offer, a business investment, or obtaining their goods. This approach is effective because it subverts a person's natural wariness by first appealing to them either through trust and/or desires [14, 65]. Afterwards, they may try to gain access to personal information or financial benefits from their attack targets. These scams are becoming more common with access to technology as they are lower risk due to their online nature [14].

#### 3.2.2 Extorting with Compromising Materials

This theme covers a form of exploitation where perpetrators use coercion and threats to compel the victim to follow the perpetrator's demands ($N = 5$). The offender manipulates or threatens the victim into sharing information or money, often weaponising any shared material with tactics of pressure, extortion, and blackmail to coerce further payments. This tactic is effective because it preys on powerful negative emotions like fear and shame, and on the victim's strong desire to protect their personal details [3].

#### 3.2.3 Exploiting Empathy and Urgency

This theme outlines scams that work by manipulating a victim's empathy or possible relation through a fabricated crisis ($N = 11$). Perpetrators typically begin by impersonating a loved one or creating a sympathetic persona to gain the victim's trust. Once a connection is made, they invent a sudden, dire situation, such as a medical emergency or an accident, creating a strong sense of urgency that pressures the victim to act immediately. The combination of an appeal to empathy and intense time pressure is designed to trigger an immediate emotional response, compelling the victim to send money before they have a chance to think critically [53, 61].

## 4 Study 2: Role-play and Interview

The three themes of Study 1 represent distinct manipulation mechanisms encountered by participants: exploiting trust and desires, fear of compromising material, and emotional pressure. While some attacks showed overlapping characteristics (e.g., romance scams involving both trust exploitation and fabricated urgency), we categorized based on the dominant manipulation mechanism driving victim engagement. We defined the dominant mechanism as the initial hook that engaged participants' attention and compliance, prior to secondary manipulation tactics. Building on these findings from study 1, study 2 used realistic role-playing scenarios and semi-structured interviews to investigate RQ2 and RQ3.

### 4.1 Methods

To understand how individuals with trait SoAnx respond to different types of social engineering attacks, participants engaged in realistic role-playing scenarios and were interviewed remotely via Microsoft Teams. These scenarios included three types of scams: job scam, romance scam, and extortion. Each selected scam matched a theme identified by participants in study 1. We designed every scenario to reflect these specific elements mapping the contextual nuances of the lived experience into the scenarios: For example, the job scam aligns with the theme *"Exploiting Trust and/or Desires"* by placing participants in a perceived professional setting with promises of financial gain. The extortion scenario reflects *"Extorting with Compromising Material"* by blackmailing participants with personal information. Finally, the romance scam represents *"Exploiting Empathy and Urgency"* by establishing trust or romantic interest before requesting financial assistance. Participants were compensated with a 10 GBP shopping voucher per hour of their time. We utilized a guided role-play methodology to examine participant reactions to realistic scam scenarios, ranging from low to high emotional intensity (job interview to romance scam). Participants were prompted to narrate their

| Common Scam Tactics from the Screening Survey | | | |
|---|---|---|---|
| **Theme** | **Description** | **N count** | **Examples** |
| **Exploiting Trust and/or Desires** | Scammers exploit a victim's trust and desire by creating a sense of deceptive familiarity, then manipulating them with promises of materialistic or financial gains to gain access to personal information or money. | 27 | *"A person had messaged me directly on a social media platform and acted friendly at the start, asking me about myself... and then presented me with a supposed business opportunity which, in reality was a scam." - P15* |
| **Extorting with Compromising Material** | A form of exploitation where perpetrators may use impersonation to build trust, intimacy, or fear. They then manipulate or threaten the victim into sharing information or money, often weaponising any shared material with tactics of pressure, extortion, and blackmail to coerce further payments. | 5 | *"Pretended to be interested in making friends with me and sent sexual images to me and had I do so in exchange, then extorted me for money." - P43* |
| **Exploiting Empathy and Urgency** | A process of emotional exploitation where perpetrators use planned deception and impersonation to gain trust and/or love from victims, only to fabricate a crisis or urgent need. By creating a scenario that elicits empathy, they bypass a victim's rational judgment to gain access to their information and financials. | 11 | *"A person contacted me over WhatsApp from a number which had a North American pre-fix code. He claimed to be my best friend... and afterwards requested me that his father was getting operated upon... and he would like me to send funds for that." - P32* |

Table 1: Three themes from study 1 developed from screening survey showing number of mentions of each theme.

thoughts and anticipated responses through semi-structured interviews. This approach provides established internal and external validity in phishing research [18, 58, 66] and allowed for the safe, ethical exploration of participant reactions without exposure to actual attack stress.

#### 4.1.1 Participants and Demographics

A total of 20 participants (nine male, 11 female; age range: 19–51) were recruited for the interviews (see Table A2 for detailed description) where 14 had experienced a social engineering attack. Study 2 recruited participants from Study 1 follow-ups (n=6) and researcher's personal network (n=14), across the US and UK sampled from the highest and lowest LSAS score. Consistent to study 1, participants were briefed with the same study information and goal of understanding the link between SoAnx and social engineering vulnerability to inform inclusive design strategies. To be eligible for inclusion, participants were required to be at least 18 years of age and capable of providing informed written consent. We included a diverse sample regarding SoAnx, ranging from those with no reported anxiety to those with mild-to-moderate levels. Participants were split into two groups based on their LSAS score: highly socially anxious (LSAS $\geq$ 55) and low socially anxious (LSAS $\leq$ 54) [57]. Low SoAnx group were participants with a LSAS between 0-54 and the high SoAnx group consisted of participants with a LSAS score above 54. This method was chosen to provide a clear cut-off for identifying differences in participant reactions, actions, and thought processes during qualitative coding. Table A2 summarises the key demographic and psychological scale information. We included participants without social engineering experience to examine potential reaction differences. Participants were recruited through study 1 follow-up (n = 6) or convenience sampling via researchers' networks (n = 14). Data collection was conducted remotely over a four-week period between July 8, 2025, and August 6, 2025. The participant pool was geographically distributed across the United States and the United Kingdom. All participants completed the LSAS, PHQ-9, and GAD-7 assessments for SoAnx and mental well-being evaluation. LSAS was used to assess participants' levels of trait SoAnx (see Table A1) [57].

#### 4.1.2 Analysis

All interviews were audio-recorded and transcribed with participant consent, with interviews ranging from 60-120 minutes. The de-identified interview transcripts were analysed using Reflective Thematic Analysis to allow themes to be informed directly from the participants' own descriptions [7, 8]. We adopting an inductive coding approach, coding at the semantic level in NVivo 14 [41], where two researcher coded the responses with multiple themes assigned per participant when appropriate. This approach enabled comparison between social anxiety and non-social anxiety groups to assess theme salience.

Aligned with reflective thematic analysis approach [9], as opposed to statistical inter-rater reliability checks, we performed conceptual checking of developed codes and themes between the two researchers involved directly in coding and in collaboration with the wider research team during regularly research meetings. We do not quantify qualitative findings, but do report trending similarities and differences between the low and high SoAnx groups to help us understand the role of SoAnx in social engineering attack scenarios; while being mindful of the limitations of this comparison, due to our small

sample size. Initial open coding generated approximately 53 codes across participants' responses for study 2, which were then grouped into 15 intermediate categories before being refined into the five final themes based on SoAnx characteristics and behaviours.

## 4.2 Study Design

Study 2 scenarios were developed through a process combining study 1 findings, prior literature, and iterative refinement to create realistic scenarios examining how SoAnx characteristics interact with different manipulation tactics. Study 1 identified three distinct manipulation mechanisms: exploiting trust/desires, extorting with compromising material, and exploiting empathy and urgency. These align with prior work suggesting differential engagement of SoAnx characteristics such as applying safety behaviours and avoidance behaviour [15, 59]. We selected one prototypical scenario per theme: job scams, extortion, and romance scams. Initial drafts extracted common elements from study 1 participant descriptions, then were refined using established social engineering frameworks [50, 73] and empirical studies of specific attack types [3, 19, 64, 65]. Fictitious persona images and initial messages were generated using Gemini 2.5 Flash [22] to ensure visual and linguistic realism. Each scenario included two parts mirroring typical attack progression: initial contact/trust-building, followed by manipulation [68]. The research team—two cybersecurity experts, one clinical psychologist, and two digital health experts—reviewed scenarios to ensure they reflected study 1 themes, contained realistic manipulation tactics, and were ethically appropriate. Participants were guided through each scenario and asked to describe how they would interact with and react to the situation. Throughout role-playing, semi-structured interview questions were used to gain insights into the participants' thoughts, reactions, emotions, and the actions they would take in each scenario. Furthermore, we asked questions on how technology could possibly support the users within each scenario.

## 4.3 Materials and Scenarios

Based on the findings from study 1 (see Table 1), we used the following scenarios that were developed around the three major exploitation themes. All three scenarios were build around a dominant exploitation strategy of study 1. None of the scenarios indicated whether the target was successfully defrauded, to allow participants to consider how they would likely respond to each scenario: **Job Scam.** Participants were presented with a two-part recruitment scam involving a job offer, interviews, submission of employment documents, and a request to transfer funds before a cheque had cleared.

**Extortion Scam.** Participants were presented with a two-part extortion scenario involving threats related to alleged possession of compromising information and escalating demands for compliance.

**Romance Scam.** Participants were presented with a two-part romance scam involving the gradual development of an online relationship, the exchange of personal information, and circumstances designed to exploit trust and emotional investment.

## 4.4 Findings Study 2: Social Anxiety as a Risk and Resilience Factor (RQ2)

We focus on patterns relevant to understanding how SoAnx characteristics may influence responses to social engineering. Some patterns appeared across both groups, suggesting universal experiences, while others were more prominent within one group, suggesting SoAnx-specific dynamics. Our analysis revealed a complex pattern between SoAnx and social engineering attack strategies. Rather than uniform effects, our exploration shows key characteristics of SoAnx can help users become more resilient while other aspects may increase their susceptibility to social engineering attacks. Following reflexive thematic analysis principles [8], we emphasise that the importance of themes is determined by their theoretical relevance to our research questions and alignment with SoAnx literature [15, 59], rather than by frequency of occurrence.

The high and low SoAnx groups were clearly separated on LSAS scores with no overlap (Low SoAnx: n = 10, M=35.6,SD = 10.9, range=17-46; High SoAnx: n = 10, M=68.3, SD = 11.5,range=55-88), supporting our comparative exploration. First we examine baseline vulnerabilities affecting all participants regardless of their levels of SoAnx. We then present SoAnx specific vulnerabilities for social engineering attacks and resilience factors against them.

### 4.4.1 Avoiding support seeking as baseline vulnerability of all participants

Barriers to support seeking manifested across all three scenarios, affecting participants regardless of SoAnx levels.

**Reporting Barriers:** All participants stated that they would avoid reporting incidents to authorities, describing multiple barriers to reporting. For some, there was a lack of knowledge and understanding around how to report such incidents, with P07 (low) stating: *"It's such a common occurrence and I wouldn't know where to start."* Beyond lack of knowledge, perceived effort was also a barrier for some, with P10 (high) stating: *"I just wouldn't, it will take more effort."* While some participants were aware of reporting processes, they avoided reporting due to scepticism over the effect it would have, as P09 (high) explained: *"I think it's not in the sense of if I'm comfortable or not, but more in the sense of will this do anything or not?"*

**Avoiding Sharing The Problem:** While both groups throughout all scenarios indicated they would avoid talking

to others (e.g., friends) about an experienced social engineering attack, the reasons for avoiding sharing experiences with others differed across the two groups. Responses from high SoAnx participants suggest that aspects of their SoAnx impacts on likelihood to share. These participants showed greater avoidance behaviours due to fear of negative evaluation by others (Theme: *Judgment and Reprimand from Social Circle*). This fear of judgment is a characteristic feature of SoAnx [5], suggesting these sharing barriers may be particularly pronounced for individuals with elevated SoAnx rather than reflecting general privacy preferences. P10 (high) explains: *"I just don't want to hear from my friends. They'll laugh at me... Because I'm already judging myself, and I will assume that others will judge me as well."* Beyond fear of judgement, high SoAnx participants within the romance scam scenario focused on the social relevance of their sharing (Theme: *Lack of Sharing Due to Social Irrelevance*), with P11 (high) saying: *"You know, it's a drama or something cool... like people will be interested in, like a story."* In contrast, low SoAnx participants' avoidance was primarily motivated by privacy-concerns (Theme: *Personal Privacy for Participant*). P12 (low) stated: *"I don't share or discuss this."* While these support-seeking barriers affected *all* participants, our analysis suggests that SoAnx characteristics such as worry about everyday activities and fear of social judgement, may create additional and distinct patterns of vulnerability in specific attack contexts.

### 4.4.2 Social Anxiety as a Potential Vulnerability Factor

Next we present themes that reflect characteristics of SoAnx like reassurance-seeking [23], attention bias to threat [15], and safety behaviours [27]. These characteristics appeared to increase susceptibility to the three social engineering attacks which we role-played in our study.

**Emotional and Social Vulnerabilities:** High SoAnx participants showed greater sensitivity to social norm violations than low SoAnx participants, particularly in the romance scam scenario. In the job scam scenario, high SoAnx participants appeared to positively respond to affirmative language (Theme: *Positive Perception of "Friendly" Language*, see Appendix Table A3). P03 (high) found phrases like *"great candidate"* provided *"a bit more comfort than a basic request for a job,"* while P04 (high) felt validated: *"I am capable, I am able to get the salary and I am good for this position."* In comparison, low SoAnx participants questioned claims using affirmative language, with P06 (low) saying: *"I'm a great candidate, but like based on what, right?"*

Affirmative language within these types of scams may increase scam vulnerability for high SoAnx participants due to their responses. This vulnerability may stem from a heightened need for social validation and the avoidance of authority figures [39], making excessive positive affirmation appealing rather than suspicious (identified by low SoAnx participants). This pattern aligns with established characteristics of SoAnx, particularly reassurance-seeking behaviours that aim to reduce anxiety through external validation [23]. Furthermore, in the extortion scenario, emotional curiosity of high SoAnx participants appeared to override safety knowledge and concerns (Theme: *Knowledge of Danger Overwritten by Emotions*, see Appendix Table A3). P09 (high) expressed this internal conflict explicitly, saying: *"I think I would be really curious to see what this supposed blackmail is... I know it's definitely a trap, but..."* This suggests that awareness of danger may be not be enough to avoid engagement when emotional curiosity is aroused, with high SoAnx individuals being more at risk due to a biased perception fuelled by trait SoAnx. For one high SoAnx participant, they recognised and managed the risk through technical means (virtual machine sandbox) while their curiosity remained aroused, P11 (high) stat: *"OK cause I have like a virtual machine on one of my old laptops. I probably click the link there and see what it is."*

These emotional vulnerabilities extend to high SoAnx participants' acute awareness of their own emotional states and potential decision-making compromises (Theme: *Desperation From Pressures of Unemployment*, see Appendix Table A3). This meta-cognitive awareness had complex effects on their vulnerability to social engineering attacks. P09 (high) recognised their potential for impulsive action: *"I feel like knowing myself, I might if I was desperate enough, I might impulsively respond first..."* P02 (high) explicitly connected emotional state to decision-making: *"It gives me such a good opportunity and if I was desperate for a job, yeah, I'd just be shocked that someone would reach out to me... I'd be tunnel focused almost."* P03 (high) similarly acknowledged vulnerability: *"I maybe would have reacted to it if I was quite desperate for a job. I would understand that in that situation."* This self-awareness represented both a protective factor and a vulnerability. As protection, it allowed participants to recognise when their judgment might be compromised and potentially adjust their behaviour accordingly. This intense focus on internal emotional states also created vulnerability by diverting attention from external threat cues towards internal emotional states.

Beyond these potential emotional vulnerabilities (see Appendix Table A3), we also observed patterns suggesting that SoAnx characteristics, especially the urge to withdraw from social situations, might influence access to external support.

**Conditions and Social Withdrawal:** While all participants indicated baseline levels of sharing avoidance, high SoAnx participants consistently expressed more barriers to sharing. They appeared to require specific conditions to be met before sharing and demonstrated greater withdrawal from social circles than low SoAnx participants, potentially limiting access to external perspectives often critical for support.

We found high SoAnx participants using humour as a protection method (Theme: *Humour as a Sharing Method*, see Appendix Table A3), providing them with a means of shar-

ing, whilst hiding their perceived failure using humour. As an example of this, in response to being asked whether they would share their experience with others P11 (high) said: *"Yeah. I might. Yeah, I guess I might if I can spin it to be less pathetic and more like funny."* Additionally, this humour condition extended to the recipient of the experience, placing more barriers to receiving the help needed to cope with social engineering attacks. P03 (high) expressed these conditions: *"Probably only my best friend, but she would need to swear to me to not use it as often as a joke."* The use of humour can be seen as a protective disclosure strategy similar to other safety behaviours of social anxious individuals. To prevent negative social outcomes, such as negative judgement, individuals apply humour as a way out [27].

These conditions that led to possible social withdrawal are not limited to sharing with others, but can extend to withdrawal from larger engagement (Theme: *General Increased Vigilance Differing Behaviour*, see Appendix Table A3). High SoAnx participants appeared more likely to completely withdraw from platforms after experiencing a social engineering attack. P02 (high) explains: *"I just wouldn't, I wouldn't take the risk on any platform, you know, similar like if it happened to me on any platform, even a platform which I have a good, like I have a good impression of"*

Further, our thematic analysis also suggested that certain SoAnx characteristics might offer potential protections, indicating a more complex relationship than uniform vulnerability.

#### 4.4.3 Social Anxiety as Potential Resilience Factor

Across our scenarios we identified several SoAnx characteristics of which support individuals' resilience against social engineering attacks. In particular we identified particularly heightened attention to social cues and norm violations [59], and social avoidance tendencies [27].

**Detection of Social Norm Violation:** Compared to low SOAnx, high SoAnx participants appeared to show heightened sensitivity to inappropriate social behaviours (Theme: *Unsolicited Information Sharing and Pet Names*, see Appendix Table A4). We found many high SoAnx participants immediately flagged premature intimacy within the romance scams scenario, while those with low SoAnx responded with less immediacy to premature intimacy. This finding highlights the relationship between characteristics of the attack context and SoAnx vulnerability, with the same emotional intensity that created vulnerability in extortion contexts appearing to act as a protective trait when perpetrators violated social boundaries. P03 (high) questioned bank detail sharing: *"I mean, why do I need to have that anyway?"* and expressed discomfort with pet names: *"I don't like being called baby."* P18 (high) noted: *"I think calling someone baby if you haven't met them in person, you know that's not a very sound judgement."* High SoAnx participants' discomfort with inappropriate intimacy and premature personal information sharing created immediate alarm responses. Furthermore, social pressure placed high SoAnx participants in a heightened state of alert (Theme: *Vigilance from External Pressure and Urgency*, see Appendix Table A4). P10 (high) elaborated: *"Yep, same day. That's a red flag. They need me to get it done as soon as possible."* High SoAnx participants felt the pressure even across written messages, placing them in a higher state of alert against the possible scam. Mirroring this similar sentiment, P09 (high): *"In my impressions of bad situations, I feel they will rush you, especially when you're in a desperate position..."* Heightened awareness of social norm violations amongst high SoAnx participants may serve as protection against social engineering attacks, as awareness of these norm violations may increase alertness to potential risks.

The apparent sensitivity to social boundary violations, potentially induced by the characteristic hyperfocus on social cues, seemed to extend to participants' potential ability to detect other forms of communication inconsistency.

**Linguistic, Tonal, and Information Inconsistency Detection:**

We found participants with high SoAnx were better able to detect social engineering attack cues within the job scam scenario (Theme: *Detection of Discrepancies and Tonal Anomalies*, see Appendix Table A4). P11 (high) detected "unprofessional" communication patterns and P02 (high) recognised the use of pressure tactics: *"It almost seems like I'm being pressured."* Furthermore, similar detection appeared in the extortion scenario related to the absence of information (Theme: *Vigilance on Missing Information*, see Appendix Table A4). P13 (high) noticed: *"They haven't even given like bank details, so you'd have to interact with them."* The potential heighten sensitivity of high SoAnx participants for linguistic, tonal, and information inconsistency may offer protection against social engineering attacks as they become more aware of the situation and question the authenticity of the communication. This enhanced detection capability is consistent with the characteristic attention bias of SoAnx, where individuals habitually scan for signs of potential social risks and signs of evaluation [15, 59].

**Avoidance of Social Interactions:** High and low SoAnx participants differed in how they sought information, with high SoAnx participants preferring indirect research over direct engagement. We identified protective restraint, with high SoAnx participants being more cautious, seeking additional information and clarification than low SoAnx participants (Theme: *Probing for Information and Clarification*, see Appendix Table A4). While low SoAnx participants actively engaged with the attacker, P06 explained: *"I would probably just ask that question, like what is the position?"*, high SoAnx participants preferred to withdraw from the interaction and use other ways to gather information . P09 (high) explains: *"I think I would probably research more about the company and the person rather than directly asking them."* Furthermore, in

the romance scam scenario, while low SoAnx participants reported feeling a sense of obligation to respond on dating apps, high SoAnx participants responded differently, reporting less obligation to respond (Theme: *Social Obligation to Respond*, see Appendix Table A4). P12 (low) noted that being on a dating app carries an expectation of interaction, explaining that if he did not reply: *"It's pretty rude."* Through these social avoidance approaches may limit their exposure to the social engineering attack due to their withdrawal tendency.

Given these patterns in our themes suggesting both potential vulnerabilities and protective factors, we explored together with our participants what technological support might be helpful during social engineering attacks.

## 4.5 Findings Study 2: Technology Design Recommendations (RQ3)

As part of the semi-structured interview we asked participants to explore how technological interventions and support systems may have been useful to them during the presented three scenarios. During initial coding, we had six codes for participant ideas; we then categorised these into the following three main themes:

### 4.5.1 Proactive and Automated Security Measures

As developed through the theme *Proactive and Automated Security Measures* (Appendix Table A5), participants with high SoAnx suggested systems that could enhance their natural detection abilities [59] while reducing cognitive load that can result in reduced attentional control (concentration). These measures included automated filtering (P04 (high): *"I think apps should be able to like maybe regularly be scanning their profiles for like scammers"*), content pre-screening to address curiosity-driven vulnerabilities (P11 (high): *"If something can click into that link for me and then relay the content to me, it would be pretty nice"*), real-time warnings (P10 (high): *"I think that there should be messages, warning messages for people."*), and active system intervention (P02 (high): *"I just wish that something intervened, I just wish that something stopped me before, I committed this action"*).

### 4.5.2 Community-Based Support Systems:

As shown through the theme *Collective Knowledge and Sharing for Support* (Appendix Table A5), participants suggested social proof which may offer support and overcome previously identified barriers. Participants suggested community feedback systems (P05 (low): *"I want to see the comments of others to this person on a dating app..."*), anonymous support networks to address isolation (P02 (high): *"I wish I would have support. But I don't know, I don't know where I would get the support from"*), and shared experience platforms without personal vulnerability (P06 (low): *"...give me some examples of what others experienced when they were cheated or hired..."*).

### 4.5.3 Educational and Accessible Support:

Ideas from this theme (Appendix Table A5) focused on both proactive and on-demand education. This also included suggestions for easily accessible and centralised knowledge resources (P01 (low): *"It would be great if there is more information about how can I deal with it and where I could report such scams."*), streamlined reporting processes to address universal barriers, and on-demand AI or human support without social exposure (P07 (low): *"To be fair, if I did find something specific, like very suspicious at the beginning and I didn't have someone to ask, I might have pasted some of what was told to me and asked ChatGPT"*). These suggestions reflected clear group differences: high SoAnx participants drove recommendations for proactive automation and anonymous community systems, tied to their vulnerability to curiosity-driven risk-taking and fear of negative evaluation, while low SoAnx participants clustered around information access and shared experiences. Together, they point to the need for security tools that account for how SoAnx characteristics heterogeneously interact with social engineering threats.

## 5 Discussion

Our findings suggest that SoAnx characteristics may increase resilience but also cause elevated susceptibility across the three explored social engineering attack scenarios. We use Spence and Rapee's evidence-based model of SoAnx to better understand how core characteristics may affect users in social engineering attacks. While our findings appear to be consistent with core aspects of SoAnx, we emphasise that our exploratory methodology limits definitive conclusions about these theoretical connections.

### 5.0.1 Enhanced Detection Of Potential Threats with Reassurance-Seeking Vulnerability

Trait socially anxious individuals have heightened attention to social threat cues as a protective mechanism [10, 15, 59]. Our findings suggest a similar pattern: High SoAnx participants focused on linguistic inconsistencies, tones, breaches of social norms, and social pressure tactics—particularly in situations with established conventions like professional job processes. However, attention bias links to excessive reassurance seeking in SoAnx [23]. Our findings suggest this creates vulnerability when attackers provide positive affirmation. High SoAnx participants responded more favorably to excessive praise than low SoAnx participants. Reassurance-seeking functions as a compensatory strategy to reduce anxiety from perceived negative evaluation [23]. When attackers provide

unsolicited affirmation, this temporarily satisfies reassurance needs, creating false safety that overrides SoAnx-enhanced threat detection. Similarly, curiosity-driven vulnerability in extortion reflects attention regulation difficulties [48]: while participants recognized link dangers, attention remained captured by potential threat content. This illustrates how attention bias manifests differently—protective when attacks exploit norm violations, but vulnerable when exploiting emotional needs or threat-verification urges. Beyond attention bias and reassurance seeking, safety behaviours may further explain our findings.

#### 5.0.2 Safety Behaviours as Protective Strategies with Limited Long-Term Effect

Safety behaviours are a prominent SoAnx characteristic, intended to prevent negative social outcomes and reduce anxiety [59]. However, over-reliance may interfere with social functioning [27]. Our findings indicate high SoAnx participants engaged in broader avoidance behaviours, including platform withdrawal and decreased social obligation to respond. They also used humour when sharing scam experiences—a protective strategy helping them communicate while reducing anxiety and deflecting judgment [27, 59]. These behaviours create a paradox: while humour may facilitate disclosure, it minimizes experience seriousness and may discourage support. Similarly, platform withdrawal protects against re-victimisation but eliminates access to platform-specific protections and community resources. Consistent with prior work [59], safety behaviours offer protection against certain social engineering risks while simultaneously reducing access to protective social resources for identifying and responding to future attacks.

#### 5.0.3 Individual Factors of Social Anxiety and Their Effects on the Strategy

Our findings suggest a complex pattern where participants with trait SoAnx showed both enhanced resilience and increased vulnerability across the three explored scenarios in study 2. This variance challenges models that position mental health needs as uniform security risk factors [31]. The same individual may appear to excel at detecting linguistic inconsistencies in one situation may also show increased susceptibility to positive affirmation due to elevated reassurance seeking. Theories about trait SoAnx emphasise both individual and environmental factors contributing to its development and maintenance. This results in different sensitivities and coping strategies. For example: some individuals may experience SoAnx in one particular social situation (e.g., giving a speech) while others may experience more general SoAnx across multiple situations. Our findings are consistent with this individualised perspective on SoAnx, where an individual may show different strategies based on their personal aspects of trait SoAnx. Taken together, our exploratory findings suggest that SoAnx characteristics may create context-dependent patterns of cybersecurity vulnerability and resilience that reflect the complex, individualised nature of trait SoAnx rather than uniform risk factors.

### 5.1 Implications for Design

While these insights require validation through systematic investigation, they indicate that effective cybersecurity interventions may need to account for how psychological factors like trait SoAnx interact with social engineering threats. Our findings reveal that trait SoAnx characteristics, such as attention bias to social cues [59], reassurance-seeking [23], fear of negative evaluation [39], and safety behaviours [27], may create context-dependent patterns of vulnerability and resilience. To support practitioners in designing more inclusive security tools, we translate identified patterns and user recommendations into three preliminary design considerations, acknowledging that specific implementation strategies require further investigation.

#### 5.1.1 Leveraging Detection Abilities While Mitigating Reassurance Vulnerabilities

Participants' suggestions for proactive support tools revealed a potential asymmetric protection problem: Security systems must account for divergent user capabilities and vulnerabilities across anxiety levels. Our findings suggest that high SoAnx participants' biased attention enabled detection of linguistic inconsistencies, but their reassurance-seeking created vulnerability to excessive praise—the same user showing both capabilities and vulnerabilities depending on attack type [23, 59]. Additionally, SoAnx characteristics such as fear of evaluation in performance situations [39] may affect how users engage with support systems themselves—for instance, reluctance to use tools requiring real-time decision-making or visible actions that might be judged by others. This suggests that systems need to accommodate for the heterogeneity of users: Highly SoAnx individuals may benefit from additional support at detecting extremely affirmative language. For example warning systems within communication platforms that flag signs of a threat could be enhanced with additional information about the tone of voice of the communication which go beyond technical anomalies.

#### 5.1.2 Community-Based Support Systems: Enabling Social Proof Without Social Exposure

Our participants' suggestions revealed an under explored tension in security design: the safety behaviours that SoAnx individuals employ for self-protection may paradoxically prevent them from accessing community-based security resources. While prior research has examined barriers to reporting a

crime and help-seeking of victims [43], our findings suggest that psychological safety behaviours may create additional, hidden obstacles to community participation: Highly SoAnx participants avoided sharing due to fear of negative evaluation rather than privacy concerns. This distinction matters for design because privacy-driven avoidance suggests different solutions than anxiety-driven avoidance: Safety behaviours (e.g., using humour to discuss scam experiences), may lead users to downplay or conceal actual attacks to avoid negative evaluation from others. Drawing on these findings, we propose the following considerations for community-based security features: First, interfaces should offer implicit social proof (e.g., aggregated threat statistics) to validate users without forcing personal disclosure. Second, system language must favour neutral framing over stigmatizing terminology (e.g., 'victim') to lower evaluation anxiety. Finally, designers should privilege asynchronous community resources over high-pressure synchronous interactions, carefully balancing user privacy with the need for collective security.

#### 5.1.3 Reducing Friction Through Automation and Non-Social Support

All participants reported barriers to formal reporting and help-seeking, with these obstacles compounding for SoAnx individuals. Beyond the social barriers discussed earlier, participants identified systemic friction points: lack of knowledge about reporting channels, skepticism about effectiveness, and high effort requirements. These universal barriers suggest that current support systems may be broadly inaccessible, not just for specific populations. However, high SoAnx participants' suggestions for automation also reflected avoidance of judgment from authority figures—suggesting automation addresses both universal friction and SoAnx-specific anxiety about social evaluation. Participants' suggestions for automated and educational tools point toward non-social pathways for security support. Automated systems that could examine suspicious content without user action might address both effort barriers and the curiosity-driven vulnerability some participants described. Educational resources accessible through AI or searchable databases could provide support without triggering the fear of judgment that prevented many participants from seeking help. This suggests potential value in AI-mediated preliminary assessment (e.g., allowing users to query suspicious content before escalating to human support) and automated evidence preservation that enables reporting when users are emotionally ready rather than requiring immediate action. Our findings reveal a temporal gap between attack occurrence and reporting readiness: high SoAnx participants reported that emotional distress often overrode judgment during incidents. Security systems should therefore offer staged support: automated mitigation during active compromise, followed by education and reporting frameworks once users have the cognitive capacity to engage with them.

### 5.2 Limitations and Future Work

Our work has several limitations. First, our small sample ($N = 20$) was gathered via convenience sampling within an academic context, risking demographic and digital literacy skews that limit generalizability. Future work should employ a broader, non-academic recruitment strategy to engage more diverse communities, utilizing larger sample sizes to evaluate how SoAnx influences vulnerabilities and resilience to social engineering attacks. Second, our role-playing design prioritized safety over active deception, sacrificing ecological validity. Evaluating explicitly labelled, hypothetical threats allows time for analytical reflection, bypassing the real-time emotional reasoning triggered by actual attacks [32]. Furthermore, the fixed scenario order may have introduced order effects, and the evaluative interview format itself may have heightened anxiety for some participants. To address these limitations ethically, future work should consider counterbalancing scenario presentation, or deploying longitudinal diary studies and retrospective analyses to capture naturalistic responses safely. Third, focusing exclusively on trait SoAnx omitted comorbid conditions or broader mental health characteristics. Future work should evaluate multiple aspects simultaneously to identify SoAnx-specific patterns. Fourth, our scenarios varied across multiple dimensions (e.g., relationship duration and progression), limiting cross-scenario comparisons and restricting our analysis to within-scenario patterns. Additionally, reporting barriers were context-dependent, varying from platform-native tools in romance scenarios to a complete lack of reporting infrastructure in extortion contexts. Future work should systematically isolate these dimensions (tactic, duration, relationship strength) and explicitly map context-specific reporting barriers to better quantify their individual impacts.

## 6 Conclusion

In this exploratory study, we examined whether and how social anxiety affects vulnerability to social engineering attacks through two studies. Study 1 gathered experiences from 42 participants who had experienced social engineering attacks, identifying common attack patterns that informed three scenarios for Study 2. Through role-playing interviews with 20 participants, we found that social anxiety characteristics create context-dependent vulnerabilities and resilience. While all participants shared barriers to reporting and seeking support, high social anxiety participants displayed unique patterns—attention bias and safety behaviours—that acted as protective or vulnerability factors depending on attack type. Participants suggested technology supports including automation and community-based systems. Our study contributes understanding of how social anxiety characteristics create both vulnerabilities and benefits in social engineering contexts, informing more inclusive security design.

## References


[1] Samar Muslah Albladi and George R. S. Weir. Predicting individuals' vulnerability to social engineering in social networks. *Cybersecurity*, 3(1):7, March 2020. ISSN 2523-3246. doi: 10.1186/s42400-020-00047-5. URL https://doi.org/10.1186/s42400-020-00047-5.

[2] Hussain Aldawood and Geoffrey Skinner. Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pages 62–68, December 2018. doi: 10.1109/TALE.2018.8615162. URL https://ieeexplore.ieee.org/document/8615162/. ISSN: 2470-6698.

[3] Sima Amirkhani, Fatemeh Alizadeh, Dave Randall, and Gunnar Stevens. Beyond Dollars: Unveiling the Deeper Layers of Online Romance Scams Introducing "Body Scam". In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, pages 1–6, New York, NY, USA, May 2024. Association for Computing Machinery. ISBN 979-8-4007-0331-7. doi: 10.1145/3613905.3651004. URL https://dl.acm.org/doi/10.1145/3613905.3651004.

[4] Katherine Amoresano and Benjamin Yankson. Human Error - A Critical Contributing Factor to the Rise in Data Breaches: A Case Study of Higher Education. *HOLISTICA – Journal of Business and Public Administration*, 14(1):110–132, June 2023. doi: 10.2478/hjbpa-2023-0007. URL https://sciendo.com/article/10.2478/hjbpa-2023-0007.

[5] American Psychiatric Association. Anxiety Disorders. In *Diagnostic and Statistical Manual of Mental Disorders*, DSM Library. American Psychiatric Association Publishing, March 2022. ISBN 978-0-89042-575-6. doi: 10.1176/appi.books.9780890425787.x05_Anxiety_Disorders. URL https://psychiatryonline.org/doi/full/10.1176/appi.books.9780890425787.x05_Anxiety_Disorders.

[6] Zinaida Benenson, Freya Gassmann, and Robert Landwirth. Unpacking Spear Phishing Susceptibility. In Michael Brenner, Kurt Rohloff, Joseph Bonneau, Andrew Miller, Peter Y.A. Ryan, Vanessa Teague, Andrea Bracciali, Massimiliano Sala, Federico Pintore, and Markus Jakobsson, editors, *Financial Cryptography and Data Security*, pages 610–627, Cham, 2017. Springer International Publishing. ISBN 978-3-319-70278-0. doi: 10.1007/978-3-319-70278-0_39.

[7] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, January 2006. ISSN 1478-0887. doi: 10.1191/1478088706qp063oa. URL https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[8] Virginia Braun and Victoria Clarke. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4):589–597, August 2019. ISSN 2159-676X. doi: 10.1080/2159676X.2019.1628806. URL https://doi.org/10.1080/2159676X.2019.1628806. Publisher: Routledge _eprint: https://doi.org/10.1080/2159676X.2019.1628806.

[9] Virginia Braun and Victoria Clarke. One size fits all? what counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology*, 18(3):328–352, 2021.

[10] Julia D Buckner, Jon K Maner, and Norman B Schmidt. Difficulty disengaging attention from social threat in social anxiety. *Cognitive therapy and research*, 34(1):99–105, 2010.

[11] Sanja Budimir, Johnny R J Fontaine, Nicole M A Huijts, Antal Haans, George Loukas, and Etienne B Roesch. Emotional Reactions to Cybersecurity Breach Situations: Scenario-Based Survey Study. *Journal of Medical Internet Research*, 23(5):e24879, May 2021. ISSN 1439-4456. doi: 10.2196/24879. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8156130/.

[12] Jan-Willem Hendrik Bullée, Lorena Montoya, Wolter Pieters, Marianne Junger, and Pieter Hartel. On the anatomy of social engineering attacks—A literature-based dissection of successful attacks. *Journal of Investigative Psychology and Offender Profiling*, 15(1):20–45, 2018. ISSN 1544-4767. doi: 10.1002/jip.1482. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jip.1482. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jip.1482.

[13] Mark Button, Chris Lewis, and Jacki Tapley. Not a victimless crime: The impact of fraud on individual victims and their families. *Security Journal*, 27(1):36–54, February 2014. ISSN 1743-4645. doi: 10.1057/sj.2012.11. URL https://doi.org/10.1057/sj.2012.11.

[14] Mark Button, Branislav Hock, David Shepherd, and Paul Gilmour. Understanding the rise of fraud in England and Wales through field theory: Blip or flip? *Journal of Economic Criminology*, 1:100012, September 2023. ISSN 2949-7914. doi: 10.1016/j.jeconc.2023.100012. URL https://www.sciencedirect.com/science/article/pii/S294979142300012X.

[15] David M. Clark and Adrian Wells. A cognitive model of social phobia. In *Social phobia: Diagnosis, assessment, and treatment*, pages 69–93. The Guilford Press, 1995.

[16] Joanne Davila and J Gayle Beck. Is social anxiety associated with impairment in close relationships? a preliminary investigation. *Behavior Therapy*, 33(3):427–446, 2002.

[17] Yue Deng, Changyang He, Yixin Zou, and Bo Li. "Auntie, Please Don't Fall for Those Smooth Talkers": How Chinese Younger Family Members Safeguard Seniors from Online Fraud. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pages 1–17, New York, NY, USA, April 2025. Association for Computing Machinery. ISBN 979-8-4007-1394-1. doi: 10.1145/3706598.3714137. URL https://dl.acm.org/doi/10.1145/3706598.3714137.

[18] Julie S. Downs, Mandy Holbrook, and Lorrie Faith Cranor. Behavioral response to phishing risk. In *Proceedings of the Anti-Phishing Working Groups 2nd Annual ECrime Researchers Summit*, eCrime '07, page 37–44, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595939395. doi: 10.1145/1299015.1299019. URL https://doi.org/10.1145/1299015.1299019.

[19] Jacqueline M. Drew and Julianne Webster. The victimology of online fraud: A focus on romance fraud victimisation. *Journal of Economic Criminology*, 3:100053, March 2024. ISSN 2949-7914. doi: 10.1016/j.jeconc.2024.100053. URL https://www.sciencedirect.com/science/article/pii/S2949791424000058.

[20] Edwin Donald Frauenstein and Stephen Flowerday. Susceptibility to phishing on social network sites: A personality information processing model. *Computers & security*, 94:101862, 2020.

[21] Robin Gandhi, Anup Sharma, William Mahoney, William Sousan, Qiuming Zhu, and Phillip Laplante. Dimensions of Cyber-Attacks: Cultural, Social, Economic, and Political. *IEEE Technology and Society Magazine*, 30(1):28–38, 2011. ISSN 1937-416X. doi: 10.1109/MTS.2011.940293. URL https://ieeexplore.ieee.org/document/5725605.

[22] Google. Google Gemini. URL https://gemini.google.com.

[23] DeMond M Grant, Matt R Judah, Adam C Mills, William V Lechner, Collin L Davidson, and LaRicka R Wingate. Rumination and excessive reassurance seeking: Mediators of the relationship between social anxiety and depression? *Journal of Psychopathology and Behavioral Assessment*, 36(3):465–474, 2014.

[24] Christopher Hadnagy. *Social Engineering: The Science of Human Hacking*. John Wiley & Sons, Inc., June 2018. ISBN 978-1-119-43338-5. URL https://onlinelibrary.wiley.com/doi/epub/10.1002/9781119433729. ISBN: 9781119433729.

[25] Tzipora Halevi, James Lewis, and Nasir Memon. A pilot study of cyber security and privacy related behavior and personality traits. In *Proceedings of the 22nd international conference on world wide web*, pages 737–744, 2013.

[26] R. G. Heimberg, K. J. Horner, H. R. Juster, S. A. Safren, E. J. Brown, F. R. Schneier, and M. R. Liebowitz. Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine*, 29 (1):199–212, January 1999. ISSN 1469-8978, 0033-2917. doi: 10.1017/S0033291798007879. URL https://www.cambridge.org/core/journals/psychological-medicine/article/psychometric-properties-of-the-liebowitz-social-anxiety-scale/6891D37D00A9BEC179E61C8BFF30F08A.

[27] Richard G. Heimberg, Faith A. Brozovich, and Ronald M. Rapee. Chapter 15 - A Cognitive Behavioral Model of Social Anxiety Disorder: Update and Extension. In Stefan G. Hofmann and Patricia M. DiBartolo, editors, *Social Anxiety (Second Edition)*, pages 395–422. Academic Press, San Diego, January 2010. ISBN 978-0-12-375096-9. doi: 10.1016/B978-0-12-375096-9.00015-8. URL https://www.sciencedirect.com/science/article/pii/B9780123750969000158.

[28] Kyung Wha Hong, Christopher M Kelley, Rucha Tembe, Emerson Murphy-Hill, and Christopher B Mayhorn. Keeping up with the joneses: Assessing phishing susceptibility in an email task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 57, pages 1012–1016. SAGE Publications Sage CA: Los Angeles, CA, 2013.

[29] Markus Huber, Martin Mulazzani, Sebastian Schrittwieser, and Edgar Weippl. Cheap and automated socio-technical attacks based on social networking sites. In *Proceedings of the 3rd ACM workshop on Artificial intelligence and security*, pages 61–64, Chicago Illinois USA, October 2010. ACM. ISBN 978-1-4503-0088-9. doi: 10.1145/1866423.1866435. URL https://dl.acm.org/doi/10.1145/1866423.1866435.

[30] Nicholas Hutchins, Andrew Allen, Michelle Curran, and Lee Kannis-Dymand. Social anxiety and online social interaction. *Australian Psychologist*, 56(2):142–153, 2021. doi: 10.1080/00050067.2021.1890977. URL https://doi.org/10.1080/00050067.2021.1890977.

[31] Eric C Ip and Daisy Cheung. Global mental health security—time for action. In *JAMA Health Forum*, volume 1, pages e200622–e200622. American Medical Association, 2020.

[32] Daniel Kahneman. *Thinking, fast and slow*. macmillan, 2011.

[33] Tara A. Karasewich and Valerie A. Kuhlmeier. Trait social anxiety as a conditional adaptation: A developmental and evolutionary framework. *Developmental Review*, 55:100886, 2020. ISSN 0273-2297. doi: https://doi.org/10.1016/j.dr.2019.100886. URL https://www.sciencedirect.com/science/article/pii/S027322971930125X.

[34] Eley Suzana Kasim, Salwa Muda, Norlaila Md Zin, Hazlina Mohd Padil, Norhidayah Ismail, and Sharifah Norzehan Syed Yusuf. Combating investment scams: insights from law enforcement and civil society toward a prevention framework. *Journal of Criminological Research, Policy and Practice*, June 2025. ISSN 2056-3841. doi: 10.1108/JCRPP-04-2025-0030. URL https://doi.org/10.1108/JCRPP-04-2025-0030.

[35] Kurt Kroenke, Robert L. Spitzer, and Janet B. W. Williams. The PHQ-9. *Journal of General Internal Medicine*, 16(9):606–613, September 2001. ISSN 1525-1497. doi: 10.1046/j.1525-1497.2001.016009606.x. URL https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

[36] Sumeet Kumar and Kathleen M. Carley. Approaches to understanding the motivations behind cyber attacks. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, pages 307–309, September 2016. doi: 10.1109/ISI.2016.7745496. URL https://ieeexplore.ieee.org/document/7745496/.

[37] Chin Kuo and Seng-Su Tsang. Constructing an Investment Scam Detection Model Based on Emotional Fluctuations Throughout the Investment Scam Life Cycle. *Deviant Behavior*, 45(2):204–225, February 2024. ISSN 0163-9625. doi: 10.1080/01639625.2023.2244115. URL https://doi.org/10.1080/01639625.2023.2244115. Publisher: Routledge _eprint: https://doi.org/10.1080/01639625.2023.2244115.

[38] Dr. Sharon L. Burton and Dr. Pamela D. (Vickerson) Moore. Pig Butchering in Cybersecurity: A Modern Social Engineering Threat. *SocioEconomic Challenges*, 8(3):46–60, October 2024. ISSN 25206621, 25206214. doi: 10.61093/sec.8(3).46-60.2024. URL https://armgpublishing.com/journals/sec/volume-8-issue-3/article-4/.

[39] Michael R. Liebowitz. Social phobia. *Modern problems of pharmacopsychiatry*, 22:141—-173, 1987. doi: 10.1159/000414022.

[40] Tian Lin, Daniel E. Capecci, Donovan M. Ellis, Harold A. Rocha, Sandeep Dommaraju, Daniela S. Oliveira, and Natalie C. Ebner. Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Transactions on Computer-Human Interaction*, 26(5):1–28, October 2019. ISSN 1073-0516, 1557-7325. doi: 10.1145/3336141. URL https://dl.acm.org/doi/10.1145/3336141.

[41] Lumivero. NVivo 14, 2023. URL https://lumivero.com/products/nvivo/.

[42] Gibson Mba, Jeremiah Onaolapo, Gianluca Stringhini, and Lorenzo Cavallaro. Flipping 419 Cybercrime Scams: Targeting the Weak and the Vulnerable. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, pages 1301–1310, Republic and Canton of Geneva, CHE, April 2017. International World Wide Web Conferences Steering Committee. ISBN 978-1-4503-4914-7. doi: 10.1145/3041021.3053892. URL https://dl.acm.org/doi/10.1145/3041021.3053892.

[43] Michael R. McCart, Daniel W. Smith, and Genelle K. Sawyer. Help seeking among victims of crime: A review of the empirical literature. *Journal of Traumatic Stress*, 23(2):198–206, 2010. doi: https://doi.org/10.1002/jts.20509. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jts.20509.

[44] Nora McDonald and Helena M. Mentis. 'Don't Fall for This': Communications about Cybersafety from the AARP. *Proc. ACM Hum.-Comput. Interact.*, 7(CSCW2):248:1–248:21, October 2023. doi: 10.1145/3610039. URL https://dl.acm.org/doi/10.1145/3610039.

[45] Beatriz Mello, Robin Welsch, Marissa Christien Verbokkem, Pascal Knierim, and Martin Johannes Dechant. Navigating the virtual gaze: Social anxiety's role in vr proxemics. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2024.

[46] Helena M. Mentis, Galina Madjaroff, and Aaron K. Massey. Upside and Downside Risk in Online Security for Older Adults with Mild Cognitive Impairment. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–13, New York, NY, USA, May 2019. Association for Computing Machinery. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300573. URL https://dl.acm.org/doi/10.1145/3290605.3300573.

[47] Thomas Mildner, Daniel Fidel, Evropi Stefanidi, Paweł W. Woźniak, Rainer Malaka, and Jasmin Niess. A Comparative Study of How People With and Without ADHD Recognise and Avoid Dark Patterns on Social Media. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pages 1–17, New York, NY, USA, April 2025. Association for Computing Machinery. ISBN 979-8-4007-1394-1. doi: 10.1145/3706598.3713776. URL https://dl.acm.org/doi/10.1145/3706598.3713776.

[48] Amanda S. Morrison and Richard G. Heimberg. Social Anxiety and Social Anxiety Disorder. *Annual Review of Clinical Psychology*, 9(Volume 9, 2013):249–274, March 2013. ISSN 1548-5943, 1548-5951. doi: 10.1146/annurev-clinpsy-050212-185631. URL https://www.annualreviews.org/content/journals/10.1146/annurev-clinpsy-050212-185631. Publisher: Annual Reviews.

[49] Francois Mouton, Mercia M. Malan, Louise Leenen, and H.S. Venter. Social engineering attack framework. In *2014 Information Security for South Africa*, pages 1–9, 2014. doi: 10.1109/ISSA.2014.6950510.

[50] Francois Mouton, Louise Leenen, and Hein S Venter. Social engineering attack examples, templates and scenarios. *Computers & Security*, 59:186–209, 2016.

[51] James Nicholson, Lynne Coventry, and Pam Briggs. Can we fight social engineering attacks by social means? assessing social salience as a means to improve phish detection. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 285–298, 2017.

[52] University of Exeter School of Psychology. The psychology of scams: Provoking and committing errors of judgement. *Office of Fair Trading*, May 2009. URL https://ore.exeter.ac.uk/repository/bitstream/handle/10871/20958/OfficeOfFairTrading%202009.pdf%3bsequence%3d1.

[53] Marilyne Ordekian, Antonis Papasavva, Enrico Mariconti, and Marie Vasek. A Sinister Fattening: Dissecting the Tales of Pig Butchering and Other Cryptocurrency Scams. In *2024 APWG Symposium on Electronic Crime Research (eCrime)*, pages 136–148, September 2024. doi: 10.1109/eCrime66200.2024.00016. URL https://ieeexplore.ieee.org/document/10896045. ISSN: 2159-1245.

[54] Aswati Panicker, Novia Nurain, Zaidat Ibrahim, Chun-Han Wang, Seung Wan Ha, Yuxing Wu, Kay Connelly, Katie A Siek, and Chia-Fang Chung. Understanding fraudulence in online qualitative studies: From the researcher's perspective. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2024.

[55] Katalin Parti and Faika Tahir. "If We Don't Listen to Them, We Make Them Lose More than Money:" Exploring Reasons for Underreporting and the Needs of Older Scam Victims. *Social Sciences*, 12(5):264, April 2023. ISSN 2076-0760. doi: 10.3390/socsci12050264. URL https://www.mdpi.com/2076-0760/12/5/264.

[56] Ruba Rum, Jessica A. Birg, Gabriella Silva, Jonathan Rottenberg, Peter E. Clayson, and Fallon R. Goodman. Social Motivation Differentiates Social Anxiety and Depression: A Daily Diary Study. *Behavior Therapy*, 55(4):698–711, July 2024. ISSN 0005-7894. doi: 10.1016/j.beth.2023.11.004. URL https://www.sciencedirect.com/science/article/pii/S0005789423001302.

[57] Nina K. Rytwinski, David M. Fresco, Richard G. Heimberg, Meredith E. Coles, Michael R. Liebowitz, Shadha Cissell, Murray B. Stein, and Stefan G. Hofmann. Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. *Depression and Anxiety*, 26(1):34–38, 2009. ISSN 1520-6394. doi: 10.1002/da.20503. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/da.20503. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/da.20503.

[58] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 373–382, 2010.

[59] Susan H. Spence and Ronald M. Rapee. The etiology of social anxiety disorder: An evidence-based model. *Behaviour Research and Therapy*, 86:50–67, 2016. doi: 10.1016/j.brat.2016.06.007.

[60] Robert L. Spitzer, Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166(10):1092–1097, May 2006. ISSN 0003-9926. doi: 10.1001/archinte.166.10.1092. URL https://doi.org/10.1001/archinte.166.10.1092.

[61] Frank Stajano and Paul Wilson. Understanding scam victims: seven principles for systems security. *Commun. ACM*, 54(3):70–75, March 2011. ISSN 0001-0782. doi: 10.1145/1897852.1897872. URL https://dl.acm.org/doi/10.1145/1897852.1897872.

[62] Dan J. Stein, Carmen C. W. Lim, Annelieke M. Roest, Peter de Jonge, Sergio Aguilar-Gaxiola, Ali Al-Hamzawi, Jordi Alonso, Corina Benjet, Evelyn J. Bromet, Ronny Bruffaerts, Giovanni de Girolamo, Silvia Florescu, Oye Gureje, Josep Maria Haro, Meredith G. Harris, Yanling He, Hristo Hinkov, Itsuko Horiguchi, Chiyi Hu, Aimee Karam, Elie G. Karam, Sing Lee, Jean-Pierre Lepine, Fernando Navarro-Mateu, Beth-Ellen Pennell, Marina Piazza, Jose Posada-Villa, Margreet ten Have, Yolanda Torres, Maria Carmen Viana, Bogdan Wojtyniak, Miguel Xavier, Ronald C. Kessler, Kate M. Scott, and WHO World Mental Health Survey Collaborators. The cross-national epidemiology of social anxiety disorder: Data from the World Mental Health Survey Initiative. *BMC Medicine*, 15(1):143, July 2017. ISSN 1741-7015. doi: 10.1186/s12916-017-0889-2. URL https://doi.org/10.1186/s12916-017-0889-2.

[63] Wenni Syafitri, Zarina Shukur, Umi Asma' Mokhtar, Rossilawati Sulaiman, and Muhammad Azwan Ibrahim. Social Engineering Attacks Prevention: A Systematic Literature Review. *IEEE Access*, 10:39325–39343, 2022. ISSN 2169-3536. doi: 10.1109/ACCESS.2022.3162594. URL https://ieeexplore.ieee.org/abstract/document/9743471.

[64] Hooi Koon Tan and Yoong David. Preying on lonely hearts: A systematic deconstruction of an Internet romance scammer's online lover persona. *Journal of Modern Languages*, 23(1):28–40, 2013. ISSN 2462-1986. URL https://jml.um.edu.my/index.php/JML/article/view/3288.

[65] Christof Ferreira Torres, Mathis Baden, and Radu State. Towards Usable Protection Against Honeypots. In *2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 1–2, May 2020. doi: 10.1109/ICBC48266.2020.9169460. URL https://ieeexplore.ieee.org/document/9169460.

[66] Raj Vardhan, Alok Chandrawal, Phakpoom Chinprutthiwong, Yangyong Zhang, and Guofei Gu. #DM-Me: Susceptibility to Direct Messaging-Based Scams. In *Proceedings of the ACM Asia Conference on Computer and Communications Security*, pages 494–508, Melbourne VIC Australia, July 2023. ACM. ISBN 979-8-4007-0098-9. doi: 10.1145/3579856.3582815. URL https://dl.acm.org/doi/10.1145/3579856.3582815.

[67] Fangzhou Wang. Victim-offender overlap: the identity transformations experienced by trafficked Chinese workers escaping from pig-butchering scam syndicate. *Trends in Organized Crime*, February 2024. ISSN 1936-4830. doi: 10.1007/s12117-024-09525-5. URL https://doi.org/10.1007/s12117-024-09525-5.

[68] Zuoguang Wang, Limin Sun, and Hongsong Zhu. Defining Social Engineering in Cybersecurity. *IEEE Access*, 8:85094–85115, 2020. ISSN 2169-3536. doi: 10.1109/ACCESS.2020.2992807. URL https://ieeexplore.ieee.org/document/9087851/.

[69] Jack M. Whittaker, Suleman Lazarus, and Taidgh Corcoran. Are fraud victims nothing more than animals? Critiquing the propagation of "pig butchering" (Sha Zhu Pan. *Journal of Economic Criminology*, 3:100052, March 2024. ISSN 2949-7914. doi: 10.1016/j.jeconc.2024.100052. URL https://www.sciencedirect.com/science/article/pii/S2949791424000046.

[70] Monica T. Whitty. The Scammers Persuasive Techniques Model: Development of a Stage Model to Explain the Online Dating Romance Scam. *The British Journal of Criminology*, 53(4):665–684, July 2013. ISSN 0007-0955. doi: 10.1093/bjc/azt009. URL https://doi.org/10.1093/bjc/azt009.

[71] Monica Therese Whitty. Is There a Scam for Everyone? Psychologically Profiling Cyberscam Victims. *European Journal on Criminal Policy and Research*, 26(3):399–409, September 2020. ISSN 1572-9869. doi: 10.1007/s10610-020-09458-z. URL https://doi.org/10.1007/s10610-020-09458-z.

[72] Ziyi Xie and Zhizhuang Duan. "Why did I fall for it?" Exploring internet fraud susceptibility in the pig butchering scam. *Security Journal*, 38(1):9, December 2024. ISSN 1743-4645. doi: 10.1057/s41284-024-00457-x. URL https://doi.org/10.1057/s41284-024-00457-x.

[73] Mohamed Zaoui, Belfaik Yousra, Sadqi Yassine, Maleh Yassine, and Ouazzane Karim. A comprehensive taxonomy of social engineering attacks and defense mechanisms: Toward effective mitigation strategies. *IEEE access*, 12:72224–72241, 2024.

[74] Yuxiang Zhai, Xiao Xue, Zekai Guo, Tongtong Jin, Yuting Diao, and Jihong Jeung. Hear Us, then Protect Us: Navigating Deepfake Scams and Safeguard Interventions with Older Adults through Participatory Design. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pages 1–19, New York, NY, USA, April 2025. Association for Computing Machinery. ISBN 979-8-4007-1394-1. doi: 10.1145/3706598.3714423. URL https://dl.acm.org/doi/10.1145/3706598.3714423.

[75] Sarah Y Zheng and Ingolf Becker. Checking, nudging or scoring? evaluating e-mail user security tools. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, pages 57–76, 2023.

# A Appendix

## A.1 Open Materials

All study materials used in Study 2 are available in our OSF repository: https://osf.io/z6j9m/?view_only=6d08aa4c874b484d8a56a6190086a501

## A.2 Additional Tables Study 2

In the following section we provide detailed overviews about the themes and results of the two qualitative studies.

### A.2.1 Participant Demographics

| Statistic | LSAS | GAD-7 | PHQ-9 |
|---|---|---|---|
| Min | 17 | 0 | 0 |
| Max | 88 | 18 | 14 |
| Median | 50.50 | 5.50 | 5 |
| Mean | 51.95 | 6.00 | 6.00 |
| SD | 23.26 | 5.34 | 4.17 |

*Note: SD represents Sample Standard Deviation.*

Table A1: Descriptive statistics for psychological scales from 20 participants from Study 2
.

| Participant ID | LSAS Score | Level of Social Anxiety Group | Age (Years) | Gender | Experienced Social Engineering Atack Previously |
|---|---|---|---|---|---|
| P01 | 37 | low | 22 | Female | Yes |
| P02 | 76 | high | 22 | Male | Yes |
| P03 | 84 | high | 19 | Female | No |
| P04 | 77 | high | 29 | Female | No |
| P05 | 18 | low | 24 | Female | Yes |
| P06 | 46 | low | 27 | Female | Yes |
| P07 | 29 | low | 23 | Female | No |
| P08 | 17 | low | 21 | Female | No |
| P09 | 57 | high | 21 | Female | No |
| P10 | 72 | high | 35 | Male | Yes |
| P11 | 65 | high | 26 | Male | Yes |
| P12 | 42 | low | 25 | Male | No |
| P13 | 62 | high | 25 | Female | Yes |
| P14 | 40 | low | 24 | Male | Yes |
| P15 | 38 | low | 31 | Male | Yes |
| P16 | 36 | low | 24 | Male | Yes |
| P17 | 17 | low | 35 | Female | Yes |
| P18 | 55 | high | 51 | Male | Yes |
| P19 | 83 | high | 26 | Female | Yes |
| P20 | 88 | high | 31 | Male | Yes |

Table A2: Participant statistics from role-playing and interview.

### A.2.2 Study 2: Themes within Baseline Vulnerabilities and Social Anxiety Vulnerability Factors

| Group | Theme | Description | N Lo | N Hi | Sample Quote |
|---|---|---|---|---|---|
| **Baseline Vulnerability: Support Seeking** | Lack of Trust, Knowledge, and Ease of Access to Authorities | Participants did not consider reporting scams due to lack of knowledge, low trust in outcomes, and perceived process difficulty. These barriers were stronger among high SoAnx participants. | 36 | 23 | *"Just knowing... what can they actually do? That's kind of putting me off." – P02 (high)* |
| | Judgment and Reprimand from Social Circle | Perceived judgment or reprimand from support systems reduced willingness to share concerns. | 3 | 16 | *"I would probably tell just one or two people that I trust. Other than that, no, I'm taking this to the grave." – P04 (high)* |
| | Personal Privacy for Participant | Personal privacy was cited as a reason for not discussing the situation with others. | 8 | 1 | *"I feel like my instinct wouldn't be to go to my friends... this is just a private conversation." – P06 (low)* |
| | Lack of Sharing Due to Social Irrelevance | Participants felt no need to share when the situation seemed socially irrelevant or low impact. | 3 | 5 | *"I would feel comfortable, but I wouldn't really feel like there's a need." – P04 (high)* |
| **SoAnx-Specific Vulnerability Factors** | Positive Perception of "Friendly" Language | Friendly or supportive language evoked comfort and confidence rather than suspicion, potentially encouraging continued interaction. | 3 | 6 | *"I find the last two sentences quite a bit of an extra nudge for comfort..." – P03 (high)* |
| | Knowledge of Danger Overwritten by Emotions | Although aware of risks (e.g., unknown links), some participants still felt compelled to click due to curiosity or emotional pull. | 1 | 3 | *"I probably click the link there and see what it is." – P11 (high)* |
| | Desperation from Pressures of Unemployment | Internal pressure or prolonged unemployment created desperation, potentially reducing critical evaluation. | 3 | 6 | *"If I've been unemployed for 14 months... I'm probably pretty desperate." – P11 (high)* |
| | Humour as a Sharing Method | Participants reframed scam experiences humorously to reduce embarrassment and manage social image. | 2 | 5 | *"If I can spin it to be less pathetic and more like funny." – P11 (high)* |
| | General Increased Vigilance / Avoidance Behaviour | While many reported increased vigilance after a scam, some high SoAnx participants described stronger avoidance of similar platforms or situations. | 0 | 6 | *"I may not give him a chance because of this negative experience... I feel I have a strong need to protect myself." – P20 (high)* |

Table A3: Themes within Baseline and Social Anxiety Vulnerability Factors, grouped by whether they reflect universal support-seeking barriers or SoAnx-specific emotional and behavioural vulnerabilities. N Lo = N Low SoAnx inputs; N Hi = N High SoAnx inputs.

### A.2.3 Study 2: Themes Within Social Anxiety as Potential Resilience Factor

| Group | Theme | Description | N Lo | N Hi | Sample Quote |
|---|---|---|---|---|---|
| **Detection of Social Norm Violations** | Unsolicited Information Sharing and Pet Names | Participants are alerted by unnecessary information sharing from the fictitious sender and the use of intimate nicknames. | 4 | 14 | *"Maybe it's just me, but like if we haven't met and they start calling me like nicknames. This is kind of kind of weird." – P11 (high)* |
| | Vigilance from External Pressure and Urgency | Participants mention feeling urgency or pressure in role-playing messages, which may put them on alert. | 3 | 7 | *"...I do think the only thing that is a little weird is that they were like, please sign it as soon as possible. I guess it's weird to rush instead of just giving a deadline." – P09 (high)* |
| **Linguistic & Information Inconsistency Detection** | Detection of Discrepancies and Tonal Anomalies | Participants noticed and questioned inconsistencies such as grammar issues and tonal shifts, leading to suspicion and wariness. | 9 | 15 | *"It feels a little bit generic. And the second one, where they asked me for money feels very direct... And then the last one, they're just like, hey, I need you to wire me this amount of money." – P11 (high)* |
| | Vigilance on Missing Information | Participants identified missing or incomplete logistical information (e.g., payment details), which raised suspicion. | 1 | 4 | *"You have one hour to give me two thousand pounds, but nowhere to send that money to. How am I supposed to give you the money?" – P10 (high)* |
| **Avoidance of Social Interactions** | Probing for Information and Clarification | While all participants conducted research on the fictitious persona, some showed more proactive probing behaviour to gain additional clarification. | 11 | 6 | *"And I would probably just ask that question, like what is the position? And maybe like, you know, what's the opening?" – P06 (low)* |
| | Social Obligation to Respond | Many participants felt a social obligation to respond due to the norms of the dating application context. | 7 | 3 | *"I would given it's a dating app, I feel like I would have been more receptive of just like a hey, how are you doing?" – P06 (low)* |

Table A4: Themes within Social Anxiety as Potential Resilience Factors, grouped by detection of social norm violations, linguistic and information inconsistencies, and avoidance of social interactions. N Lo = N Low SoAnx inputs; N Hi = N High SoAnx inputs.

### A.2.4 Study 2: Themes About Technology Suggestions

| Group | Main Theme | Theme Description | Sub-theme | Sub-theme Description | Sample Quote |
|---|---|---|---|---|---|
| **Automation and Risk Detection** | **Proactive and Automated Security Measures** | Participant's vigilance and knowledge are strategies used to avoid scams; however, emotions may override logical decision-making. | Automated Filtering | The system should automatically filter or block suspicious messages, similar to an email spam folder. | *"I think apps should be able to like maybe regularly be scanning their profiles for like scammers or like regularly be or at least bots." – P04 (high)* |
| | | | Pre-Screening and Information Summary | The technology could verify user profiles, scan for reused photos, analyze unknown links, and provide summaries including credibility, activity, and sender details. | *"If something can click into that link for me and then relay the content to me, it would be pretty nice." – P11 (high)* |
| | | | Risk Indicators and Warnings | The system should provide real-time warnings about suspicious behavior, such as unusual language or sudden financial requests. | *"I think that there should be messages, warning messages for people. Letting people don't ask for visa gift cards. This isn't normal." – P10 (high)* |
| **Community Support** | **Collective Knowledge and Sharing for Support** | This theme highlights participants' desire for social proof and shared experiences to help identify scams and reduce the mental burden of verifying concerns independently. | – | No sub-themes identified. | *"I want to see the comments of others to this person on a dating app like 90% of the person give the some for this person and like a 10% of person think she is a genuine person." – P05 (low)* |
| **Education and Access** | **Educational and Accessible Support** | This theme highlights the desire for educational resources and accessible support mechanisms to identify and respond to scams. | Knowledge-Building Resources | The platform should provide educational guidance, such as how to engage safely with job applications, expected salary ranges, and reporting channels. | *"It would be great if there is more information about how can I deal with it and where I could report such scams... That would be really helpful..." – P01 (low)* |
| | | | On-Demand Support | A simple way to access help from a human or AI model when feeling uncertain about a suspicious interaction. | *"To be fair, if I did find something specific, like very suspicious at the beginning and I didn't have someone to ask, I might have pasted some of what was told to me and asked ChatGPT." – P07 (low)* |

Table A5: Merged themes and sub-themes identified for technology suggestions from participants, grouped by automation and risk detection, community support, and education and access.

## A.3 Declaration of AI use

During the preparation of this work, the authors used generative AI tools, including OpenAI GPT-5 and Claude Sonnet 4, to assist with editing, such as drafting sentences from bullets or refining text, or code and polishing grammar and style. Additionally, Gemini 2.5 Flash, was used to develop images and initial messages within Study 2. The authors take full responsibility for verifying, reviewing, and editing all generated content, including text, code, citations, and any outputs derived from these tools, and attest that the final publication reflects their original intellectual contributions and is free from falsified, plagiarized, or misrepresented content.