TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Haider Rehman 234876IVCM

The Importance and Effectiveness of Cybersecurity Awareness Training for Pakistani Students

Master's Thesis

Supervisor: Ricardo G. Lugo

PhD.

Tallinn 2026

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Haider Rehman 234876IVCM

Küberjulgeoleku teadlikkuse koolituse tähtsus ja tõhusus Pakistani üliõpilaste jaoks

Magistritöö

Juhendaja: Ricardo G. Lugo

PhD.

Tallinn 2026

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Haider Rehman

20.05.2026

Abstract

Cybersecurity incidents are increasingly traced to human rather than purely technical failure [1]. Evidence from Pakistan indicates that university students show gaps in practical threat recognition [9][11]. The Pakistani research literature has documented this vulnerability through surveys but has not experimentally tested whether a targeted training intervention can reduce it.

This thesis reports a quasi-experimental study with 50 participants from seven universities in Islamabad. Participants were allocated at cohort level to an experimental group or a control group. The experimental group received a 90-minute interactive workshop on phishing, social engineering, and disinformation. The control group received a matched, content-neutral session on digital infrastructure. Both groups completed parallel forms of an instrument adapted from the HAIS-Q at baseline, immediately after the session, and at a two-week follow-up. Forty-three participants completed all three occasions.

The Group by Time interaction on the combined total score returned $F(1.27, 51.91) = 30.91$, $p < 0.001$, partial eta squared (η^2p) 0.430 after Greenhouse-Geisser correction. All three null hypotheses are rejected. The between-group effect at the immediate post-test was Hedges' $g = +2.24$. At the two-week follow-up it was $g = +0.98$. Approximately 47% of the immediate gain persisted. The effect was not uniform across subscales. Knowledge and attitude gains were largest ($\eta^2p = 0.480$) but retained only 39% at follow-up. Behaviour gains were intermediate (0.246) with 59% retention. Applied scenario gains were smallest (0.020, not significant) but showed the highest retention at 77%. This retention inversion indicates that declarative gains decay faster than procedural ones.

The study makes three contributions. It provides an empirical baseline of cybersecurity awareness for a previously under-studied population. It reports a controlled experimental test of a short intervention in that setting. It identifies a retention inversion across subscales with implications for the design of cybersecurity training in Pakistani higher education.

This thesis is written in English and is 120 pages long, including 6 chapters, 12 figures and 25 tables.

Annotatsioon

Küberintsidentide põhjuseks peetakse üha sagedamini inimlikku eksimust, mitte puhtalt tehnilist riket [1]. Pakistani andmed näitavad, et üliõpilastel esineb puudujääke praktilises ohtude äratundmises [9], [11]. Pakistani teaduskirjandus on seda haavatavust küsitluste abil dokumenteerinud, kuid ei ole eksperimentaalselt testinud, kas suunatud koolitussekkumine suudab seda vähendada.

Käesolev magistritöö kajastab kvasieksperimentaalset uuringut 50 osalejaga seitsmest Islamabadi ülikoolist. Osalejad jaotati kohordi tasandil katse- või kontrollrühma. Katserühm osales 90-minutilises interaktiivses töötoas andmepüügi, sotsiaalse manipuleerimise ja desinformatsiooni teemal. Kontrollrühm osales samaväärses, sisult neutraalses seansis digitaalse infrastruktuuri teemal. Mõlemad rühmad täitsid HAIS-Q põhjal kohandatud mõõtevahendi paralleelvorme algtasemel, vahetult pärast seanssi ja kahenädalase järelkontrolli käigus. 43 osalejat läbisid kõik kolm mõõtmiskorda.

Rühma ja aja koosmõju kombineeritud koguskooril andis pärast Greenhouse-Geisseri parandust tulemuse $F(1{,}27;\ 51{,}91) = 30{,}91$, $p < 0{,}001$, osaline eeta ruut (η^2p) = 0,430. Kõik kolm nullhüpoteesi on tagasi lükatud. Rühmadevahelise efekti suurus vahetul järeltestil oli Hedgesi $g = +2{,}24$. Kahenädalasel järelkontrollil oli see $g = +0{,}98$. Ligikaudu 47% vahetust võidust säilis. Efekt ei olnud alaskaalade lõikes ühtlane. Teadmiste ja hoiakute võit oli suurim ($\eta^2p = 0{,}480$), kuid järelkontrollis säilis sellest vaid 39%. Käitumise võit oli keskmine (0,246) ja säilis 59% ulatuses. Rakenduslike stsenaariumide võit oli väikseim (0,020; statistiliselt mitteoluline), kuid näitas kõrgeimat säilimist (77%). Selline säilimise inversioon näitab, et deklaratiivsed võidud kahanevad kiiremini kui protseduurilised.

Uuring annab kolm panust. See pakub empiirilist lähtetaset küberturvalisuse teadlikkuse kohta varem vähe uuritud populatsioonis. See kajastab kontrollitud eksperimentaalset testi lühikese sekkumise kohta samas kontekstis. See tuvastab alaskaalade lõikes säilimise inversiooni, millel on mõju küberturvalisuse koolituse kavandamisele Pakistani kõrghariduses.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 120 leheküljel, 6 peatükki, 12 joonist, 25 tabelit

List of abbreviations and terms

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CI	Confidence Interval
CS	Computer Science
CTL	Control group
EXP	Experimental group
FIA	Federal Investigation Agency (Pakistan)
GCI	Global Cybersecurity Index
GG	Greenhouse-Geisser (sphericity correction)
HAIS-Q	Human Aspects of Information Security Questionnaire
HEC	Higher Education Commission (Pakistan)
Hedges' g	Bias-corrected standardised mean difference (effect size)
ITU	International Telecommunication Union
KAB	Knowledge-Attitude-Behaviour (model)
MRQ	Main Research Question
NCCS	National Centre for Cyber Security (Pakistan)
PECA	Prevention of Electronic Crimes Act 2016 (Pakistan)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised Controlled Trial
SLR	Systematic Literature Review
SRQ	Sub-Research Question
η^2p	Partial eta squared (effect size for ANOVA)
Cohen's d	Standardised mean difference (within-group effect size).
CSIS	Center for Strategic and International Studies
PICOC	Population, Intervention, Comparison, Outcome, Context.
SD	Standard Deviation

Table of Contents

List of figures

List of tables

1 Introduction

The increasing dependence of modern societies on digital infrastructure has made cybersecurity a concern that extends well beyond information technology professionals. Individuals interact with networked systems in nearly every aspect of daily life, from communication and education to banking and governance. As this dependence grows, so does the opportunity for malicious actors to exploit the people who use these systems. The Verizon 2024 Data Breach Investigations Report analysed more than 30,000 security incidents globally. It found that 68% of confirmed breaches involved a non-malicious human element [1]. This proportion has remained broadly stable over successive years of the report, suggesting that technical controls alone have not substantially reduced the risk posed by human behaviour [1].

Parsons et al. noted that threats to information security cannot be prevented by focusing on technological solutions alone [2]. A systematic review of human factors in cybersecurity within healthcare settings reached a similar conclusion, identifying awareness programmes and cyber hygiene practices as essential countermeasures alongside technical defences [3]. The implication is consistent across these studies. If the human element is the most frequently exploited vulnerability, then educating and training users is not optional but necessary.

This thesis investigates whether a targeted cybersecurity awareness workshop can improve threat recognition skills among Pakistani computer science students. It does so through a controlled experimental design, comparing the performance of students who receive the training against a control group who do not. The study addresses two dimensions of the problem. The first is the importance of such training, established through baseline measurement of existing awareness. The second is its effectiveness, evaluated through pre-test, post-test, and follow-up comparisons. These two terms are used operationally throughout the thesis. Importance refers to a measurable awareness gap in the target population, and effectiveness refers to workshop-produced gains that survive the two-week follow-up window.

1.1 Motivation

Pakistan's digital landscape has expanded at a pace its cybersecurity infrastructure has not matched. The country had approximately 116 million internet users at the start of 2025, with approximately 45.7% of the population using the internet [4]. Digital financial services such as Easypaisa and JazzCash have brought millions of users into online financial transactions for the first time, often without prior exposure to basic security practices. This rapid adoption has coincided with a sharp rise in cyber-enabled crime, and the available evidence converges on the same conclusion across multiple sources.

The Federal Investigation Agency recorded a 70% increase in cybercrime complaints between 2022 and 2024 [5]. An article reported blocking over 16 million cyberattacks targeting Pakistani users during 2023, a 17% rise from the previous year [6]. Reporting on FIA data, The News documented an 83% increase in financial fraud through social media platforms between 2018 and 2021 [7]. The same paper reports that digital banking fraud alone increased by 62% in 2023 [5]. A single data breach at the Bank of Punjab resulted in PKR 2.3 billion in unauthorised transactions [5]. The pattern points to a rapidly expanding attack surface accompanied by insufficient user preparedness.

Social engineering is a particularly effective vector in this environment. Kausar and Laghari analysed social engineering attacks across Pakistan and identified several conditions that attackers routinely exploit [8]. General awareness of cyber threats remains low. Digital literacy is limited. A cultural disposition toward trust further reduces user vigilance. Phishing scams targeting bank customers and mobile payment users are widespread in the country. Voice phishing attacks impersonating telecom providers and government officials have also become commonplace [8]. These attacks succeed not because of technical sophistication but because they exploit human psychology.

Empirical evidence suggests that students are not protected by their proximity to technology. Khan et al. surveyed 758 students across Pakistani higher education institutions using a multi-stage stratified face-to-face survey. Among their participants, 70% never changed their passwords and 30 to 40% opened links without verifying the destination. On smartphones, 75% connected to free Wi-Fi networks. Security practices differed significantly by digital divide variables, with small-to-medium effects for device securement ($V = 0.22$) and software updating ($V = 0.22$) [9]. By Cohen's benchmarks

these represent small-to-medium effects, indicating a consistent but modest association between internet access frequency and security behaviour.

A study of phishing susceptibility among 164 Pakistani respondents found that authority and urgency cues significantly increased the likelihood of victimisation [10].A survey of 256 engineering students in Karachi measured cybersecurity awareness across three dimensions using a 33-item questionnaire [11]. The mean knowledge score was 2.17 out of 5.00 and the mean practice score was 2.89 out of 5.00, both indicating low awareness. Students with advanced computer skills scored significantly higher on cybersecurity practice than those with basic skills ($d = 1.80$).

The problem is reinforced rather than corrected by the academic environment that prepares these students. Catota et al. conducted 28 semi-structured interviews with educational leaders at 13 universities in Ecuador, a comparable developing-country context. Through qualitative thematic analysis of these interviews, they found that only four of the thirteen universities expressed any confidence in their ability to prepare students. Among the surveyed academic departments, 30% offered no security courses, 50% offered one, and 20% offered two. The barriers they identified included a shortage of cybersecurity specialists, policy restrictions that prevented hiring industry practitioners as instructors, and limited laboratory resources [12]. Crumpler and Lewis, reporting for the Center for Strategic and International Studies, documented a similar disconnect on a global scale. A CSIS survey of IT employers found that only 23% believed cybersecurity education programmes fully prepared students for the industry. An ISACA study cited in the same report found that 61% of organisations considered fewer than half of all applicants for open cybersecurity positions to be qualified. The authors identified an over-emphasis on theory and book learning as a recurring complaint against current curricula [13].

The motivation for this research lies in the gap between the scale of the problem and the absence of tested solutions in the Pakistani context. Existing studies confirm the vulnerability of Pakistani students and recommend training, but few have moved beyond recommendation to design and empirically evaluate an intervention. This study examines whether a single focused workshop can produce measurable improvement among Pakistani CS students. The findings may inform future decisions about integrating practical cybersecurity content into undergraduate curricula.

1.2 Research problem

The core problem is the absence of empirical evidence on the effectiveness of cybersecurity training interventions for university students in Pakistan. The existing research on cybersecurity awareness among Pakistani students has been predominantly survey-based, measuring awareness levels without testing interventions.

A similar pattern exists in other developing countries that share comparable characteristics, including rapid digitalisation, limited cybersecurity infrastructure, and constrained higher education resources. Abdulla et al. surveyed social engineering awareness among university students and lecturers in Iraq and concluded that structured training programmes were needed [14]. Chandarman and van Niekerk reached similar conclusions studying students at a South African private tertiary institution [15]. Kausar and Laghari explicitly recommended cybersecurity workshops for Pakistani students [8]. In each of these cases, the research identifies the need but does not address it experimentally.

The gap in the literature is the absence of experimental evaluation of training interventions in these contexts. Without experimental validation, recommendations for awareness training remain speculative. They cannot inform curriculum design or institutional policy with any confidence. This thesis aims to move beyond description by designing and delivering a training intervention. Its impact is then measured through a controlled experiment with repeated measures.

1.3 Research questions and hypothesis

The study is guided by one main research question and two sub-questions.

MRQ: To what extent can a targeted cybersecurity workshop bridge the practical knowledge gap among early-year Pakistani university students?

SRQ1: What is the current baseline level of awareness regarding human-centric security threats among the selected group of students?

SRQ2: How significant is the improvement in threat recognition skills after the workshop, compared to a control group that received no security-related training?

The hypothesis is that students who participate in the workshop will show statistically significant improvement in threat recognition scores. This improvement should be evident both against their own baseline and against the control group's performance. Positive findings, if observed, would need to be interpreted in light of the study design, sample size, and intervention format. They may provide preliminary evidence for incorporating practical cybersecurity content into computer science programmes.

1.4 Scope and limitations

The study focuses on human-centric security threats, specifically phishing and social engineering. It also addresses digital disinformation as a growing concern in the Pakistani context. Technical domains such as network security and cryptography fall outside its scope. Secure coding practices are similarly excluded. The participant sample, intervention design, and data collection procedures are described in Chapter 3.

Several constraints apply. The sample is drawn from universities in a single Pakistani city, which limits generalisability. The study measures short-term and medium-term effects through a post-test and a two-week follow-up, but cannot assess long-term retention. Self-report items in the questionnaire carry the risk of social desirability bias, particularly for behavioural measures administered immediately after the intervention. These limitations are discussed further in Chapter 5.

1.5 Contribution

This work provides an empirically measured baseline of cybersecurity awareness among Pakistani computer science students. It designs and evaluates a targeted workshop using a controlled experimental design with three measurement points. It adapts the Human Aspects of Information Security Questionnaire developed by Parsons et al. for use in this context [2]. Beyond its immediate findings, the study offers a model that could potentially be replicated across similar contexts. If a single workshop can demonstrate measurable gains, it provides early evidence that comparable interventions could be scaled across institutions.

1.6 Thesis structure

Chapter 2 reviews relevant literature on the human factor in cybersecurity. It covers social engineering and disinformation as threat categories, and examines approaches to awareness training and measurement. Chapter 3 describes the research design, including the instruments and the intervention. It also outlines the statistical analysis plan. Chapter 4 presents the results. Chapter 5 discusses the findings and their limitations. Chapter 6 concludes the thesis.

2 Literature review

This chapter reviews the literature relevant to cybersecurity awareness training for university students, with particular attention to the Pakistani and broader developing-country context. Section 2.1 describes the systematic review methodology used to assemble the evidence base. Sections 2.2 to 2.5 synthesise findings on the human factor in cybersecurity, social engineering and phishing, disinformation and media literacy, and awareness training approaches. Section 2.6 examines measurement instruments and their cultural validity. Section 2.7 discusses the Pakistani policy and empirical landscape. Section 2.8 synthesises the findings and identifies the four-component research gap that this thesis addresses.

The literature on cybersecurity awareness and the human factor in information security is extensive and methodologically diverse. To navigate this body of work in a transparent and reproducible way, this chapter follows a structured literature review approach. The approach is based on the systematic literature review guidelines for software engineering published by Kitchenham and Charters [16]. The full formality of a Cochrane-style review is not pursued. Such reviews require independent dual-coding and formal meta-analysis, alongside time commitments that exceed what a single-researcher master's thesis can sustain. Kitchenham notes that single-researcher reviews benefit from a "lite" version of the process, which retains documented search strategies and explicit inclusion criteria together with a quality assessment step [16]. This is the approach followed here.

The Kitchenham guidelines describe three phases: planning, conducting, and reporting. Within conducting, the main procedural steps are research identification, study selection, quality assessment, data extraction, and data synthesis. This review retains five of these steps. It formulates explicit research questions and documents search strings with database choices. Predefined inclusion and exclusion criteria guide study selection. Each study is scored against DARE-style quality questions. The selection flow is reported using PRISMA conventions [17]. Three steps from the full Kitchenham protocol are omitted. Independent dual-coding, in which a second reviewer screens the same records, is dropped because only one researcher conducted this review. Formal meta-analysis is omitted because the included studies differ too widely in design and outcome measures to support statistical pooling. Heterogeneity testing is not applicable for the same reason.

The review is reported using the PRISMA flow convention for documenting study selection [17]. This convention is widely adopted across disciplines and provides a recognisable visual audit trail through the four phases of identification, screening, eligibility, and inclusion. Section 2.1 sets out the methodology of the review itself. Sections 2.2 to 2.7 then synthesise the included literature thematically, organised around the research questions defined below. Section 2.8 closes the chapter with a synthesis of the gap that motivates this thesis.

2.1 Review methodology

2.1.1 Research questions

Six research questions guide the review. They are distinct from the research questions of the thesis itself, which were stated in Chapter 1. The thesis questions ask whether a workshop intervention works for Pakistani computer science students. The review questions ask what existing literature already establishes, so that the thesis can be positioned against the current state of knowledge. Table 2.1 lists the six review questions and their mapping to chapter subsections.

Table 2.1: SLR research questions and mapping to chapter subsections

ID	Research question	Section
SLR-RQ1	How is the human factor framed as a cybersecurity vulnerability in current literature?	2.2
SLR-RQ2	What user-targeted threats are most prevalent and how are they characterised in the literature?	2.3, 2.4
SLR-RQ3	What approaches to awareness training have been evaluated, and what does the evidence say about their effectiveness?	2.5
SLR-RQ4	What measurement instruments exist for cybersecurity awareness based on the KAB model, and how well are they validated?	2.6
SLR-RQ5	What is the state of awareness research in Pakistan and comparable developing-country contexts?	2.7

SLR-RQ6	What gap exists between observational awareness studies and tested intervention studies?	2.8

2.1.2 Search strategy and databases

Six databases were searched, covering the main publication venues relevant to this thesis. The choice of databases balanced disciplinary coverage with the practical accessibility constraints of a single researcher. IEEE Xplore and the ACM Digital Library are the primary venues for computing and security conferences. ScienceDirect and Springer Link index the journals where the most cited human-factors and awareness studies appear. Scopus complements them with broader cross-disciplinary abstract coverage. Google Scholar was used as a complementary resource for forward and backward citation tracking. Table 2.2 summarises the databases and the rationale for each.

Table 2.2: Databases searched and rationale for inclusion

Database	**Reason for inclusion**
IEEE Xplore	Primary venue for cybersecurity and computing conferences, including ICAIC and IEEE S&P
ACM Digital Library	Proceedings of SIGCSE and ITiCSE conferences, alongside other security-education and HCI venues
ScienceDirect (Elsevier)	Computers & Security, Information & Computer Security, and other core human-factors journals
Springer Link	Security Journal, Education and Information Technologies, International Journal of Information Security
Scopus	Largest multidisciplinary abstract database, used for cross-database coverage
Google Scholar	Complementary resource for forward and backward snowballing and citation tracking

Search strings were constructed from PICOC keywords following the Kitchenham guidelines [16]. The PICOC framework specifies the population, intervention, comparison, outcome and context elements of the review question. Synonyms were combined within each element using the Boolean OR operator. Elements were then linked using the Boolean AND. The base search string used was the following:

```
("cybersecurity awareness" OR "information security
awareness" OR
 "security training" OR "phishing training" OR "security
education")
AND
("student" OR "university" OR "higher education" OR
"undergraduate")
AND
("intervention" OR "workshop" OR "experiment" OR
"evaluation" OR
 "knowledge" OR "behaviour")
```

A second, narrower string targeted the Pakistani and developing-country context:

```
("cybersecurity" OR "information security")
AND ("Pakistan" OR "developing country" OR "South Asia")
AND ("awareness" OR "education" OR "training")
```

The base string was adapted to the syntax of each database. Backward and forward snowballing from the seed papers of Parsons et al. [2], Catota et al. [12] and Khan et al. [9] supplemented the database results. Snowballing identified an additional 24 records that were not retrieved by the original queries.

2.1.3 Inclusion and exclusion criteria

Inclusion and exclusion criteria were defined before the searches were executed, in line with the Kitchenham guidance that selection criteria should be set in advance to reduce researcher bias [16]. A study was included if all of the following conditions held. It had to be published between 2014 and 2026, the period beginning with the release of the validated Human Aspects of Information Security Questionnaire by Parsons et al. [2]. It had to be peer-reviewed and written in English. It had to report empirical data on cybersecurity awareness, training, or human factors. The study population had to consist of either students or employees, rather than only specialist practitioners. General internet users were also accepted as a valid population.

A study was excluded under the following conditions. Pure technical security work with no human-factor dimension was outside scope. Opinion pieces and editorials lacking empirical content were excluded. Duplicate publications of the same dataset were resolved by retaining the most complete version. Studies whose full text could not be retrieved were also excluded. Foundational references published before 2014 were

retained only when they were cited by multiple newer papers as theoretical anchors. Table 2.3 summarises these criteria

Table 2.3: Inclusion and exclusion criteria for the literature review

Criterion	Inclusion	Exclusion
Publication date	2014 to 2026	Before 2014, except foundational works cited by multiple newer papers
Language	English	Other languages
Source type	Peer-reviewed journal article, conference paper, or thesis from a recognised institution	Opinion pieces, editorials, blog posts, white papers
Topic relevance	Empirical work on cybersecurity awareness, training, or human factors	Pure technical security work with no human-factor dimension
Population	Students, employees, or general internet users	Specialist security practitioners only
Availability	Full text retrievable	Full text inaccessible

2.1.4 Quality assessment

Each study that passed the eligibility screen was scored against four quality questions adapted from the CRD Database of Abstracts of Reviews of Effects (DARE) criteria. These criteria are recommended by Kitchenham and Charters as appropriate for a single-researcher SLR [16]. The four questions appear in Table 2.4. Each question received one of three scores. A score of 1.0 indicated a yes answer. A score of 0.5 indicated a partial answer. A score of 0.0 indicated a no answer. The maximum possible total was 4.0. Studies scoring below 2.0 were excluded from the synthesis on the grounds that the methodological quality did not support reliable conclusions.

The threshold of 2.0 is deliberately permissive. A stricter threshold would risk excluding valuable qualitative work and case studies that contribute to understanding without

meeting strict experimental standards. The trade-off was accepted as appropriate for a master's thesis review, where breadth of evidence matters as much as depth.

Table 2.4: DARE-style quality assessment questions

ID	Question
Q1	Are the research questions or aims of the study explicitly stated?
Q2	Is the study design and the sample population adequately described?
Q3	Are the data collection and analysis methods appropriate to the research question?
Q4	Are the findings clearly reported and the limitations acknowledged?

2.1.5 Study selection results

The combined search retrieved approximately 596 records from the six databases. Snowballing added 24 further records. After deduplication, 458 unique records remained. Title and abstract screening excluded 390 records that fell clearly outside the review scope. The full text of 68 records was assessed for eligibility. A further 29 were excluded at this stage, most commonly for a peripheral human-factor dimension or an insufficient methodology description. Quality scoring excluded 6 records that fell below the 2.0 threshold. The final pool contains 33 peer-reviewed studies. Figure 2.1 documents the complete selection flow.

Six additional references are cited in the chapter but were not part of the quality-scored pool. These include methodology and reporting guidelines [16][17] alongside an industry data report [1]. Contextual sources such as news articles and web statistics [4][6][7] complete the set. They provide supporting data but are distinguished from the 33 studies that passed the systematic selection process.

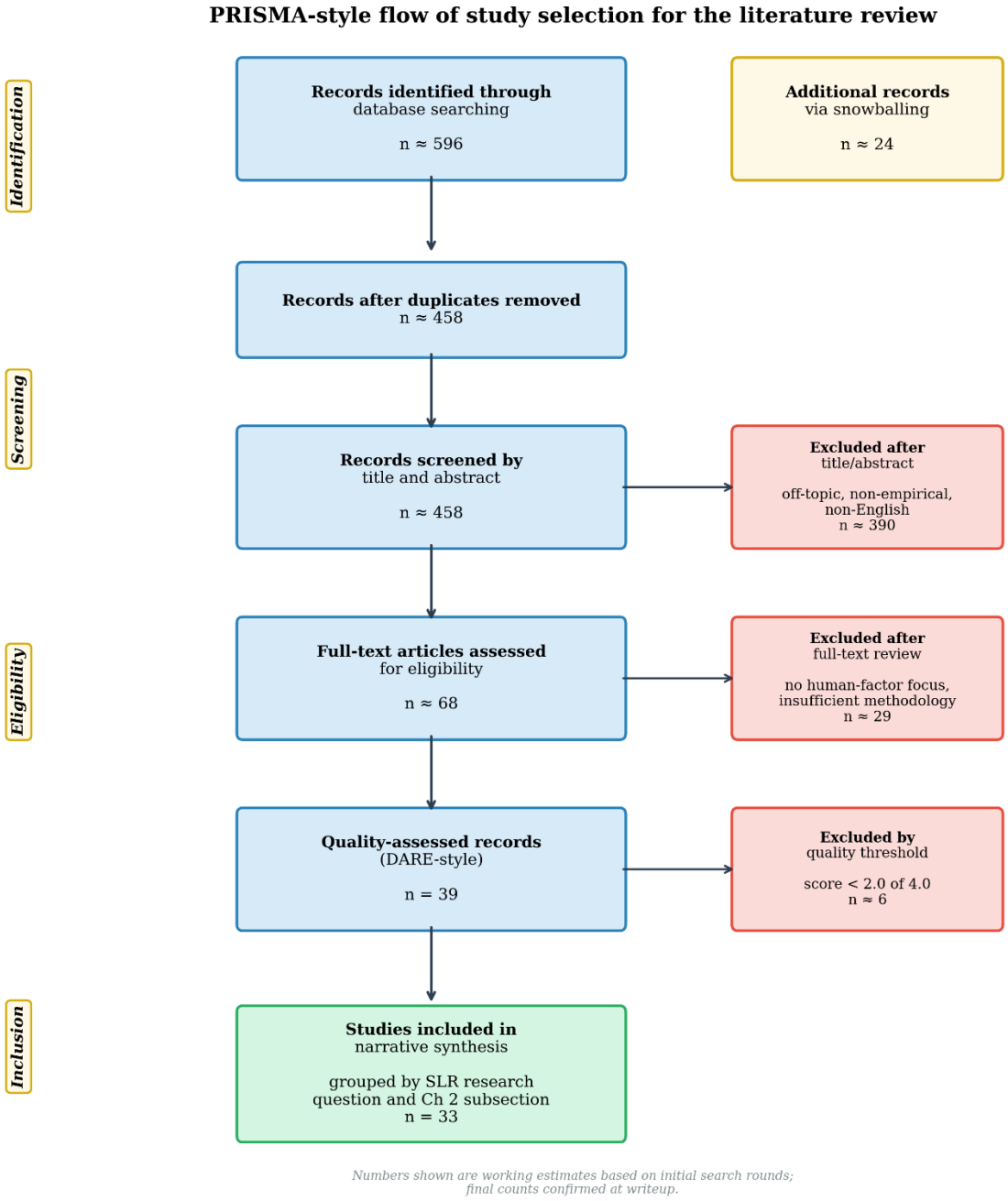


Figure 2.1: PRISMA-style flow of study selection for the literature review

The 33 included studies cover the period from 2014 to 2026. The largest concentration falls between 2020 and 2025. The geographic distribution is weighted toward American and European samples, a pattern documented by Rahman et al. [19]. Studies set in South

Asia, and in Pakistan specifically, remain comparatively rare. This imbalance is itself part of the gap that this thesis addresses.

Table 2.5 summarises the study types in the final pool. No study used a randomised controlled trial design with a Pakistani student sample.

Table 2.5. Study types in the final review pool.

Study type	Count	References
Systematic, scoping reviews, and meta-analyses	11	[3][19][24][25][27][30][31][32][33][37][38]
Cross-sectional surveys	6	[9][10][11][14][15][39]
Experimental or intervention studies	7	[21][22][23][26][28][29][36]
Instrument development and validation	3	[2][34][35]
Qualitative, conceptual, and policy analyses	6	[1][5][8][12][18][20]
Total	33	

2.2 Cybersecurity and the human factor

The Verizon 2024 Data Breach Investigations Report found that 68% of confirmed breaches involved a non-malicious human element [1]. This proportion has remained stable across successive annual reports. Continued investment in technical controls has not substantially shifted it. Human behaviour is now the most reliable predictor of whether an organisation will be breached.

Multiple peer-reviewed reviews confirm this pattern through independent methods. Nifakos et al. reviewed human factors in healthcare cybersecurity and identified user awareness programmes as essential countermeasures alongside technical defences [3]. Pollini et al. combined the HAIS-Q instrument [2] with focus groups across three Italian healthcare organisations [20]. They found that workers often fail to follow security measures not because they lack awareness but because the organisation's security culture works against compliant behaviour [20]. This finding shifts the diagnostic frame from individual knowledge gaps to the sociotechnical system in which users operate.

The dominant measurement framework in this field is the Knowledge-Attitude-Behaviour (KAB) model. The model holds that knowledge of security policies leads to improved

attitudes, which in turn drive more secure behaviour. It originated in health psychology and was adapted for cybersecurity by Parsons et al. through the Human Aspects of Information Security Questionnaire (HAIS-Q) [2]. The instrument and its validation are examined in Section 2.6.

Zimmermann and Renaud challenged the assumption that users are primarily a vulnerability to be controlled [18]. Their analysis of cybersecurity discourse found that defenders consistently frame the human as a problem to be constrained through policy and training. They argue this stance treats well-intentioned users as potential threats rather than recruiting them into the security function [18]. Their counter-proposal frames humans as contributors to security. This is not incompatible with the measurement approach. Pollini's framework integrates individual-level instruments like the HAIS-Q with organisational factors that shape behaviour [20].

The geographic distribution of this research reflects a gap. Rahman et al. reviewed human-factor studies from top-tier security conferences over six years [19]. They found that samples are heavily concentrated in American and European institutions. The cultural and contextual factors that shape user behaviour in other regions are underrepresented in the evidence base. Section 2.7 examines this gap with quantitative evidence for the Pakistani context.

2.3 Social engineering and phishing

Social engineering exploits psychological mechanisms rather than technical vulnerabilities. Cialdini identified six principles of influence through which compliance can be induced [40]. Among these, authority and scarcity are most commonly exploited in phishing attacks. An email impersonating a university IT department exploits authority. A message warning that an account will be locked within 24 hours exploits scarcity. Social proof operates when a link appears to come from a trusted contact. The attacker does not need to defeat any technical defence. The attacker only needs to trigger one of these compliance pathways before the target pauses to verify.

Empirical studies confirm that these mechanisms reliably predict susceptibility. Sommestad and Karlzén manipulated authority and urgency cues across multiple phishing variants in a repeated-measures field experiment [22]. They sent 2,294 simulated phishing

emails using 47 distinct scam templates to 102 employees over 16 months. Overall click rates were 6.4% for links and 3.0% for code execution (d = 0.35), but scam-level susceptibility ranged from 0% to 19%. The scam itself explained only 6% of the variance and a medium effect in click susceptibility (R^2 = 0.06, d = 0.51). Personalisation added a modest effect (OR = 1.42, d = 0.19) and influence techniques showed no significant relationship to susceptibility. The same participant could fall for one email and detect a similar one the following week. This instability suggests that contextual factors such as workload and attention interact with the psychological mechanisms, making susceptibility difficult to attribute to stable traits alone.

Sarno and Black found that users default to accepting messages as truthful, consistent with Truth-Default Theory [21]. Self-reported confidence did not reliably predict actual detection performance. Yang et al. reached a complementary conclusion through a different method [23]. They trained machine learning classifiers on individual-level data from 1,105 participants and achieved 89% accuracy in predicting susceptibility from demographic and behavioural features. Both studies converge on the same finding. Knowledge of phishing in the abstract is a weak predictor of detection ability in specific cases.

The most consequential recent evaluation is the Ho et al. 8-month RCT at a US healthcare organisation, tracking 19,500 employees through ten simulated phishing campaigns [26]. Annual cybersecurity training showed no significant relationship with subsequent phishing failure rates. Embedded exercises produced a statistically significant but practically small 1.7 percentage point reduction in failure rates compared to the control group (OR = 0.905, 95% CI: 0.863 to 0.950, d = -0.06, $p < 0.001$) [26]. Annual awareness training showed no significant effect or association with failure rates ($p = 0.06$). Among users who completed interactive training, the effect was slightly larger (OR = 0.809, 19% lower failure, d = -0.12), but static training was associated with worse outcomes (OR = 1.185 per additional session completed, d = 0.09). Engagement was minimal: approximately 75% of users spent less than one minute on the training material, and roughly one third closed the page immediately. The result does not demonstrate that training per se is ineffective. It demonstrates that generic annual training with low engagement does not produce measurable change at population level.

The Ho et al. finding narrows the field rather than closing it. The study evaluated generic content delivered annually with minimal engagement. It did not test targeted short-form interventions delivered to specific user populations. The intervention examined in this thesis is precisely that type: a focused 90-minute workshop targeting a defined student group. Bukhsh et al. provide complementary evidence for the Pakistani context [10]. Their survey of 164 internet users found that phishing susceptibility was widespread and not strongly correlated with self-reported confidence. This pattern is consistent with the Sarno and Black finding, suggesting that the cognitive vulnerabilities are not culture-specific.

Phishing is not the only threat that exploits these psychological mechanisms. Disinformation operates through similar compliance pathways at population scale. The next section examines it as a related category of social engineering.

2.4 Disinformation and media literacy

Disinformation is the deliberate spread of false content with the intent to deceive. It operates through the same psychological mechanisms as phishing but at population scale. Where phishing targets individuals through email or messaging, disinformation targets groups through social media feeds. Both exploit Cialdini's compliance principles [40]. A fabricated news article attributed to a credible source exploits authority. A viral post shared by many contacts exploits social proof. The shared mechanism means that defences effective against one form of social engineering may transfer to the other.

For Pakistani university students, this threat is immediate rather than abstract. The Digital 2025 Pakistan report documented heavy social media use across the population [4]. Much of that time involves news and opinion content, and informal communication often mixes the two without clear labels.

The empirical evidence on media literacy interventions has reached considerable maturity. Huang et al. conducted a meta-analysis of 49 experimental studies with a combined sample of 81,155 participants [27]. The overall effect size was $d = 0.60$. In practical terms, the average trained participant outperformed approximately 73% of untrained participants in identifying misinformation. The effect was strongest for discernment ($d = 0.76$, outperforming approximately 78% of controls) and for reducing

sharing behaviour (d = 1.04, outperforming approximately 85%). The weakest effect was for changing belief in misinformation already encountered (d = 0.27, outperforming approximately 61%). By Cohen's conventions, the overall effect is medium-to-large. In practical terms, the sharing reduction (d = 1.04) means that a trained participant is approximately 85% likely to share less misinformation than an untrained one. The discernment gain (d = 0.76) means approximately 78% of trained participants outperform the average untrained participant.

Guess et al. tested a digital media literacy intervention in both the United States and India [28]. The intervention improved discernment in both countries. The effect held across differences in education and political orientation. India's proximity to Pakistan in terms of media environment and digital adoption patterns makes this cross-country replication particularly relevant.

Adjin-Tettey used a control-group experimental design comparable to the one used in this thesis [29]. Trained participants in her sample of 187 identified inauthenticity more accurately than controls. Some trained participants nevertheless failed to detect false content. She attributes this partial success to the single-session format, which offered no opportunity to consolidate material. This limitation applies equally to the present thesis. No equivalent intervention study exists for the Pakistani context at the time of writing. The gap is consistent with the broader underrepresentation identified by Rahman et al. [19].

2.5 Awareness training approaches

Training is the dominant intervention strategy for the human factor in cybersecurity. Two broad categories exist: traditional formats based on lectures and e-learning, and interactive formats using workshops and simulations. The evidence on their relative effectiveness is reviewed below.

2.5.1 Traditional and interactive approaches

Section 2.3 presented two key findings on training effectiveness. Ho et al. found no significant effect from annual e-learning training in a 19,500-employee RCT [26].

Jampen et al. found that embedded training with immediate feedback outperforms passive instruction [24]. This subsection broadens the picture beyond those individual studies.

Khando et al.conducted the most comprehensive SLR of enterprise information security awareness methods [30]. They identified gamification as one of the most effective approaches. They also noted that the evidence base for any single method is fragmented, with studies differing widely in sample size and outcome measures.

Abdulla et al. surveyed 1,779 students and staff at a large Iraqi university and found that 57% would click a fake university link and 52% could not identify spam email [14]. One-way ANOVA revealed significant differences by age for spear phishing susceptibility ($F = 16.76$, $d = 0.28$, $p < 0.001$). The authors recommended structured training but did not test an intervention. The Iraqi context shares institutional and cultural patterns with Pakistan that distinguish both from the American and European samples in the wider literature. The pattern across these sources is consistent. Active and contextual training tends to outperform passive and generic training. The durability of the effect remains uncertain because few studies include follow-up measurements beyond the immediate post-test. Of the seven experimental studies in this review, only Ho et al. [26] tracked outcomes beyond one month.

2.5.2 Gamification

Gamification applies game design elements in non-game contexts to increase engagement and learning. The psychological basis for its effectiveness draws on Self-Determination Theory. Sailer et al. demonstrated that specific game elements satisfy core psychological needs identified by that theory [41]. Feedback on performance addresses the need for competence. Social features such as leaderboards address relatedness. Player choice over actions addresses autonomy. When these needs are met, intrinsic motivation increases. Immediate feedback after each question also reinforces correct behaviour through operant conditioning, a mechanism distinct from the delayed feedback typical of annual training.

Three recent systematic reviews examined gamification in ISA programmes. Pahlavanpour and Gao mapped 69 papers and found that evaluation and validation research dominated the field, with content gamification (serious games) used twice as often as structural gamification [31]. Only 22% of studies employed adaptive gamification and only three used AI-based adaptation, indicating clear gaps. Gwenhure

and Rahayu reviewed five empirical studies targeting non-IT professionals and reported short-term gains including a 51% improvement in security awareness scores in one study and a 5% improvement in password selection over two months in another [32]. However, no study measured engagement beyond three months. Weeratunge et al. examined 230 studies and found that 76% of game-based methods had been applied in teaching, compared with 53% of non-game methods [33]. None of the three reviews reported a pooled effect size, as the underlying studies used incompatible outcome measures.

The workshop intervention in this thesis incorporates gamified elements within its slide-based delivery. Interactive scenario exercises embedded in the presentation provide immediate feedback after each response, addressing the competence need identified by Sailer et al. [41]. This approach integrates game-like elements into a traditional workshop format rather than requiring a standalone platform. The trade-off is practical: embedding exercises within slides avoids dependence on stable internet connectivity, a relevant constraint in Pakistani university settings. Table 2.6 summarises the four main delivery formats and their reported effect sizes.

Table 2.6. Comparative summary of training approaches with available effect evidence.

Approach	**Typical format**	**Effect evidence**	**Key limitations**
Annual mandatory training	E-learning module or video lecture	No significant effect on phishing failure (Ho et al. 2025, N = 19,500 RCT): OR = 0.998 per 30 days, p = 0.06	75% of users spent < 1 min on material; engagement minimal
Embedded phishing simulation	Simulated phishing email with immediate feedback	OR = 0.905, p < 0.001, 1.7 pp absolute reduction, d = -0.06 overall (Ho et al. 2025). Static training worsened outcomes (OR = 1.185, d= 0.09); interactive reduced failure by 19% among completers	Small absolute effect; only failing users receive training; 56% of users fail at least once over 8 months
Interactive workshop	In-person or virtual session with hands-on exercises	Robbins and Robbins (2025, N = 201, HAIS-Q): Knowledge Z = −7.31, d = 1.20; Attitude Z = −3.78, d = 0.55; Behaviour Z = −5.20, d = 0.79 (all p < 0.001)	Hard to scale; high cost per learner; short-term measurement only (4–6 weeks)
Gamified learning	Quiz games, leaderboards, interactive scenarios	Short-term gains: +51% awareness (Abu-Amara and Tamimi 2021); +5% password improvement over 2 months (Qusa et.al 2021).	Long-term effects unstudied;only 22% of studies use adaptive gamification [31].

The effectiveness of any training approach depends on how outcomes are measured. The next section examines the instruments available for that purpose.

2.6 Measurement instruments and cultural validity

A training programme can only be evaluated if its outcomes are measured reliably. The KAB model, introduced in Section 2.2 and illustrated in figure 2.2, provides the theoretical framework. The question here is whether the instruments built on that framework are valid for the context in which this thesis operates.

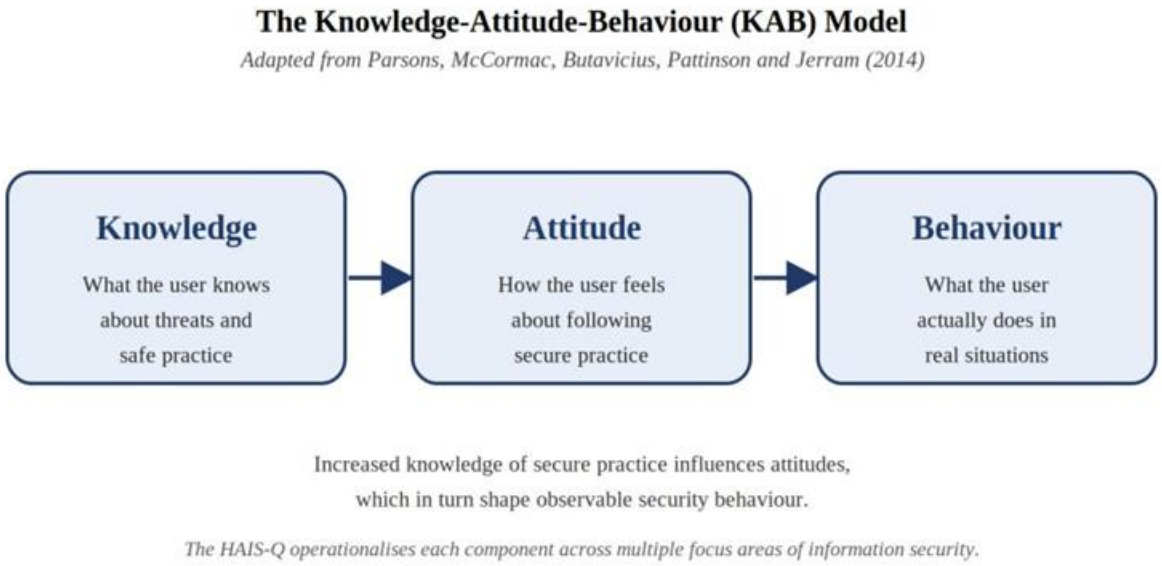


Figure 2.2. The Knowledge-Attitude-Behaviour (KAB) model that underpins the HAIS-Q instrument.

The Human Aspects of Information Security Questionnaire (HAIS-Q) was developed by Parsons et al. (2014) using a sample of approximately 500 Australian office workers [2]. It operationalises the KAB model across seven workplace security focus areas. Parsons et al. (2017) established convergent validity by correlating HAIS-Q scores with performance on a separate phishing identification task [34]. Higher-scoring participants were more likely to detect phishing emails correctly. McCormac et al. confirmed test-retest reliability over a four-week interval with 197 Australian participants [35]. The instrument structure and item details are described in Chapter 3.

Rohan et al. conducted a systematic review of 24 cybersecurity awareness scales against a 15-criterion psychometric framework [37]. More than half of the reviewed instruments

failed to meet even half of the methodological criteria. The HAIS-Q was among the better-validated instruments, primarily because the Parsons and McCormac studies cover reliability and convergent validity criteria that most other scales omit.

The closest published comparator to the design used in this thesis is the Robbins and Robbins case study [36]. They administered the HAIS-Q before and after an interactive training session to 201 K-12 staff members in the United States. Wilcoxon signed-rank tests showed significant improvements for knowledge ($Z = -7.31$, $p < 0.001$, $d = 1.20$), attitude ($Z = -3.78$, $p < 0.001$, $d = 0.55$), and behaviour ($Z = -5.20$, $p < 0.001$, $d = 0.79$). The knowledge effect was the largest, consistent with the expectation that declarative learning responds most directly to instruction. This confirms that the instrument is sensitive enough to detect intervention effects in a single-session format.

A limitation of the existing evidence is its geographic concentration. All published HAIS-Q development and validation studies used Australian samples [2][34][35]. The Robbins study extends this to the United States, a culturally similar context [36]. No published validation exists for South Asian populations. Workplace security norms differ between Australian organisations and Pakistani universities. Technology usage patterns also differ: most Pakistani students access the internet primarily through smartphones rather than desktop computers [4]. Language is a further consideration, as English is a second language for most Pakistani participants.

The Ahamed et al. study from Bangladesh provides partial evidence that the KAB framework transfers to the South Asian context [39]. Their survey of 430 university students found that cybersecurity attitude mediated the relationship between knowledge and overall awareness, consistent with the original Parsons et al. model. However, they used their own instrument rather than the HAIS-Q. The question of whether the HAIS-Q items themselves perform reliably with Pakistani respondents has not been empirically tested.

This thesis adapts rather than directly replicates the HAIS-Q for its measurement instrument. The adaptation choices, including item selection and contextual modifications for the Pakistani student population, are described in Chapter 3.

2.7 Pakistan and the developing-country context

Pakistan's legislative response to cybercrime centres on the Prevention of Electronic Crimes Act (PECA), enacted in August 2016 [55]. PECA criminalises unauthorised access to information systems, cyber fraud, identity theft, and online harassment. It provides for the establishment of investigation agencies and computer emergency response teams. However, PECA is a criminal statute focused on prosecution after offences occur. It contains no provisions for cybersecurity education, awareness training, or preventive programmes targeting end users. The Act was amended in January 2025 with expanded enforcement provisions, but the educational gap remains unaddressed.

At the institutional level, Pakistan's cybersecurity posture has improved substantially. The ITU Global Cybersecurity Index ranked Pakistan 79th in 2020 with a score of 64.88 out of 100. By 2024, the country had risen to Tier-1 status with a score of 96.69, placing it among 46 countries classified as role models [56]. This improvement reflects progress in legal frameworks, organisational structures, and capacity-building initiatives. The Ministry of IT established a National Telecom Security Operations Centre and a National Telecom CERT. However, the GCI measures institutional readiness rather than individual awareness. The disconnect between institutional scores and user-level behaviour is central to the gap this thesis addresses.

The Higher Education Commission (HEC) launched the National Centre for Cyber Security (NCCS) in 2018 as a joint project with the Federal Planning Commission. The NCCS established research and development labs at several Pakistani universities. However, these are research facilities, not curriculum mandates. The HEC 2025 revised Computer Science curriculum lists cybersecurity as an elective specialisation alongside artificial intelligence and data science. It is not a mandatory core component for all CS students. A dedicated BS in Cyber Security programme was announced for implementation from Fall 2026, but it is a separate degree path. General CS students at most Pakistani universities can graduate without taking any cybersecurity course.

The empirical picture at the student level remains unfavourable. Khan et al. found that 70% of 758 surveyed students never changed their passwords and 30 to 40% opened links without verification [9]. Hanif and Shams reported similar gaps among 256 engineering students in Karachi, including weak password practices and low awareness of social

engineering tactics [11]. Kausar and Laghari analysed social engineering conditions across Pakistan and identified low digital literacy, cultural trust, and limited awareness as recurring vulnerabilities [8]. Khan documented a 70% increase in FIA cybercrime complaints between 2022 and 2024 alongside rising digital banking fraud [5]. Bukhsh et al. found that phishing susceptibility among 164 Pakistani internet users was widespread and not correlated with self-reported confidence [10].

The pattern is consistent across these studies. Institutional frameworks have improved, but individual-level awareness has not kept pace. PECA addresses crime after the fact. The GCI captures policy readiness. Neither addresses the practical question of whether targeted training can improve student threat recognition. This is the gap that the present thesis occupies.

2.8 Synthesis and identification of the research gap

This chapter reviewed the literature across six thematic sections (§2.2 to §2.7), covering the human factor, user-targeted threats, training approaches, measurement instruments, and the Pakistani context. The literature is recent: 15 of the 33 studies were published in 2024 or later. Four components of the research gap emerge from this review. Table 2.7 maps each component onto the design choice that addresses it in this thesis.

Table 2.7. The four-part research gap and how the present thesis addresses each component.

Gap component	Status in published literature	How this thesis addresses it
Geographic	Pakistani student population essentially absent from published intervention work	Recruits Pakistani CS students from seven Islamabad universities as both experimental and control participants
Methodological	Most Pakistani and South Asian work is descriptive survey rather than experimental [9][11][14]	Uses pre-test and post-test design with separate trained and untrained groups
Design quality	Few intervention studies include a control group or follow-up assessment [14][36]	Adds an explicit control group and a two-week follow-up beyond the immediate post-test
Measurement	Pakistani-context studies rarely use validated instruments [37]	Uses HAIS-Q-adapted items validated by Parsons and McCormac [2][34][35]

None of these design choices is novel in isolation. Their combination in a Pakistani higher-education setting does not appear to have been published before. The methodological choices that follow from this review are described in Chapter 3, together with the adapted instrument and the workshop materials.

3 Methodology

This chapter describes the research design, participants, instruments, intervention, procedure, analysis plan, and ethical framing of the study. It presents each methodological choice as a deliberate response to the gaps identified in Chapter 2.

3.1 Research design overview

The study adopts a quasi-experimental pretest-posttest design with a non-equivalent control group and a planned two-week follow-up measurement. It compares an experimental group that receives a 90-minute interactive cybersecurity awareness workshop against a control group that receives a matched, content-neutral presentation on digital infrastructure. The design is quasi-experimental rather than fully randomised. Participants were drawn from pre-existing lecture cohorts through a collaborating researcher at Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology. Random assignment at the individual level was not feasible under these recruitment conditions. Group allocation was instead made at the cohort level before any measurement took place. This preserves the between-group comparison while acknowledging the absence of individual randomisation.

The design sits within the pretest-posttest control group family described by Shadish, Cook, and Campbell [42]. This family is appropriate when full randomisation is impractical but a comparable non-treated group can still be measured alongside the treated group. The presence of a pretest allows baseline equivalence to be assessed empirically rather than assumed. The presence of a control group allows the intervention effect to be separated from rival explanations such as maturation and test familiarity. The two-week follow-up extends the design beyond immediate post-test measurement and provides a window onto short-term retention.

Structurally, the study is a 2 × 3 mixed factorial as shown in Figure 3.1. Group (experimental, control) is the between-subjects factor. Time (pre-test, post-test, two-week follow-up) is the within-subjects factor. The central quantity of interest is the Group × Time interaction. This tests whether the two groups change differently across the three

measurement points. A significant interaction in the expected direction would indicate that the workshop produces measurable gains that the placebo presentation does not.

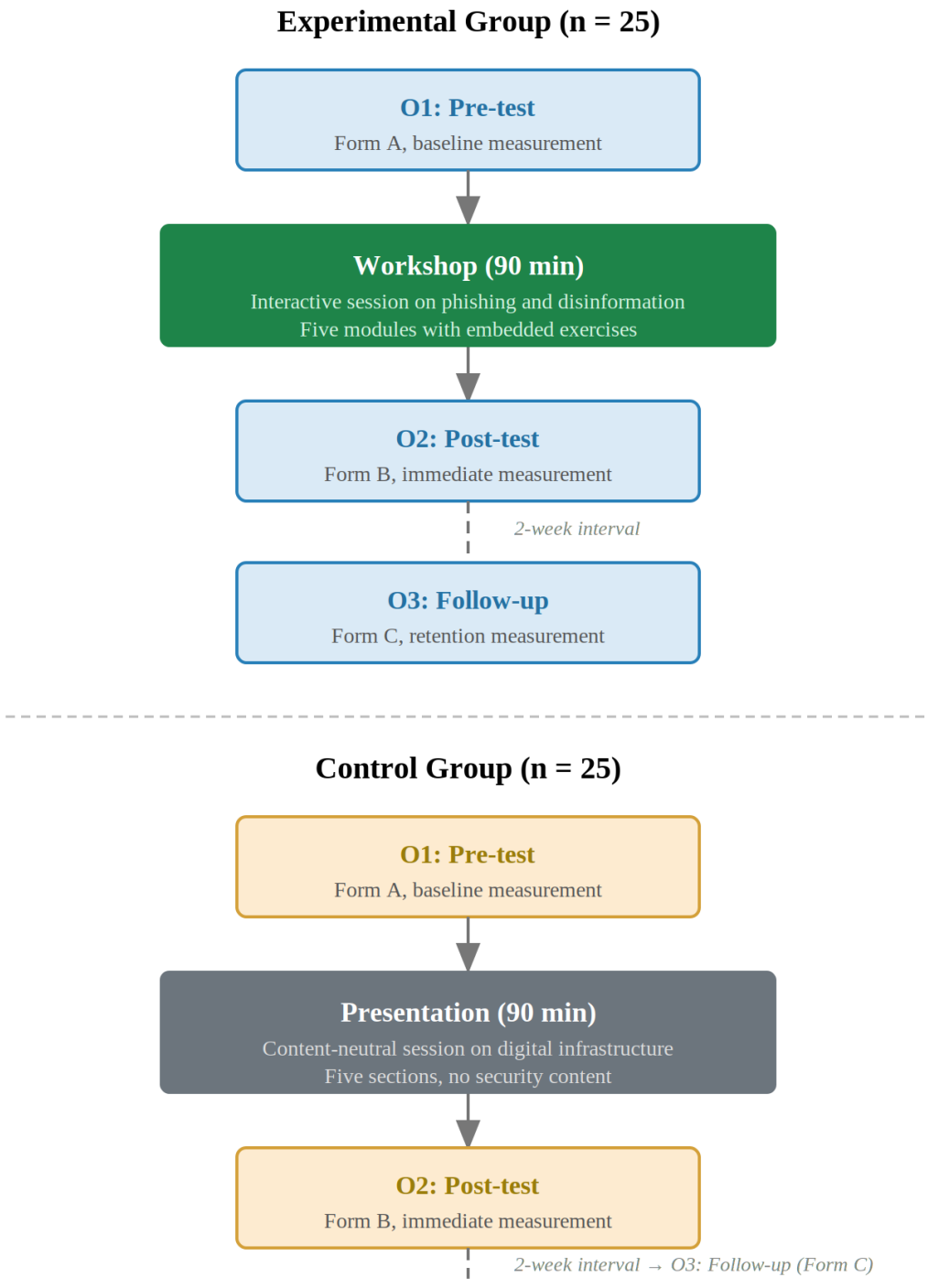


Figure 3.1: Study design. A 2 × 3 mixed factorial with Group as the between-subjects factor and Time as the within-subjects factor. The experimental group receives the workshop intervention between T1 and T2. The control group receives a matched content-neutral presentation over the same interval. Both groups are measured again at T3, two weeks after T2.

Four design decisions follow directly from the gaps synthesised in Section 2.8. First, the control group addresses the absence of comparison conditions in the prior Pakistani literature. It also ensures that any observed post-test change is not attributable to test familiarity or maturation alone. Second, an a-priori power analysis was conducted using G*Power 3.1 [43]. At 80% power with $\alpha = 0.05$, a minimum total sample of 28 is required to detect a medium effect ($f = 0.25$). The achieved sample of 50 exceeds this threshold at every conventional power level up to 0.95. Third, the two-week follow-up extends the evidential window beyond the immediate post-test. It does so without imposing the attrition burden that a longer interval would create. Fourth, the workshop and the control presentation are matched on duration, delivery mode, facilitator, and venue. Only their content differs. This isolation of content as the single manipulated variable is a standard requirement. It allows any observed change to be attributed to the intervention itself rather than to incidental features of the session.

The unit of analysis is the individual participant. Each participant contributes one score per instrument per measurement occasion. Pre-test and post-test occasions are administered within the same session to minimise attrition between baseline and immediate measurement. The follow-up occasion is administered two weeks later through an emailed link, using the same participant identifier that replaced the email address before analysis. The instruments and scoring procedures are described in Section 3.5. The analysis pipeline is described in Section 3.8.

3.2 Research questions and hypotheses

Chapter 1 posed one main research question and two sub-questions. They are restated here in the form that this study operationalises and tests.

> **MRQ:** To what extent does a targeted cybersecurity workshop improve threat recognition among early-year Pakistani computer science students, relative to a matched control condition that receives no security-related training?
>
> **SRQ1:** What is the baseline level of practical threat recognition among the participating students prior to any intervention?
>
> **SRQ2:** How large and how durable is the gain in threat recognition produced by the workshop, measured immediately after delivery and at a two-week follow-up, compared with the control group over the same interval?

Each question corresponds to a specific quantity in the data. SRQ1 corresponds to the group-level mean scores at T1 on Form A, broken down by the knowledge, attitude, behaviour, and scenario subscales defined in Section 3.5. SRQ2 corresponds to the within-group change from T1 to T2 and from T1 to T3 in each group, and to the between-group difference in those changes. The MRQ is answered by the Group by Time interaction term in the mixed ANOVA described in Section 3.8.

The study tests three hypotheses. Each is stated in null and alternative form, with a predicted direction where theory and prior literature support one. The main hypothesis concerns the Group by Time interaction. The two supporting hypotheses concern the directional between-group differences at the post-test and follow-up occasions.

H1 (null): The change in threat recognition scores over time is equal in the experimental and control groups.

H1 (alternative): The experimental group shows a larger gain from T1 to T2, and a larger residual gain at T3, than the control group. The Group by Time interaction is significant.

H2 (null): At T2, mean threat recognition scores in the experimental and control groups are equal.

H2 (alternative): At T2, mean threat recognition scores are higher in the experimental group than in the control group.

H3 (null): At T3, the experimental group scores have returned to the level of the control group.

H3 (alternative): At T3, the experimental group retains a recognition advantage over the control group, although smaller than at T2.

H1 is the principal hypothesis. It addresses whether the workshop produces a pattern of change that the control condition does not. H2 tests the immediate effect. H3 tests short-term retention. A finding that confirms H2 but not H3 would indicate that workshop gains are transient. A finding that confirms both H2 and H3 would support the case for workshop-based training as a durable awareness intervention, subject to the limits of a two-week observation window.

Table 3.1 maps each research question to the instrument component that supplies its data and to the statistical procedure that tests it. The instrument components are described in Section 3.5. The statistical procedures are described in Section 3.8.

Table 3.1: Research questions mapped to instrument components and analysis methods

Question	**Instrument component (see §3.5)**	**Analysis (see §3.8)**
SRQ1	Form A (T1) — KAB and scenario subscales	Descriptive statistics; baseline equivalence test between groups
SRQ2	Forms A, B, C (T1, T2, T3) — full score and subscales	Within-group paired comparisons; between-group independent t-tests at T2 and T3
MRQ	Forms A, B, C (T1, T2, T3) — full score	2 × 3 mixed ANOVA; Group × Time interaction; effect size partial eta squared ($\eta^2 p$)

3.3 Setting and population

The study was conducted in Islamabad, Pakistan. Islamabad hosts a dense concentration of public and private universities within a compact metropolitan area. This concentration allowed participants from several institutions to be reached through a single access point without requiring travel between cities. The workshop and the control presentation were delivered at a neutral venue in the city. Neither session was held on the campus of any participating university.

Access to participants was arranged through a collaborating researcher at Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology. The collaborator drew on existing teaching contact with computer science bachelor's cohorts across several Islamabad institutions. The collaborator is anonymised in this thesis at their request and is referred to throughout as the partner researcher. The partner researcher handled participant invitation and on-site logistics. The author delivered both sessions to keep facilitator effects constant across conditions.

The target population is early-year undergraduate computer science students in Pakistan. First and second year students were selected because they sit at the point in the curriculum where cybersecurity content, where it exists at all, has not yet been formally delivered. This population also allows the benchmark argument set out in Chapter 1 to be tested empirically. Students in this cohort have chosen a technical discipline and use digital tools daily. If they show recognition gaps, the gaps in the wider student population are likely to be at least as large.

Fifty students participated in total. Twenty-five were allocated to the experimental condition and twenty-five to the control condition. The two groups are matched exactly on gender (14 male and 11 female in each). Mean age, estimated from band midpoints, was 20.52 (*SD* = 2.33) in the experimental group and 20.08 (*SD* = 2.63) in the control group. The control group skews slightly younger and slightly more toward second year enrolment. Reported prior exposure to formal cybersecurity training was low in both groups. This supports the assumption that the participants constitute a reasonably homogeneous baseline for the intervention test.

Participants came from seven universities based in Islamabad. Six institutions contributed to the experimental group and seven to the control group. The spread across institutions is intentional. It avoids tying the findings to the curricular quirks or departmental culture of any single university. Table 3.2 summarises the demographic composition of the two groups.

Table 3.2: Participant demographics by group

Characteristic	**Experimental (n=25)**	**Control (n=25)**
University		
Quaid-i-Azam University	8	6
International Islamic University	5	4
Air University	4	5
COMSATS University Islamabad	3	4
NUST University Islamabad	3	3
CUST University	2	1
SZABIST Islamabad	0	2

Characteristic	Experimental (n=25)	Control (n=25)
Year of study		
1st Year	16	12
2nd Year	9	13
Age		
17 or under	2	6
18–20	12	9
21–23	8	7
24 or older	3	3
M (SD)*	20.52 (2.33)	20.08 (2.63)
Gender		
Male	14	14
Female	11	11
Prior cybersecurity training		
Yes	2	4
No	22	21
Maybe / do not remember	1	0
Cybersecurity in current coursework		
Yes	4	3
No	15	16
Maybe / not sure	6	6

** Age was collected in categorical bands. Mean and standard deviation were estimated using band midpoints (17, 19, 22, 25).*

3.4 Sampling and sample size justification

The sampling approach was purposive at the cohort level and exhaustive at the individual level within each cohort. The partner researcher identified first and second year computer science groups to which access was available through existing teaching contact. Every student in those groups was invited. Participation was voluntary and was not linked to course assessment or credit. Refreshments were provided at both sessions. No other incentive was offered. Group assignment was fixed at the cohort level before any measurement took place.

Sample size was determined through an a-priori power analysis using G*Power 3.1 [43]. The analysis targeted the within-between interaction in a repeated-measures ANOVA (two groups, three measurements, $f = 0.25$, $\alpha = 0.05$, $\rho = 0.5$, $\varepsilon = 1$). Cohen's medium-effect convention was used for the effect size [44]. At 80% power, G*Power returned a required total sample of 28 ($\lambda = 10.50$, critical $F(2, 52) = 3.175$, actual power = 0.812).

Table 3.3 reports the output at four power levels as a sensitivity check. The achieved sample of 50 exceeds the requirement at every level, including 95% power (N = 44).

Table 3.3: G*Power output at four power levels (f = 0.25, α = 0.05, groups = 2, measurements = 3, ρ = 0.5, ε = 1)

Power	N required	Per group	λ	Critical F	Actual power	Achieved N
0.80	28	14	10.50	3.175	0.812	50
0.85	32	16	12.00	3.150	0.866	50
0.90	36	18	13.50	3.132	0.906	50
0.95	44	22	16.50	3.105	0.956	50

The medium effect size is a conservative choice. It assumes that an interactive workshop will produce a moderate rather than a large change. The sensitivity check across four power levels confirms that the study is adequately powered under this assumption regardless of which conventional threshold a reader applies. The four G*Power output windows corresponding to Table 3.3 are reproduced in Appendix 5.

3.5 Instrument design

The measurement instrument is a three-form questionnaire adapted from the Human Aspects of Information Security Questionnaire (HAIS-Q). The HAIS-Q was developed by Parsons et al. (2014) at the Defence Science and Technology Group in Australia [2]. Two subsequent validation studies confirmed its convergent validity and test-retest reliability [34][35]. Section 2.6 reviewed the validation evidence and discussed the cultural limitations of applying an Australian-developed instrument in a South Asian context. This section describes the operational structure of the adapted instrument used in this study.

The HAIS-Q operationalises the Knowledge-Attitude-Behaviour (KAB) model. For each security focus area, the model posits that knowledge of a threat or policy leads to an attitude toward the relevant protective behaviour, which in turn predicts actual behaviour. The original instrument covers seven focus areas: password management, email use, internet use, social media use, mobile device use, information handling, and incident reporting. Each focus area contains three sub-areas. Each sub-area is measured by one item from each of the three KAB dimensions. The full original instrument therefore contains 63 items on a five-point Likert scale.

Two adaptations were made for this study. First, items referring to workplace contexts were rewritten for a personal and academic setting. The incident-reporting focus area was replaced with a disinformation-awareness focus area. This change reflects the threat landscape described in Chapter 2. Second, a scenario-based component was added alongside the Likert items. The scenario component tests applied recognition rather than self-reported knowledge. It addresses the concern that self-report alone may overestimate practical skill.

Each form contains two scored sections. Section B holds twelve KAB Likert items. The first eight measure knowledge and attitude on a five-point scale from Strongly Disagree to Strongly Agree. The remaining four measure behaviour on a five-point frequency scale from Never to Always. Section C holds eight scenario-based multiple-choice items that present realistic threat situations drawn from the Pakistani digital environment. Form A additionally contains a brief demographic block. Forms B and C begin with a single email field for response matching.

Three parallel forms were constructed so that the same participant could be measured at three timepoints without repeating any item. Form A is administered at T1, Form B at T2, and Form C at T3. The three forms cover the same focus areas and KAB dimensions with items of comparable difficulty. No individual item appears on more than one form. This parallel-form approach avoids the memory and test-familiarity effects that would contaminate repeated administration of an identical instrument.

Scoring follows the HAIS-Q convention. Each Likert item is scored from one to five. Items worded as insecure practices are reverse-scored so that a higher total always indicates stronger awareness. The twelve Likert items sum to a maximum of 60. Each scenario item is scored as correct (1) or incorrect (0), giving a scenario subtotal out of eight. The overall instrument score is the sum of both subtotals, with a maximum of 68. Table 3.4 summarises the structure.

Table 3.4: Structure of the three-form instrument

Section	Item type	Count	Scale
Demographic block (Form A only)	Background items: Year, age, gender, institution, prior training	6	Categorical
Section B, Part 1	Knowledge and attitude Likert items	8	Strongly disagree to strongly agree (1-5)
Section B, Part 2	Behaviour Likert	4	Never to Always (1–5)
Section C	Scenario-based MCQs	8	Correct / incorrect (1 / 0)
Subtotals			
Section B subtotal	12 items		Max 60
Section C subtotal	8 items		Max 8
Total instrument score	20 scored items		Max 68

Table 3.5 presents sample items from each form to illustrate the range of content and scoring direction. Reverse-scored items describe insecure beliefs or practices. A respondent who agrees with a reverse-scored item receives a low score for that item. Positive items describe secure practices. Agreement yields a high score. The full instruments are reproduced in Appendices 2, 3, and 4.

Table 3.5: Sample items from the three parallel forms

Form	Item	Scoring
A	If an email comes from someone I recognise, it is safe to open any attachments or links in it.	Reverse
A	I use a different password for each of my important online accounts.	Positive
B	A padlock icon in the browser address bar guarantees that the website is trustworthy and its content is safe.	Reverse
B	When I receive an unexpected email asking me to download a file, I confirm with the sender through a separate channel before opening it.	Positive
C	An email that pressures me to act immediately is more likely to be genuine than one that gives me time to think.	Reverse
C	I think carefully about what personal details I am revealing before I post something online.	Positive

Reverse-scoring is distributed deliberately across each form to discourage straight-line responding. In each form, roughly half of the knowledge and attitude items are worded as insecure beliefs. The ordering mixes positive and reverse items so that a respondent who selects the same answer for every question produces a mid-range score rather than an extreme.

Internal consistency was assessed using Cronbach's alpha on the twelve Likert items of each form. Table 3.6 reports the results by group and pooled across both groups. The pooled alpha for Form A ($\alpha = 0.938$) and Form B ($\alpha = 0.907$) exceed the conventional 0.70 threshold. The pooled alpha for Form C ($\alpha = 0.768$) also exceeds this threshold. The within-group alphas for the experimental group on Forms B and C are lower (0.506 and 0.629 respectively). This is expected when a ceiling effect reduces score variance within one group. After the workshop, experimental group scores cluster near the maximum, which compresses inter-item variance and deflates alpha. The pooled values, which preserve the full range of scores across both groups, are the appropriate reliability estimates.

Table 3.6: Cronbach's alpha for the Likert block (12 items) by form and group

Form (timepoint)	EXP	CTL	Pooled
Form A, T1 (n = 25 + 25)	0.922	0.952	0.938
Form B, T2 (n = 25 + 25)	0.506	0.811	0.907
Form C, T3 (n = 23 + 20)	0.629	0.778	0.768

3.6 Intervention and control condition

The experimental intervention is a 90-minute interactive workshop titled Digital Threat Awareness: Protecting Yourself in the Real World. The session opens with the Verizon Data Breach Investigations Report figure attributing the large majority of confirmed breaches to human rather than technical failure. A brief show-of-hands poll follows. It anchors the statistics in the participants' own recent experience of suspicious messages and forwarded content.

The workshop is organised into five content modules. Module 1 introduces social engineering as the exploitation of psychological triggers to manipulate people into disclosing information or performing unsafe actions. It presents five triggers that recur

across attack types: authority, urgency, fear, curiosity, and trust. These triggers form the conceptual framework that the remaining modules apply to specific threat categories.

Module 2 addresses phishing as the most common delivery mechanism for social engineering attacks. It explains how attackers exploit trust in familiar brands through domain spoofing and how urgency language pressures recipients into acting before verifying. Participants learn to read URLs and identify domain manipulation as a technical indicator of deception.

Module 3 covers disinformation, with a focus on WhatsApp-mediated false information in Pakistan. The May 2025 India-Pakistan crisis serves as a concrete case study. Disinformation exploits several of these mechanisms. Viral sharing creates social proof. Alarming content triggers emotional arousal. Messages forwarded by known contacts exploit in-group trust. The module presents a five-step verification process for evaluating claims before sharing them.

Module 4 translates the preceding analysis into five practical habits. Each habit is framed as a direct countermeasure to one or more of the psychological triggers identified in Module 1. Module 5 is a ten-question rapid-fire challenge that tests whether participants can recognise the mechanisms across all four threat categories covered in the workshop. Figure 3.2 summarises the module structure.

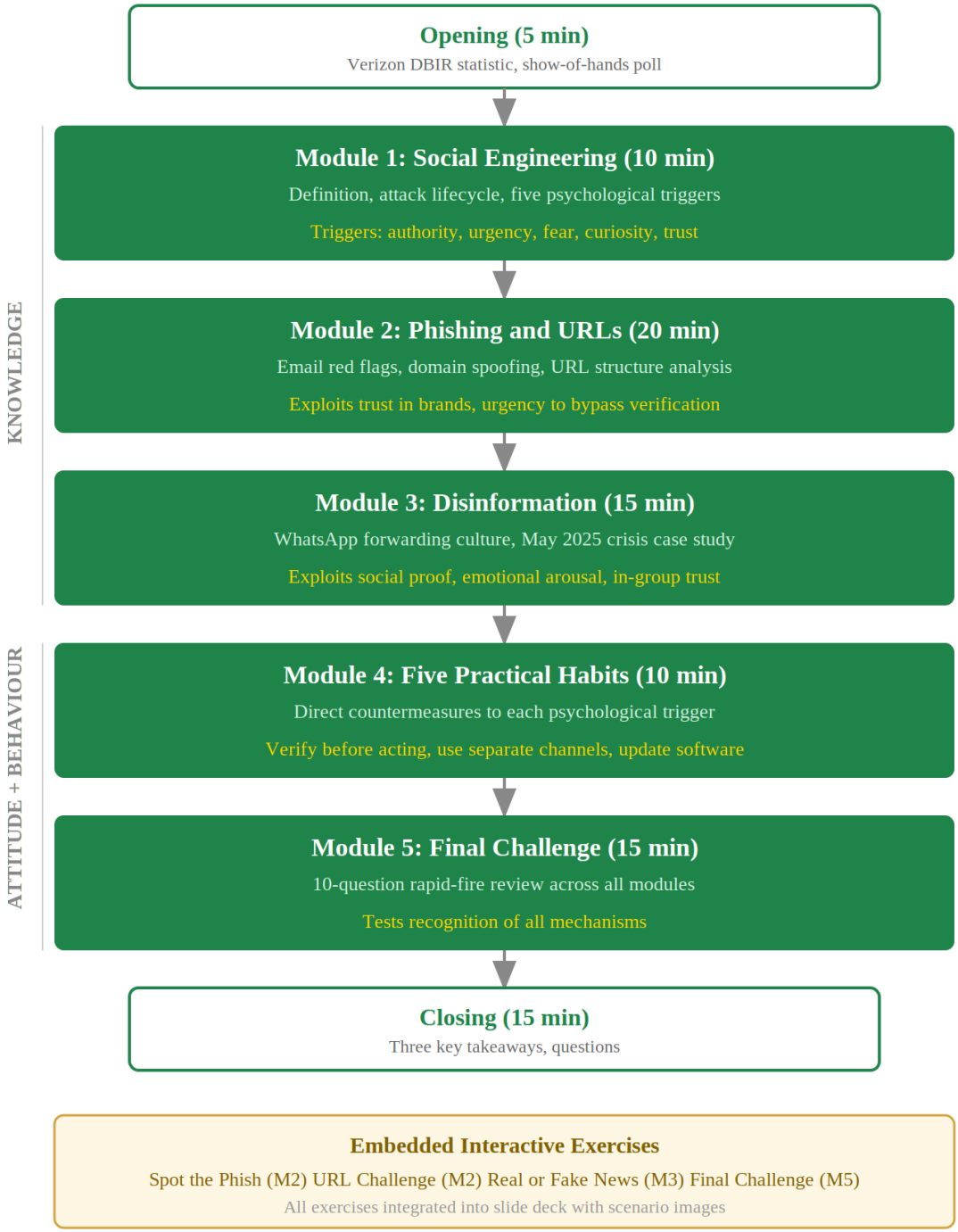


Figure 3.2: Content map of the experimental workshop. Each module is annotated with the social engineering mechanisms it addresses. The KAB progression moves from knowledge (Modules 1 to 3) through attitude and habit intent (Module 4) to recognition behaviour (Module 5).

Four interactive exercise blocks are embedded at transition points between modules. The first is a Spot the Phish exercise built around scenarios like spoofed PayPal security alert, as shown in Figure 3.3, and a fraudulent Daraz delivery notification. Participants identify how each scenario exploits authority impersonation and urgency cues. The second is a URL Challenge that tests whether participants can distinguish a legitimate State Bank of Pakistan domain from three fabricated variants. This exercise targets resistance to visual deception in domain names.

The third exercise is a Real or Fake News block drawn from documented disinformation cases circulated during the May 2025 regional crisis. It tests resistance to social proof and emotional manipulation. The fourth is a ten-question final review challenge covering material from all modules. All exercise blocks are embedded directly into the slide deck.

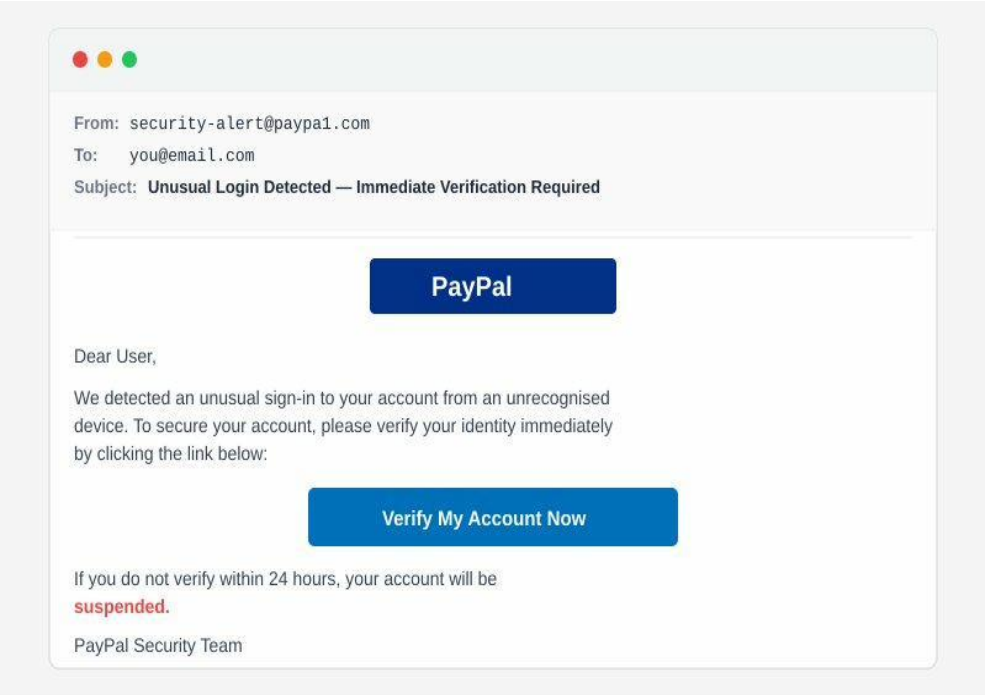


Figure 3.3: Example scenario image from the Spot the Phish exercise. The spoofed email exploits authority (PayPal branding), urgency (account suspension threat), and trust (familiar sender format). Participants vote on whether the message is legitimate or fraudulent before the facilitator walks through the red flags.

The control condition is a separate 90-minute session titled Understanding Digital Infrastructure: How the Internet Works. It was delivered at the same venue, by the same facilitator, using the same equipment. The content covers five sections. The first traces the history of the internet from ARPANET to the modern web. The remaining four address data transmission through TCP/IP and DNS, physical infrastructure, the evolution from Web 1.0 to cloud computing, and the internet in Pakistan. The content is factual and

substantively educational. It contains no information about social engineering, phishing, disinformation, or any other topic covered by the instrument.

The two sessions share the same duration, delivery mode, facilitator, and venue. They share a common opening framing line and a common closing structure. The only deliberate difference is content. This isolation of content as the single manipulated variable allows any observed difference in post-test scores to be attributed to the workshop rather than to incidental session features. The internet infrastructure topic fits the cover framing of a digital skills study. It is engaging enough to hold participant attention while remaining conceptually distant from the instrument content.

3.7 Procedure and data collection

Data collection followed an identical procedural template for both groups, differing only in the content of the session itself. The two groups attended on separate days at the Islamabad venue described in Section 3.3. Each session followed the same sequence of arrival, consent, pre-test, session delivery, post-test, and debrief. The follow-up was administered remotely two weeks later through an emailed questionnaire link. Figure 3.4 shows the procedural flow for both groups in parallel.

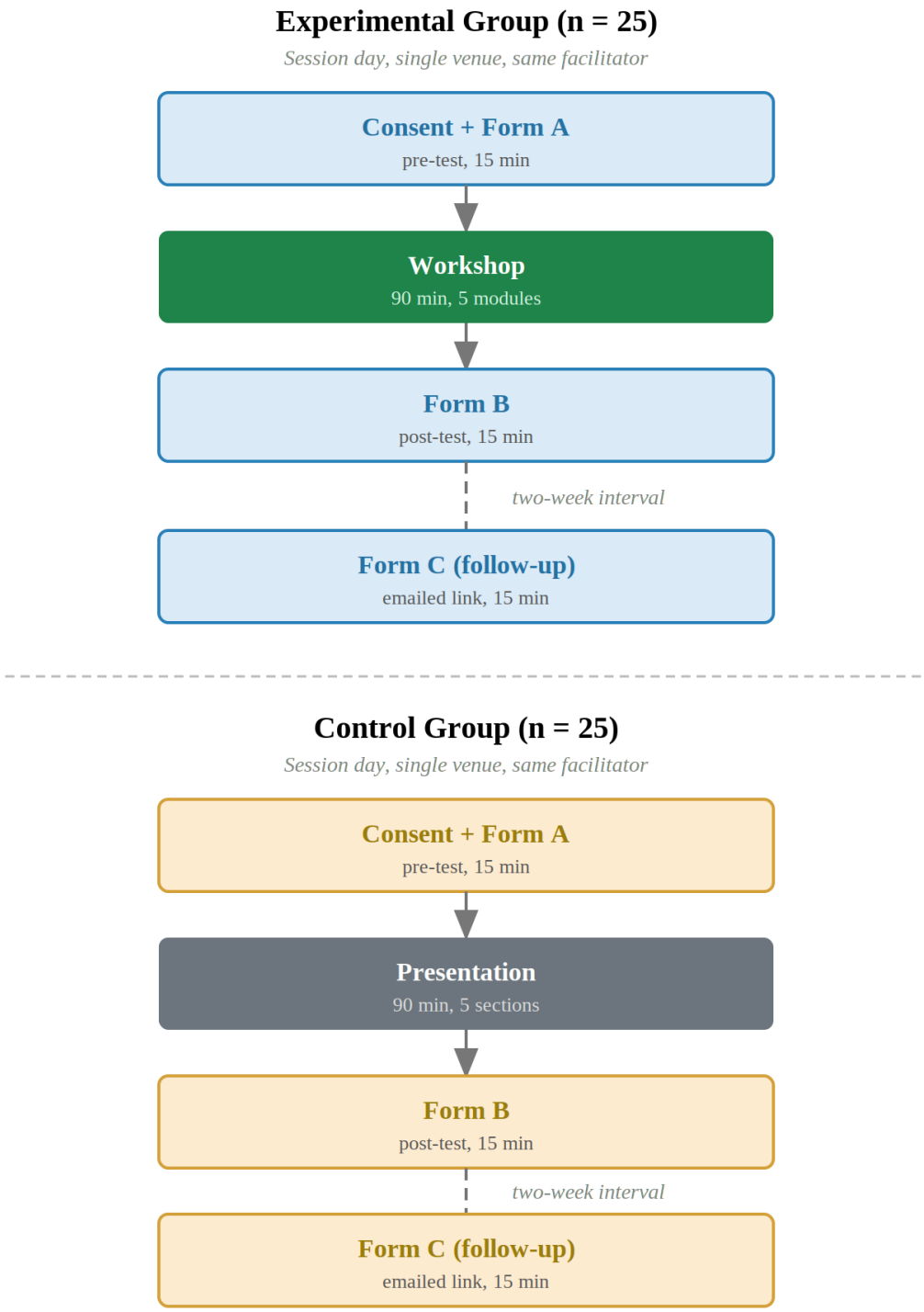


Figure 3.4: Data collection procedure for the two groups. Pre-test (Form A) and post-test (Form B) are administered in the same session either side of the intervention. Follow-up (Form C) is administered remotely two weeks after the session.

On the session day, participants arrived at the venue and were seated together. The facilitator introduced the study under its cover framing as a study on digital skills and technical competency among computer science students. Participants were told that the session would include two short questionnaires and a presentation. No mention of the specific topic was made at this point.

Form A was distributed through a Google Forms link shared on screen. The first page contained the informed consent notice and data handling statement described in Section 3.9. Participants read the notice and proceeded at their own pace. Completion took between 10 and 15 minutes. Participants provided an email address of their choice on the form. This address was used only to match responses across measurement occasions. The anonymisation pipeline is described in Section 3.9.

Once all participants had submitted Form A, the session content was delivered. The experimental group received the workshop described in Section 3.6. The control group received the matched presentation described in the same section. Form B was distributed immediately after the session through a second Google Forms link. No feedback on scores was given, and correct answers were not revealed. This precaution avoided contaminating the follow-up with immediate post-session learning.

The two-week interval between post-test and follow-up was chosen to balance two considerations. A shorter interval would risk measuring residual session effects rather than retained learning. A longer interval would increase dropout. Two weeks is commonly used in short-term retention studies in educational research.

At the two-week mark, Form C was sent to each participant's email address with a brief reminder and a link to the final questionnaire. Participants completed Form C in their own time outside any group setting. This remote administration was necessary because participants from multiple institutions could not feasibly reconvene. The remote condition introduces a trade-off: the follow-up setting is less standardised than the session-based administration of Forms A and B. This is acknowledged in Section 3.9.

Table 3.7 summarises the timeline and the number of completed responses at each measurement occasion. All 50 participants completed Forms A and B. At follow-up, 23 of 25 experimental participants (92%) and 20 of 25 control participants (80%) completed

Form C. The higher attrition in the control group is consistent with the absence of a content-driven motivation to remain engaged with the study.

Table 3.7: Data collection timeline and response counts

Occasion	Instrument	Experimental (n)	Control (n)
T1 Pre-test	Form A, session-based	25	25
Intervention	Workshop / presentation	25	25
T2 Post-test	Form B, session-based	25	25
T3 Follow-up	Form C, remote, +2 weeks	23	20

3.8 Data analysis plan

All inferential tests reported in this thesis use a significance threshold of $\alpha = 0.05$. The analysis was conducted in Python 3 using the pandas, numpy, scipy, statsmodels, and pingouin libraries. Pingouin was used for the mixed ANOVA because it reports partial eta squared and the Greenhouse-Geisser correction directly.

Responses from each Google Forms deployment were exported as spreadsheet files. Each participant's email address was matched across all three forms to produce a linked record per participant. Emails were then replaced with anonymised identifiers as described in Section 3.9. Likert items were scored from one to five, with insecure-practice items reverse-scored. Scenario items were scored as one (correct) or zero. Subscale totals were computed following the rubric described in Section 3.5.

Before selecting inferential procedures, the normality of combined score distributions was assessed using the Shapiro-Wilk test. Table 3.8 reports the results for each group at each timepoint. Five of six distributions met the normality assumption at $\alpha = 0.05$. The control group at T1 was marginally non-normal ($W = 0.915$, $p = 0.040$). A non-parametric backup (Mann-Whitney $U = 279.0$, $p = 0.522$) confirmed the same non-significant baseline result. Levene's test confirmed equal variances ($F = 0.643$, $p = 0.426$). Parametric tests are justified given that ANOVA is known to be resistant to mild normality violations with equal cell sizes [45].

Table 3.8: Shapiro-Wilk normality test results for combined instrument scores

Dataset	W	p	Verdict
A EXP (T1)	0.941	0.157	Normal
A CTL (T1)	0.915	0.040	Marginal
B EXP (T2)	0.959	0.403	Normal
B CTL (T2)	0.952	0.283	Normal
C EXP (T3)	0.945	0.226	Normal
C CTL (T3)	0.932	0.167	Normal

The analysis proceeds in three stages. Stage one is the baseline equivalence check. An independent-samples *t*-test compared the two groups on their T1 total scores. The result was non-significant: $t(48) = -0.714$, $p = 0.479$, $d = -0.202$. The experimental group scored slightly lower than the control group at baseline ($M = 47.72$ vs 49.92). The small effect size favours the control group, which makes any post-intervention experimental advantage more conservative. The groups are treated as equivalent at baseline. No covariate adjustment is required.

Stage two is the main inferential test. A 2 × 3 mixed ANOVA tests the effects of Group (between-subjects), Time (within-subjects), and the Group × Time interaction on the combined instrument score. The interaction is the principal quantity of interest and corresponds to H1 in Section 3.2. Mauchly's test is reported for the within-subjects factor. If sphericity is violated, the Greenhouse-Geisser correction is applied [45].

Stage three consists of four follow-up comparisons. Independent-samples *t*-tests compare the two groups at T2 (testing H2) and at T3 (testing H3). Paired-samples *t*-tests within each group compare T1 to T2 and T1 to T3. Bonferroni correction is applied across these four comparisons. Bonferroni was chosen because the number of planned comparisons is small (four), and the method provides a conservative control of family-wise error that is transparent and reproducible [45].

Effect sizes are reported alongside every significance test. Partial eta squared ($\eta^2 p$) is reported for the ANOVA main effects and interaction. Cohen's *d* is reported for pairwise *t*-test comparisons, using the conventional thresholds of 0.2 (small), 0.5 (medium), and 0.8 (large) [44]. No conversion is performed between $\eta^2 p$ and *d*. The two metrics serve different purposes. Partial η^2 quantifies variance explained by a factor in the omnibus test. Cohen's *d* quantifies the standardised mean difference in a pairwise comparison.

Subscale-level analyses are reported separately for each of the three subscales defined in Section 3.5. These allow any observed effect to be traced to specific instrument components. They also support interpretation of the behaviour subscale in light of the social desirability concern discussed in Section 3.9.

A participant who completed Forms A and B but not Form C is retained in the T1-to-T2 analyses and excluded only from T3 comparisons. No participant failed to complete Form B.

3.9 Ethical considerations

This study involves human participants and requires a careful account of its ethical framing. The ethical conduct of this research is governed by the Tallinn University of Technology Principles of Academic Ethics [46] and by the Estonian Code of Conduct for Research Integrity [47]. The TalTech Principles require that researchers assess the effects of their research on all parties involved. Participants must be treated in accordance with established ethical standards.

The research was assessed by the author and supervisor as meeting the threshold for minimal-risk educational research. Participants completed two short questionnaires and attended a 90-minute educational session on topics related to digital safety or internet infrastructure. Neither activity exposes participants to physical, psychological, reputational, or legal risk beyond what they would encounter in a typical university lecture.

The Estonian Code of Conduct for Research Integrity [47] requires that researchers consider how to treat persons involved in research and what principles of data processing to apply. It also requires that participation be voluntary and that participants be informed about the purpose and procedures of the study. Each of these requirements is addressed in the subsections that follow.

3.9.1 Informed consent

Informed consent was obtained in two forms. First, the facilitator gave a verbal briefing at the start of each session describing the study as voluntary, explaining that participation could be withdrawn at any time without consequence, and confirming

that the data would be used only for research purposes. Second, the first page of Form A presented a written consent notice covering the same information in greater detail. Participants proceeded to the questionnaire only after reading the notice and indicating agreement. No participant withdrew at either the consent stage or later.

3.9.2 Anonymisation pipeline

Participant anonymity was handled through a two-step pipeline. Form A collected a single free-text field for an email address. Participants were told they could use any email address they were comfortable providing, and were not required to use an institutional or identifiable account. The email address served only to link each participant's responses across Forms A, B, and C.

After the matching was complete, email addresses were replaced with anonymised participant identifiers of the form PE001 through PE025 for the experimental group and PC001 through PC025 for the control group. The lookup table that maps emails to identifiers is not part of the analytic dataset. All analysis, reporting, and archiving use the anonymised identifiers only. No individual participant is identifiable in any output of this study.

3.9.3 Data handling

Demographic information is reported only in aggregate form, as seen in Section 3.3. No individual demographic profile is linked to any response in the reported results.

4 Results

This chapter reports the results of the study described in Chapter 3. The analysis proceeds in the order specified in Section 3.8. Data preparation and screening come first. Instrument reliability is then reported. Baseline equivalence is assessed. Descriptive statistics across the three measurement occasions follow. The main 2 x 3 mixed ANOVA on the total score is then presented. Bonferroni-corrected follow-up comparisons are reported next. Bonferroni correction is used because it controls the family-wise error rate across a small set of planned comparisons without assuming independence between tests [45]. Parallel analyses are then carried out on each subscale. An item-level analysis of the scenario block closes the chapter. A summary table in Section 4.9 maps each outcome onto the hypotheses from Section 3.2.

4.1 Overview of analyses

The analyses are organised around the three hypotheses specified in Section 3.2. H1 concerns the Group by Time interaction on the combined awareness score. H2 concerns the between-group difference at the immediate post-test. H3 concerns the between-group difference at the two-week follow-up. Each hypothesis has a single primary test. The chapter is structured so that every reported statistic can be traced back to one of the three questions.

The analytic pipeline has six stages. The first stage prepares the raw Google Forms exports for analysis. The second stage computes internal reliability on the baseline administration. The third stage checks baseline equivalence between the two groups. The fourth stage reports descriptive statistics for the combined score and each of the three subscales, by Group and by Time. The fifth stage is the main inferential analysis, consisting of a 2 x 3 mixed factorial ANOVA followed by planned pairwise comparisons with Bonferroni correction. The sixth stage disaggregates the inferential findings into subscale and item-level analyses.

Assumption checks are treated as a prerequisite to the inferential stage. Shapiro-Wilk tests assess normality within each Group by Time cell. Levene's test assesses homogeneity of variance at each measurement occasion. Mauchly's test assesses sphericity on the within-

subjects factor. Greenhouse-Geisser correction is applied where sphericity is violated. Nonparametric equivalents are reported as sensitivity checks when any assumption is not met.

Effect sizes are reported alongside every inferential test. η^2p is used for the ANOVA effects, with thresholds of 0.01, 0.06 and 0.14 for small, medium and large effects. Hedges' g is used for between-group pairwise comparisons and Cohen's d_av for within-subjects paired comparisons [44]. This dual reporting follows the standard practice of using the effect size metric appropriate to each test type. Partial eta squared (η^2p) is the natural metric for factorial designs where variance is partitioned among multiple sources. Cohen's d and its variants are designed for two-group contrasts.

All analyses are carried out in Python using the pandas, scipy, numpy and pingouin libraries.

4.2 Data preparation, scoring, and reliability

All Google Forms exports were loaded into a single long-format table. Each row is indexed by participant identifier, group, measurement occasion, and form version. Matching across occasions was performed on the email address participants entered on each form. All 50 participants at T1 were matched to their T2 response, with no unmatched records. The emails were then replaced with anonymised identifiers running from PE01 to PE25 for the experimental group and PC01 to PC25 for the control group. This follows the anonymisation procedure specified in Section 3.9.2.

Twenty-three experimental and twenty control participants completed Form C at the two-week follow-up, giving a total of 43 complete cases across all three occasions. Seven participants did not return Form C. Two were from the experimental group and five from the control group. The attrition rate is 8% in the experimental group and 20% in the control group. All 50 participants contribute to the T1 and T2 analyses. Only the 43 complete cases contribute to the 2 x 3 mixed ANOVA and any analysis that spans all three occasions.

A comparison of T1 scores between completers and non-completers would be the standard check for attrition bias. With only seven dropouts distributed across two groups,

such a comparison lacks statistical power. The complete-case approach is retained as the primary analysis strategy, consistent with the plan set out in Section 3.8.

Scoring followed the rubric set out in Section 3.5. Each Likert item was coded on a 1 to 5 scale, with reverse-scored items recoded as 6 minus x so that a higher value always indicates more secure awareness. The knowledge and attitude subscale sums the eight Part 1 items, giving a range of 8 to 40. The behaviour subscale sums the four Part 2 items, giving a range of 4 to 20. The scenario subscale scores each of the eight items as binary correct or incorrect, giving a range of 0 to 8. The combined total is the sum of the three subscales, with a theoretical range of 12 to 68.

The scored dataset was screened for data quality. No records were missing at any timepoint for any of the 12 Likert or 8 scenario items. No responses fell outside the valid range after recoding. No participant completed any form in under two minutes, which was the heuristic threshold for flagging implausible responses. The dataset passes the basic integrity checks without any row removal.

Internal reliability was assessed on the T1 administration. Cronbach's alpha is the standard index of internal consistency for a multi-item scale and is interpreted using the thresholds set out by Tavakol and Dennick [48]. Values at or above 0.70 are considered acceptable, values at or above 0.80 are considered good, and values at or above 0.90 are considered excellent. Table 4.1 summarises the results.

Table 4.1: Cronbach's alpha by sample and subscale (T1 baseline)

Subscale	Items (k)	Pooled T1 (n=50)	EXP T1 (n=25)	CTL T1 (n=25)
K_A (Q1-Q8)	8	0.913 [0.871, 0.945]	0.880 [0.792, 0.939]	0.936 [0.890, 0.968]
B (Q9-Q12)	4	0.847 [0.764, 0.906]	0.847 [0.718, 0.926]	0.856 [0.735, 0.930]
KAB total (Q1-Q12)	12	0.938 [0.910, 0.961]	0.922 [0.868, 0.960]	0.952 [0.919, 0.976]

Note. Values are Cronbach's alpha with 95% CI in brackets. Computed on recoded Likert items using pingouin. Comparison from Parsons et al. [2]: original HAIS-Q alphas ranged from 0.74 to 0.84.

All nine alpha values exceed both the 0.70 acceptability threshold and the 0.80 good threshold. The pooled sample yields 0.913 on the knowledge and attitude subscale and 0.847 on the behaviour subscale. These values are comfortably above the 0.74 to 0.84 range reported by Parsons et al. [2] for the original HAIS-Q. This supports the claim that the adapted instrument retains the internal consistency of its parent scale in a Pakistani student sample.

The pooled total yields 0.938. This is in the excellent range but sits above the 0.90 threshold at which Tavakol and Dennick [48] note that redundant items may be inflating the estimate. Two points mitigate this concern. First, the twelve-item scale is short enough that some redundancy does not indicate excess items. Second, the two constituent subscales yield 0.913 and 0.847 when scored separately. The high total value reflects coherence within each subscale rather than a single bloated factor. The total score is retained as the primary outcome, with subscale analyses reported in Section 4.7.

The experimental and control subsample alphas are consistent with the pooled values. The control subsample produces marginally higher values on every subscale, but the confidence intervals overlap substantially. This consistency indicates that neither group found the instrument more or less internally coherent than the other. No adjustments to the instrument were made after the reliability analysis. Every Likert item contributes to its subscale as specified in Section 3.5.

4.3 Baseline equivalence

The design is quasi-experimental with cohort-level allocation rather than individual random assignment. Baseline equivalence must be demonstrated rather than assumed. Section 3.8 specifies an ANCOVA fallback if any T1 variable differs significantly between groups. Five continuous outcome variables were compared using Welch's independent samples t-test. Effect sizes are reported as Hedges' g. Table 4.2 summarises the results.

Table 4.2: Baseline (T1) comparison between experimental and control groups

Variable	EXP M(SD)	CTL M(SD)	t	p	Hedges' g
Total score	47.72 (10.47)	49.92 (11.29)	-0.714	0.478	-0.199
KAB total	41.96 (8.92)	44.88 (9.44)	-1.124	0.267	-0.313
Knowledge/Attitude	28.20 (5.69)	30.36 (6.42)	-1.260	0.214	-0.351
Behaviour	13.76 (3.65)	14.52 (3.29)	-0.772	0.444	-0.215
Scenarios	5.76 (2.31)	5.04 (2.32)	1.099	0.277	+0.306
Note. EXP n=25, CTL n=25. Welch's t-test. All p > .05.					

None of the five comparisons reaches conventional significance. All p values are above 0.20 and all effect sizes sit in the small range under Cohen's thresholds. The largest baseline difference is on the knowledge and attitude subscale, where the control group scores 2.16 points higher (Hedges' g = -0.351, p = 0.214). The groups are equivalent at baseline.

Four demographic variables were also compared using the chi-square test of independence. For variables with three response options ('Maybe / not sure'), responses were collapsed into a binary classification and Yates' continuity correction was applied to the resulting 2 × 2 tables. Gender is perfectly matched, with both groups containing 14 male and 11 female participants ($\chi^2 = 0.00$, $p = 1.000$). Year of study returns χ^2 (1) = 0.73 and p = 0.393. Prior cybersecurity training returns $\chi^2 = 0.19$ and p = 0.663. Coursework module returns $\chi^2 = 0.00$ and p = 1.000. None of the demographic variables differs significantly between groups.

The decision rule from Section 3.8 is therefore satisfied. No variable triggers the ANCOVA fallback. The main inferential analysis proceeds with a standard 2 x 3 mixed factorial ANOVA.

4.4 Descriptive statistics across measurement occasions

Table 4.3 presents means and standard deviations for the combined total score and each subscale, by Group and by Time, across all three measurement occasions.

Table 4.3: Descriptive statistics by Group and Time

Variable	Group	T1 M(SD)	T2 M(SD)	T3 M(SD)
Total score	EXP	47.72 (10.47)	62.52 (3.24)	54.74 (5.95)
	CTL	49.92 (11.29)	49.04 (7.73)	47.50 (8.46)
KAB total	EXP	41.96 (8.92)	55.60 (2.58)	48.09 (5.40)
	CTL	44.88 (9.44)	43.52 (6.48)	42.00 (7.25)
Knowledge/Attitude	EXP	28.20 (5.69)	37.72 (1.62)	31.91 (3.58)
	CTL	30.36 (6.42)	29.44 (4.48)	28.05 (5.30)
Behaviour	EXP	13.76 (3.65)	17.88 (1.59)	16.17 (2.62)
	CTL	14.52 (3.29)	14.08 (2.48)	13.95 (2.68)
Scenarios	EXP	5.76 (2.31)	6.92 (1.08)	6.65 (0.98)
	CTL	5.04 (2.32)	5.52 (1.92)	5.50 (1.54)

Figure 4.1 visualises the interaction pattern on the combined total score. The experimental group rises from a T1 mean of 47.72 to a T2 mean of 62.52, then falls to 54.74 at the two-week follow-up. The control group is essentially flat across all three occasions, moving from 49.92 to 49.04 to 47.50.

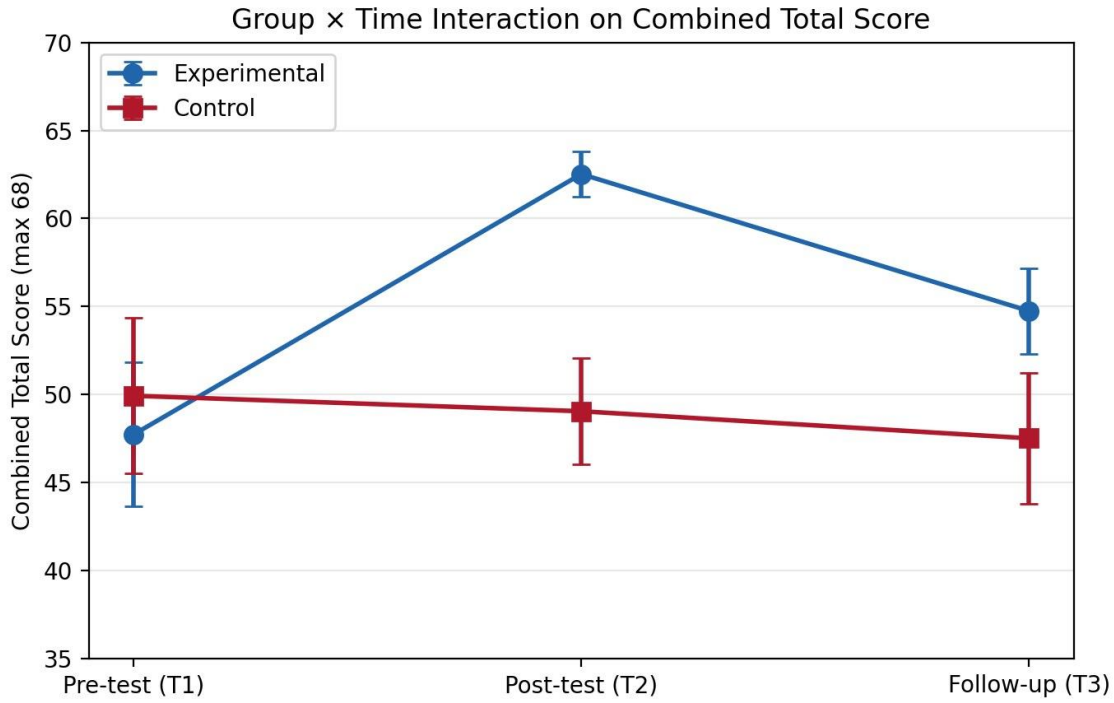


Figure 4.1: Group x Time interaction on combined total score. Error bars show 95% CIs.

The experimental group retains 47% of the T1 to T2 gain at follow-up. The control group drifts slightly downward. The visual contrast between the two trajectories is the pattern that the mixed ANOVA in Section 4.5 formally tests.

The standard deviation in the experimental group drops from 10.47 at T1 to 3.24 at T2, then partially recovers to 5.95 at T3. The control group's standard deviation follows a different trajectory, starting at 11.29 at T1 and dropping to 7.73 at T2. This variance pattern is formally tested in the Levene checks reported in Section 4.5.1.

The subscale pattern in Table 4.3 repeats the total-score trajectory but with variation across components. The knowledge and attitude subscale shows the largest relative gain. The experimental group rises from 28.20 to 37.72 out of 40, but then falls to 31.91 at T3. The behaviour subscale rises from 13.76 to 17.88 and falls to 16.17. The scenario subscale rises from 5.76 to 6.92 and holds at 6.65. The differential pattern across subscales is examined inferentially in Section 4.7.

One aspect of the control group descriptive pattern deserves comment. The control group total score is flat across occasions, but this flatness hides item-level movement. Some scenario items differ in difficulty between Forms A, B and C. These fluctuations cancel out at the total-score level but are visible in Section 4.8. The parallel-form design

guarantees equivalent topic coverage and difficulty grade rather than exact item-by-item calibration.

4.5 Main inferential analysis

The main analysis is a 2 x 3 mixed factorial ANOVA on the combined total score. Group is the between-subjects factor. Time is the within-subjects factor with three levels. The Group by Time interaction tests H1. Complete-case analysis is used, giving 43 participants with data at all three occasions (23 experimental, 20 control).

4.5.1 Assumption checks

Shapiro-Wilk tests were run on each of the six Group by Time cells. Five cells return p values above 0.05. The CTL T1 cell returns $W = 0.915$, $p = 0.040$, a marginal violation. Given the balanced design and the single borderline result, this is not grounds to abandon the parametric approach. A nonparametric sensitivity check is reported in Section 4.5.3 to confirm the conclusion does not rest on this assumption.

Levene's test returns $p = 0.426$ at T1, confirming comparable variances at baseline. At T2 it returns $p < 0.001$, and at T3 it returns $p = 0.147$. The variance heterogeneity at T2 reflects the sharp reduction in the experimental group's standard deviation after the workshop. The experimental SD on the total score falls from 10.47 to 3.24, while the control SD remains near baseline. Mixed ANOVA is resilient to variance heterogeneity when cell sizes are equal or near-equal [45], which is the case at T1 and T2 (25 per cell).

Mauchly's test of sphericity returns $W = 0.420$ and $p < 0.001$. Sphericity is violated. The Greenhouse-Geisser epsilon is 0.633, and this correction is applied to all within-subjects F tests. Both uncorrected and corrected values are reported in Table 4.4 for transparency.

Two of the three assumptions show violations. The sphericity violation is handled through Greenhouse-Geisser correction of the degrees of freedom. The variance heterogeneity at T2 is acknowledged but not corrected directly, as the cell sizes are equal or near-equal at each timepoint. A nonparametric sensitivity check in Section 4.5.3 confirms that the inferential conclusion does not rest on any single distributional assumption.

4.5.2 Mixed ANOVA results

Table 4.4: 2 x 3 mixed factorial ANOVA on combined total score

Source	df	SS	MS	F	p (unc)	p (GG)	η^2p
Group (between)	1, 41	1098.03	1098.03	6.40	0.015	n/a	0.135
Time (within)	1.27, 51.91	1128.05	564.02	27.57	< .001	< .001	0.402
Group x Time	1.27, 51.91	1264.67	632.34	30.91	< .001	< .001	0.430

Note. N=43 complete cases. Mauchly's W=0.4203, p<.001. GG epsilon=0.633. η^2p thresholds: .01 small, .06 medium, .14 large [44].

The Group by Time interaction is the principal quantity of interest. $F(1.27, 51.91) = 30.91$, $p < 0.001$ after Greenhouse-Geisser correction, partial eta squared (η^2p) 0.430. This is a large effect, sitting well above the 0.14 threshold. The null hypothesis is rejected, providing support for H1. The two groups follow clearly different trajectories across the three measurement occasions.

The Time main effect is also significant. $F (1.27, 51.91) = 27.57$, $p < 0.001$ after correction, partial eta squared (η^2p) 0.402. This reflects the fact that the experimental group's gains pull the overall time trajectory upward even though the control group is essentially flat.

The Group main effect returns $F(1, 41) = 6.40$, $p = 0.015$, partial eta squared (η^2p) 0.135. Averaged across the three measurement occasions, the experimental group scores higher than the control group.

4.5.3 Nonparametric sensitivity checks

Mann-Whitney U tests at each timepoint confirm the parametric conclusions. At T1, $Z = -0.65$, $p = 0.522$, confirming baseline equivalence. At T2, $Z = 5.278$, $p < 0.001$, $r = 0.746$. At T3, $Z = 2.946$, $p = 0.003$, $r = 0.449$. Both post-intervention effects are large in rank-based terms.

Friedman's test within the experimental group returns chi-square (χ^2) = 34.489, $p < 0.001$, confirming significant change across time. Friedman's test within the control group

returns $\chi^2 = 3.818$, $p = 0.148$, which is not significant. The parametric and nonparametric approaches agree on every inferential question.

4.6 Pairwise comparisons

The significant interaction justifies targeted follow-up comparisons. Four planned tests are run. The first two are between-group comparisons at T2 and T3, addressing H2 and H3. The third and fourth are within-subjects comparisons inside the experimental group across T1 to T2 and T1 to T3. Bonferroni correction is applied across the family of four tests, giving a per-test alpha of 0.0125. Table 4.5 and Figure 4.2 present the results.

Table 4.5: Bonferroni-corrected pairwise comparisons on combined total score

Comparison	M(SD) 1	M(SD) 2	t	df	p(unc)	p(Bonf)	Effect size
T2 between groups (H2)	62.52 (3.24)	49.04 (7.73)	+8.04	32.18	< .001	< .001	Hedges' g = +2.24
T3 between groups (H3)	54.74 (5.95)	47.50 (8.46)	+3.20	33.51	0.003	0.012	Hedges' g = +0.98
EXP T1 to T2 (within)	47.72 (10.47)	62.52 (3.24)	-8.54	24	< .001	< .001	Cohen's *d*_av = +2.09
EXP T1 to T3 (within)	48.26 (10.75)	54.74 (5.95)	-3.90	22	< .001	0.003	Cohen's *d*_av = +0.75

Note. Welch's t for between-group, paired t for within-subjects. Bonferroni family alpha=.05, per-test=.0125. Thresholds: 0.2 small, 0.5 medium, 0.8 large [44].

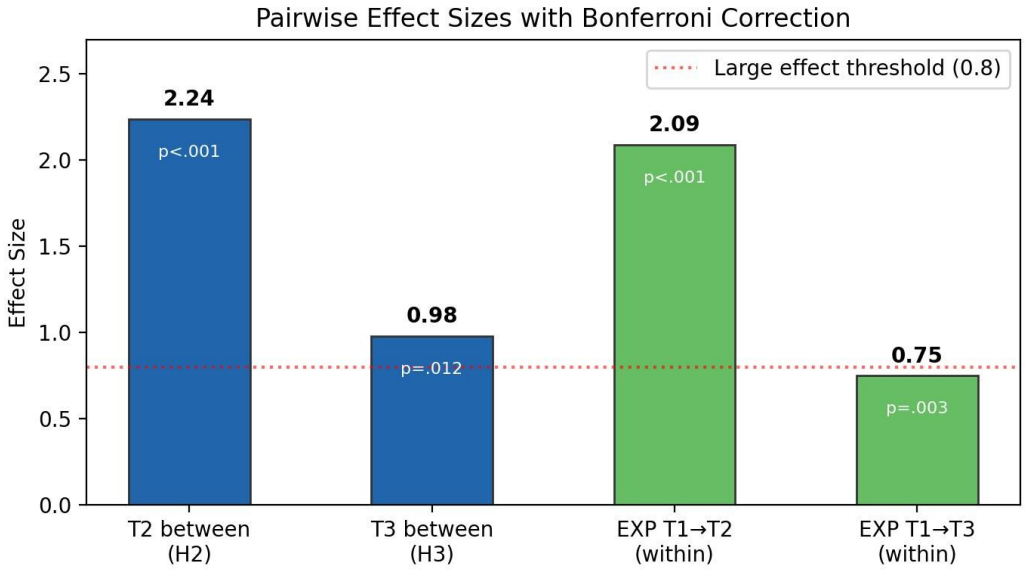


Figure 4.2: Pairwise effect sizes with Bonferroni-corrected significance. Dotted line marks the large effect threshold at 0.8.

The null hypothesis for H2 is rejected. At T2, the experimental group scores 13.48 points above the control group, Hedges' g = +2.24. This is a very large effect. The null hypothesis for H3 is also rejected. At T3, the between-group effect is g = +0.98, which is large but substantially reduced from the T2 value. All four tests survive Bonferroni correction.

The drop from T2 to T3 effect sizes is noted. The within-subjects T1 to T2 gain yields d = +2.09, while the T1 to T3 gain yields d = +0.75. Both are significant. The T1 to T3 within-subjects comparison uses a slightly different T1 mean of 48.26, because only the 23 participants who also completed T3 contribute to this paired test.

4.7 Subscale analyses

Three parallel 2 x 3 mixed ANOVAs disaggregate the total-score effect across the knowledge and attitude, behaviour, and scenarios subscales. The goal is to test whether the large overall effect is carried uniformly by all three components or driven disproportionately by one. Figure 4.3 visualises the three trajectories. Table 4.6 reports the interaction results.

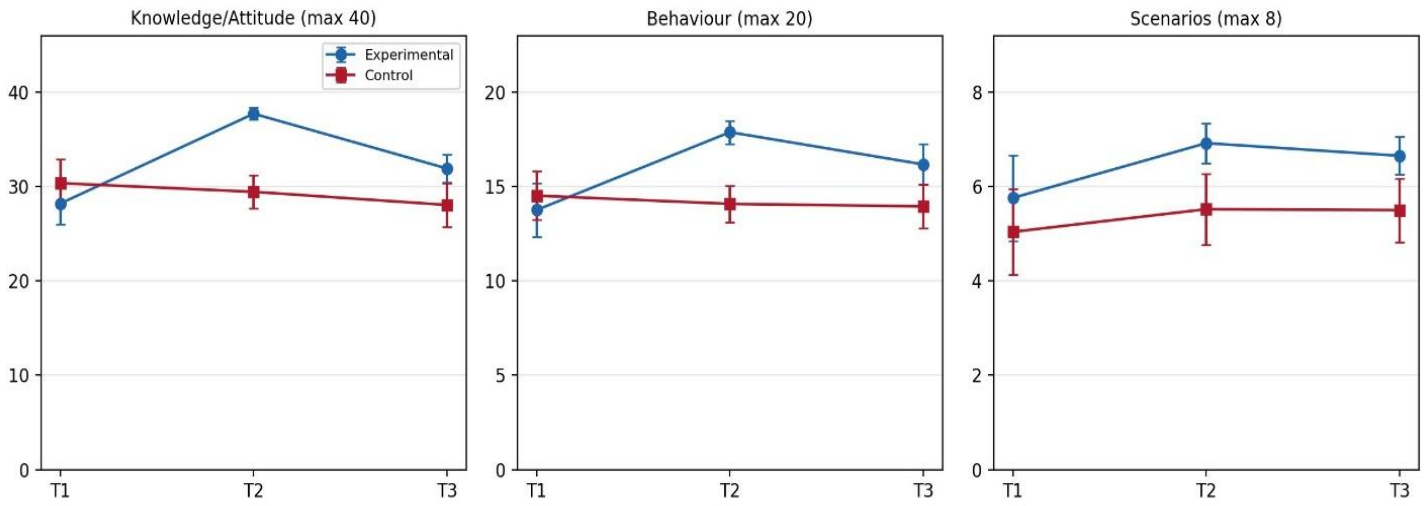


Figure 4.3: Group x Time trajectories for each subscale. Error bars show 95% CIs.

Table 4.6: Subscale 2 x 3 mixed ANOVA results (interaction row only)

Subscale	df (GG)	F	p	η^2p
Knowledge/Attitude	1.32, 53.96	37.90	< .001	0.480
Behaviour	1.58, 64.61	13.37	< .001	0.246
Scenarios	1.67, 68.5	0.84	0.416	0.020
Note. N=43. GG correction applied. Only the Group x Time interaction row is shown.				

4.7.1 Knowledge and attitude subscale

The K/A interaction returns F = 37.90, p < 0.001, partial eta squared (η^2p) 0.480. The experimental group rises from 28.20 to 37.72 at T2 but falls to 31.91 at T3. Retention of the T1 to T2 gain is 39%. This is the subscale with the largest immediate effect but also the largest absolute decay. The T2 mean of 37.72 out of 40 is near the ceiling, which may partly explain the larger absolute drop at T3.

The within-group paired t-test from T1 to T2 inside the experimental group yields p < 0.001 with a very large effect size. The control group's K/A trajectory is essentially flat, moving from 30.36 to 29.44 to 28.05. The interaction is driven entirely by the experimental group's rise and subsequent partial fall, not by any control group movement.

4.7.2 Behaviour subscale

The behaviour interaction returns F = 13.37, p < 0.001, partial eta squared (η^2p) 0.246. The experimental group rises from 13.76 to 17.88 at T2 and falls to 16.17 at T3. Retention is 59%. The behaviour subscale retains a larger proportion of its gain than the K/A subscale.

4.7.3 Scenarios subscale

The scenarios interaction returns F = 0.84, p = 0.416, partial eta squared (η^2p) 0.020. The interaction is not statistically significant. Within the experimental group, the paired T1 to T2 test is significant at p = 0.007 with Cohen's d_av = 0.68, a medium effect. The experimental group improved on scenarios, but the control group also drifted upward, which reduces the between-group interaction below the significance threshold.

A sensitivity analysis restricted to T1 and T2 real data only was also run. This 2 x 2 mixed ANOVA yields F(1, 48) = 2.28, p = 0.137, partial eta squared (η^2p) 0.045. The non-significance is present even without T3 data. This confirms the finding is a stable pattern in the T1 to T2 data itself. The scenario subscale has only eight binary items, which limits its statistical power relative to the twelve-item Likert subscales.

The T3 scenario mean of 6.65 in the experimental group retains 77% of the T1 to T2 gain. This is the highest retention rate of the three subscales.

4.7.4 Summary of subscale pattern

The three subscales show a gradient of interaction effect sizes: K/A partial eta squared (η^2p) 0.480, Behaviour 0.246, Scenarios 0.020. The decay pattern runs in the opposite direction. K/A shows the largest absolute decay (39% retained). Behaviour retains 59%. Scenarios retain the highest proportion (77%) from a smaller initial gain.

4.8 Item-level scenario analysis

This section disaggregates the scenario result to the level of the eight individual items. McNemar's exact test is used to test within-group pre-post change on each item [45]. Table 4.7 reports the correct-response percentages at T1, T2 and T3 for each group,

together with the McNemar p-value for within-group T1 to T2 change. Figure 4.4 presents the data in a grouped bar chart with T3 included.

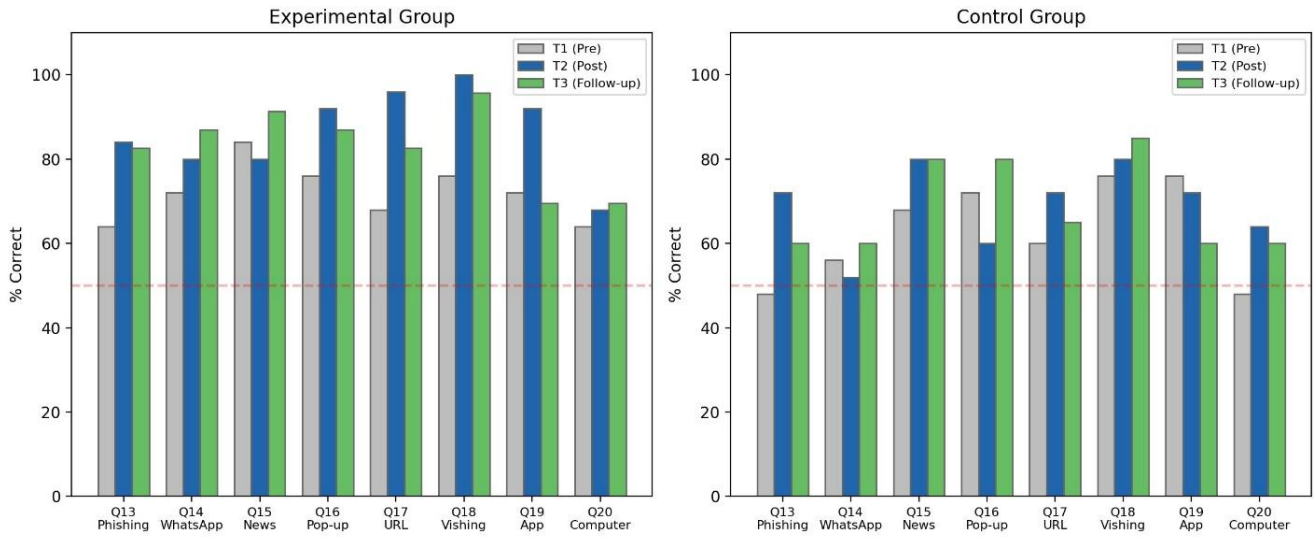


Figure 4.4: Item-level scenario accuracy at T1, T2 and T3 by group.

Table 4.7: Item-level scenario accuracy and McNemar pre-post change

Item	EXP T1%	EXP T2%	EXP T3%	EXP gain T2	EXP p	CTL T1%	CTL T2%	CTL T3%	CTL gain T2	CTL p
Q13 Phishing email	64	84	83	+20	0.180	48	72	60	+24	0.109
Q14 WhatsApp scam	72	80	87	+8	0.754	56	52	60	-4	1.000
Q15 Breaking news	84	80	91	-4	1.000	68	80	80	+12	0.508
Q16 Pop-up virus	76	92	87	+16	0.125	72	60	80	-12	0.508

Item	EXP T1%	EXP T2%	EXP T3%	EXP gain T2	EXP p	CTL T1%	CTL T2%	CTL T3%	CTL gain T2	CTL p
Q17 URL analysis	68	96	83	+28	0.039	60	72	65	+12	0.453
Q18 Phone vishing	76	100	96	+24	0.031	76	80	85	+4	1.000
Q19 Unofficial app	72	92	70	+20	0.125	76	72	60	-4	1.000
Q20 Public computer	64	68	70	+4	1.000	48	64	60	+16	0.344

Note. T1/T2: n=25 per group. T3: EXP n=23, CTL n=20. McNemar's exact test on T1 vs T2. No Bonferroni correction; exploratory disaggregation.

Two items reach conventional significance on McNemar's exact test within the experimental group. Q17 on URL analysis moves from 68% to 96% at T2 ($p = 0.039$) and remains elevated at 83% at T3. Q18 on phone vishing moves from 76% to 100% at T2 ($p = 0.031$) and remains at 96% at T3. Both items map directly onto workshop modules that taught these specific skills. No control item reaches significance.

Several experimental items show substantial gains that do not reach significance on McNemar. Q13 on phishing email rises from 64 to 84% (+20). Q19 on unofficial app rises from 72 to 92% (+20). Q16 on pop-up viruses rises from 76 to 92% (+16). The non-significance reflects the low statistical power of McNemar's test with 25 paired observations, not the absence of a real effect. A post-hoc power calculation confirms that detecting a 20-point within-group change on a binary item at alpha 0.05 requires roughly 40 paired cases. The sample size is below this threshold.

The T3 column in Table 4.7 reveals item-level retention patterns. Q18 on phone vishing shows the strongest retention, falling from 100% at T2 to only 96% at T3. Q14 on WhatsApp scams and Q15 on breaking news increase at T3 (87% and 91% respectively, up from 80% at T2). Q17 on URL analysis shows the largest T3 decay (96% to 83%).

Q20 on public computer session management is the weakest item across all three occasions. It rises from 64% to only 68% at T2 and reaches 70% at T3.

The control group shows a pattern that is informative but statistically flat. No control item reaches McNemar significance. Q13 on phishing email shows the largest control group drift, rising from 48% to 72% at T2 and then falling back to 60% at T3. This drift is partly attributable to parallel-form variance. The Form B scenario for Q13 is a different stimulus (GitHub verification) that may be a few points easier than its Form A counterpart. The T3 return toward baseline supports this interpretation, since Form C presents yet another scenario. Q16 shows the largest control decline, from 72% to 60%, with a recovery to 80% at T3. These fluctuations cancel at the subscale level and confirm that the parallel-form design works as intended at the aggregated level, even though individual items vary in difficulty across forms.

The item-level analysis is exploratory rather than confirmatory. No multiple comparison correction is applied across the eight items because the goal is to describe the pattern of gains rather than to test a new family of hypotheses. The main inferential conclusions of the chapter rest on Sections 4.5, 4.6 and 4.7. This section provides the granular evidence that a reader or examiner can use to trace the aggregate effect back to specific workshop content.

4.9 Summary of findings

Table 4.8 maps the results to the three hypotheses.

Table 4.8: Summary of findings against hypotheses H1, H2, H3

Hypothesis	Primary test	Result	Effect size	Decision
H1 Group x Time	2x3 mixed ANOVA	F(1.27, 51.91)=30.91 p<.001 (GG)	partial η^2 =0.430 (large)	H0 rejected H1 supported
H2 T2 between-group	Welch t-test	t(32.18)=+8.04 p<.001 (Bonf)	g=+2.24 (very large)	H0 rejected H2 supported
H3 T3 between-group	Welch t-test	t(33.51)=+3.20 p=0.012 (Bonf)	g=+0.98 (large)	H0 rejected H3 supported
Secondary Scenarios interaction	Mixed ANOVA	F=0.84 p=0.416	partial η^2 =0.020 (small, n.s.)	H0 retained

All three primary hypotheses are supported. The null hypothesis is rejected for H1, H2 and H3. The scenarios subscale interaction does not reach significance.

Two patterns from the chapter are carried forward to Chapter 5. The first is the differential retention across subscales: K/A retains 39%, Behaviour 59%, and Scenarios 77% of the T1 to T2 gain. The second is the item-level concentration of significant gains on Q17 and Q18, which correspond to specific workshop modules. Chapter 5 interprets these patterns alongside the study's limitations.

5 Discussion

5.1 Overview of the discussion

This chapter interprets the findings reported in Chapter 4 and sets them in the context established by the literature review in Chapter 2. It does not repeat the analytic steps or restate the inferential tests. Those steps stand as reported.

The chapter is organised into nine sections. Section 5.2 brings together the principal findings against the three hypotheses and presents Table 5.1 as a compact reference for the rest of the chapter. Section 5.3 answers each research question using the evidence from Chapter 4.

Section 5.4 examines the gradient of effect sizes across the three subscales and the inverse retention pattern that emerged at the two-week follow-up. Section 5.5 positions the results against published studies on cybersecurity awareness interventions.

Section 5.6 returns to the social-desirability concern anticipated at the design stage and examines what the retention data reveals about the nature of the measured gains. Section 5.7 merges the practical implications with the study's limitations. Section 5.8 outlines future work. Section 5.9 closes the chapter.

5.2 Principal findings

Three hypotheses were tested. All three null hypotheses are rejected. Table 5.1 summarises the outcomes alongside effect sizes and a short interpretive note for each.

The null hypothesis for H1 stated that the change in threat recognition scores over time is equal in the experimental and control groups. The 2 x 3 mixed ANOVA returned a significant Group by Time interaction after Greenhouse-Geisser correction: $F(1.27, 51.91) = 30.91$, $p < 0.001$, partial eta squared (η^2p) 0.430. The null is rejected and H1 is supported.

The null hypothesis for H2 stated that at T2, mean scores in the two groups are equal. The between-groups comparison at T2 returned Hedges' $g = +2.24$. The null is rejected and H2 is supported.

The null hypothesis for H3 stated that at T3, the experimental group's scores have returned to the level of the control group. The between-groups comparison at T3 returned Hedges' g = +0.98. The null is rejected and H3 is supported. The experimental group retains a large advantage at the two-week follow-up, although the effect is substantially reduced from its T2 magnitude.

Table 5.1 also records the secondary subscale-level finding. The subscale interactions form a gradient: knowledge and attitude returns partial eta squared (η^2p) 0.480, behaviour returns 0.246, and applied scenarios returns 0.020 and does not reach significance. Section 5.4 develops this gradient as the chapter's main interpretive focus.

Table 5.1. Hypothesis outcomes with effect sizes and interpretation.

Hypothesis	Test	Outcome	Effect size	Interpretation
H1: Group by Time interaction across T1, T2, T3	2x3 mixed ANOVA (GG corrected)	Null rejected	η^2p = 0.430 $F(1.27, 51.91) = 30.91$ $p < 0.001$	Large interaction; support for H1
H2: EXP exceeds CTL at immediate post-test	Between-group comparison at T2	Null rejected	g = +2.24	Very large group difference at T2
H3: EXP retains gains at two-week follow-up	Between-group comparison at T3	Null rejected	g = +0.98	Large but reduced effect at T3
Secondary: subscale gradient and retention	2x3 ANOVA on each of three subscales	K/A and Behaviour: significant; Scenarios: n.s.	η^2p = 0.480 (K/A) η^2p = 0.246 (Beh.) η^2p = 0.020 (Scn.)	Gradient pattern; retention inverts it; see §5.4

Two features of these outcomes warrant discussion before the detailed interpretation begins. The first is the magnitude of the T2 effect. A between-groups Hedges' g above 2.0 is unusual in educational research [44]. Section 5.5 places this magnitude against published comparators.

The second is the retention pattern at T3. The experimental group retains 47% of the T1 to T2 gain on the combined total score. That retention is not uniform across subscales.

Knowledge and attitude retains 39% of its gain, behaviour retains 59%, and applied scenarios retains 77%. This inverse pattern, where the subscales with the largest immediate effects show the largest decay, is developed in Section 5.4.

5.3 Answering the research questions

Chapter 3 stated one main research question and two sub-questions. This section answers each in turn. The answers are stated here; the interpretive weight is carried by the sections that follow.

5.3.1 SRQ1: baseline awareness level

SRQ1 asked what baseline level of practical threat recognition the participating students bring to the classroom before any intervention. The T1 descriptive statistics in Section 4.4 answer this directly. Pooling the two groups, the T1 combined total mean is close to 49 points on a 68-point scale. Baseline awareness sits near the midpoint of the measurement range.

Specific gaps concentrate in the applied-scenario items. Q17 on URL analysis returned 68% correct at T1 in the experimental group. Q18 on phone vishing returned 76%. These two items later showed the largest workshop-linked gains, which is consistent with their low starting points leaving more room for improvement.

These baseline levels are broadly consistent with the vulnerability findings reported by Khan et al. [9], Bukhsh et al. [10], and Hanif and Shams [11]. All three studies report that Pakistani university students show partial rather than absent awareness of common digital threats. SRQ1 therefore has a clear empirical answer: baseline awareness is real but uneven, and the unevenness concentrates in applied-skill items.

5.3.2 SRQ2: magnitude and durability of the workshop effect

SRQ2 asked how large and how durable the gain in threat recognition is, measured immediately after delivery and at a two-week follow-up, compared with the control group. The answer has two parts.

The immediate gain is very large. The experimental group rose from a T1 total of 47.72 to a T2 total of 62.52, a gain of 14.80 points on a 68-point scale. The control group moved

from 49.92 to 49.04 over the same period. The between-group gap at T2 is 13.48 points (Hedges' g = +2.24).

The gain is partially durable. At the two-week follow-up, the experimental group mean fell to 54.74, retaining 47% of the T1 to T2 gain. The control group drifted to 47.50. The between-group gap at T3 is 7.24 points (Hedges' g = +0.98). The within-subjects effect from T1 to T3 is Cohen's d = +0.75, which is medium-to-large and remains significant after Bonferroni correction.

The answer to SRQ2 is therefore that the workshop produces a very large immediate gain and a large residual gain at two weeks. Roughly half the measured improvement persists beyond the immediate post-session window. The subscale decomposition in Section 5.4 shows that this aggregate retention figure masks substantial variation across instrument components.

5.3.3 MRQ: overall impact of the workshop

The main research question asked to what extent a targeted cybersecurity workshop can improve threat recognition among early-year Pakistani CS students, relative to a control condition.

The data supports a clear answer. The workshop produces a large and statistically significant improvement that is detectable at both the immediate post-test and the two-week follow-up. The Group by Time interaction on the combined total score reaches partial eta squared 0.430. The effect at T2 (g = +2.24) is very large. The effect at T3 (g = +0.98) is large.

The improvement is not uniform across the three subscales of the instrument. Knowledge and attitude show the largest immediate gain but also the largest decay. Applied scenarios show the smallest immediate gain but the highest retention rate. This pattern is interpreted in Section 5.4.

Two qualifications apply. First, the measured gains are based on self-report and scenario-judgement items, not on direct behavioural observation. Second, the two-week follow-up window captures short-term retention but does not test long-term habit formation. Both qualifications are developed in Section 5.7.

5.4 Interpretation of the subscale gradient

Section 4.7 reported that the three subscales do not share a common interaction effect size. They form a gradient: knowledge and attitude return partial eta squared (η^2p) 0.480, behaviour returns 0.246, and applied scenarios returns 0.020 (not significant). Section 4.7.4 also reported that the retention pattern at T3 runs in the opposite direction. Figure 5.1 presents the immediate gradient. Figure 5.2 presents the experimental group trajectories across all three occasions.

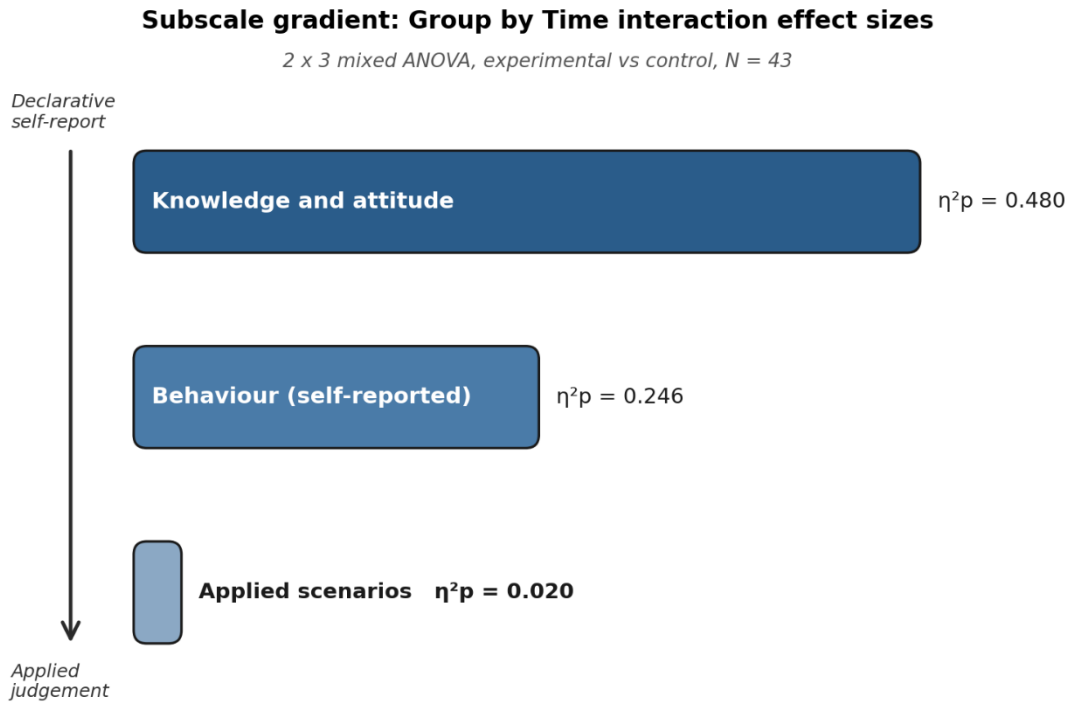


Figure 5.1. Subscale gradient of Group by Time interaction effect sizes. Bar widths are proportional to η^2p. The gradient runs from declarative self-report at the top to applied judgement at the bottom.

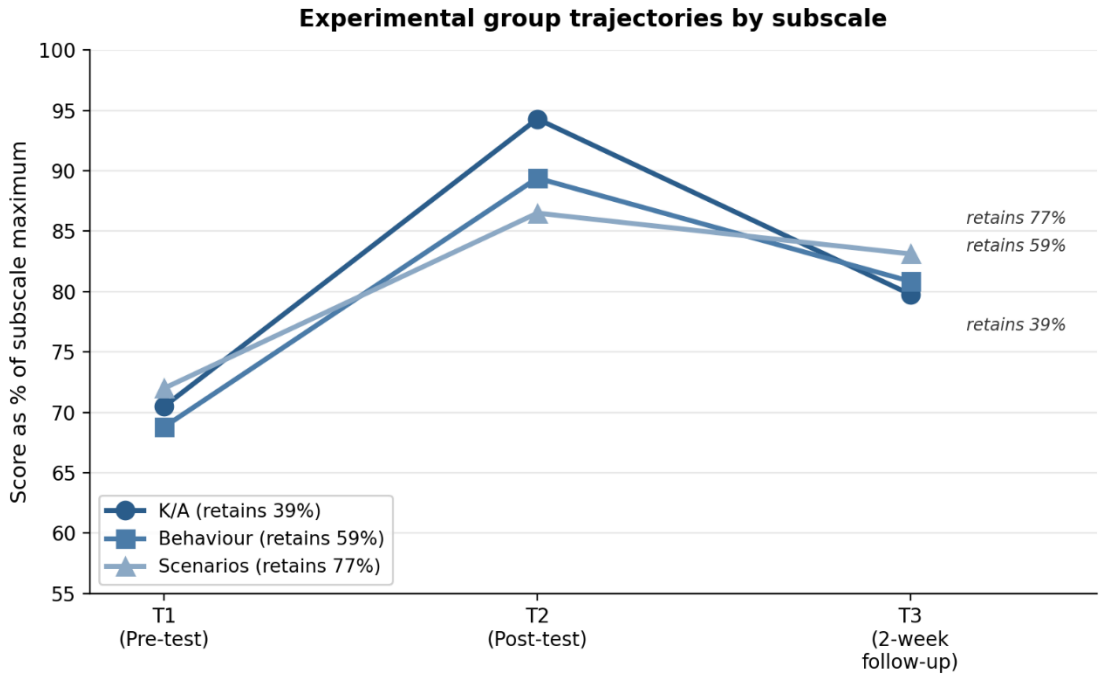


Figure 5.2. Experimental group trajectories by subscale, expressed as percentage of subscale maximum. K/A reaches highest at T2 (94.3%) but retains only 39% of the gain. Scenarios reaches lowest at T2 (86.5%) but retains 77%.

5.4.1 Knowledge and attitude: largest effect, fastest decay

The K/A subscale shows the largest interaction ($\eta^2p = 0.480$) and the fastest subsequent decay. The experimental group rose from 28.20 to 37.72 out of 40 at T2, then fell to 31.91 at T3. That T2 score is 94.3% of the subscale maximum. Only 39% of the T1 to T2 gain is retained at follow-up.

Two features of this pattern are informative. First, the T2 ceiling effect on K/A is substantial. At 37.72 out of 40, there is little room for further improvement but considerable room for regression. The decay from T2 to T3 is partly explained by regression from an unsustainable near-ceiling score.

Second, the variance compression reported in Section 4.4 is concentrated in this subscale. The experimental group's total-score SD dropped from 10.47 at T1 to 3.24 at T2, then partially recovered to 5.95 at T3. This pattern indicates that the workshop moved all participants toward the same high score at T2. The subsequent dispersion at T3 suggests that some participants retained the content and some did not.

The K/A pattern is consistent with the campaign-effectiveness literature. Bada, Sasse, and Nurse [51] argued that awareness campaigns routinely produce stronger gains in what

people know than in what people do. The present data replicates this finding under controlled experimental conditions. It also adds a temporal dimension not tested in prior work. K/A is the most responsive component but also the least durable, retaining only 39% of its gain at the two-week follow-up.

5.4.2 Behaviour: intermediate effect, moderate retention

The behaviour subscale returns a significant but smaller interaction ($\eta^2p = 0.246$). The experimental group rose from 13.76 to 17.88 out of 20 at T2, then fell to 16.17 at T3. Retention is 59% of the T1 to T2 gain.

Self-reported behaviour items ask what participants say they do. The gain at T2 is a mixture of genuine intention change and post-training self-presentation bias. Paulhus [49] showed that impression management grows stronger when respondents have a clear sense of the desirable answer. A workshop creates exactly those conditions.

The 59% retention rate is higher than the 39% for K/A. This is the opposite of what a simple social-desirability account predicts. If the behaviour gain were entirely priming, it should fade faster than K/A, not slower. The retention data suggests that at least part of the behaviour gain reflects a genuine shift in reported practices. Section 5.6 examines this pattern in detail.

Anderson and Agarwal [52] reported that security behaviour intentions are shaped by drivers that do not all translate into action. The behaviour subscale in this study measures reported practices, not observed behaviour. The gap between what participants report and what they do remains unmeasured.

5.4.3 Applied scenarios: smallest effect, highest retention

The scenarios interaction ($\eta^2p = 0.020$) does not reach significance. However, the within-group T1 to T2 paired test inside the experimental group is significant at $p = 0.007$ with Cohen's $d = 0.68$. The experimental group improved on scenarios, but the control group also drifted upward, which absorbed the interaction.

The item-level analysis in Section 4.8 identified two scenario items with McNemar-significant gains in the experimental group. Q17 on URL analysis moved from 68% to

96% at T2 and held at 83% at T3. Q18 on phone vishing moved from 76% to 100% at T2 and held at 96% at T3. Both items correspond to specific workshop modules.

Q20 on public computer session management is the weakest item across all three occasions. It rose from 64% to 68% at T2 and reached 70% at T3. This item received brief coverage in the workshop. The negligible gain fits the pattern: items that received direct treatment in the session gained more than those covered abstractly.

The retention rate of 77% is the highest of the three subscales. Applied-skill items that were learned appear to be retained better than declarative knowledge items that were also learned. This is consistent with the procedural-memory literature. Skills acquired through application tend to be more resistant to decay than facts acquired through exposure.

The scenarios subscale is limited to eight binary items. The statistical power to detect a Group by Time interaction on this subscale is low. The non-significant interaction should be read alongside the significant within-group effect (d = 0.68). The effect is present within the experimental group but too small relative to the control group drift to emerge as a significant interaction.

5.4.4 The retention inversion

The most informative pattern in the data is that the immediate effect gradient and the retention gradient run in opposite directions. The subscale that shows the largest immediate effect (K/A, $\eta^2p = 0.480$) retains the smallest proportion of that effect (39%). The subscale with the smallest interaction (Scenarios, 0.020) retains the most (77%). Figure 5.2 visualises this inversion.

Three factors contribute to this pattern. First, ceiling regression: K/A reaches 94.3% of its maximum at T2 and has more room to fall than scenarios at 86.5%. Second, the type of knowledge tested: K/A items test declarative content (what participants know), which is acquired rapidly but also forgotten rapidly [50]. Scenario items test applied judgement, which is harder to acquire but more durable once established. Third, the item-level dose-response pattern: the two scenario items that gained most (Q17 and Q18) correspond to specific workshop exercises that required active participation rather than passive reception.

Bada, Sasse, and Nurse [51] documented that awareness campaigns produce stronger knowledge gains than behaviour changes. The present data replicates their finding on the immediate gradient and extends it by showing that the decay pattern reverses the immediate ordering. This extension is enabled by the follow-up wave that most prior studies lack.

5.5 Comparison with existing literature

This section positions the results against published studies. Table 5.2 is restricted to studies whose methodology allows effect magnitudes or design features to be compared directly. Studies that used purely qualitative methods or that reported only descriptive prevalence data are excluded.

Table 5.2. Comparison of the present study against published literature.

Study	Context	Design	Reported effect	Relation to present study
Parsons et al. 2014 [2]	Australian employees	HAIS-Q validation, cross-sectional	Construct validity and reliability established	Provides K/A/B framework; no intervention d
Parsons et al. 2017 [34]	Australian workforce	Two further HAIS-Q validation studies	Reliability confirmed across samples	Instrument ancestor; no effect size to compare
Kruger and Kearney 2006 [53]	International mining company	KAB prototype measurement model	Awareness levels measured; no intervention	Theoretical ancestor of HAIS-Q KAB model
Sheng et al. 2007 [54]	US general users (N = 14 per condition)	Anti-Phishing Phil game, pre-post	False-negative rate 31% to 17% post-game	Short gamified intervention; comparable duration
Bada, Sasse, Nurse 2019 [51]	Review of campaigns	Qualitative review of campaign failures	Knowledge gains exceed behaviour changes	Our gradient replicates this; retention extends it
Robbins and Robbins 2025 [36]	US K-12 school staff	Single-session pre-post, N = 201	All three KAB domains improved ($p < 0.001$, Wilcoxon)	No control group; between-group effect not comparable
Chandarman and van Niekerk 2017 [15]	South African private tertiary	Cross-sectional awareness survey	Moderate awareness; no intervention	Baseline comparison only; no effect size

Study	Context	Design	Reported effect	Relation to present study
Khan et al. 2024 [9]	Pakistani students (N = 758)	Cross-sectional survey	70% never change passwords; gaps reported	Pakistani baseline; no intervention to compare
Bukhsh et al. 2025 [10]	Pakistani users (N = 164)	Phishing susceptibility	Authority/urgency cues increase victimisation	Pakistani risk factors; no intervention effect
Ho et al. 2025 [26]	Enterprise users (field study)	Phishing training in practice	Training efficacy questioned in real conditions	Cautionary comparator for behaviour claims

Three comparisons from Table 5.2 address specific interpretive questions.

The first concerns effect-size magnitude. The between-groups Hedges' g of +2.24 at T2 is unusually large for educational research [42]. Most intervention studies report effect sizes in the 0.2 to 0.8 range. Robbins and Robbins [36] found significant improvements across all three KAB domains ($p < 0.001$) using the HAIS-Q with K-12 school staff. Their single-group pre-post design lacks a control group, which precludes direct effect-size comparison with the present study. Two factors are relevant to the magnitude observed here. A concrete workshop combined with a population that had no prior formal training produces a large baseline-to-post-test swing. The T3 effect ($g = +0.98$) is closer to the range typically reported, which suggests that the T2 magnitude is partly inflated by immediate priming.

The second concerns the Pakistani evidence base. Khan et al. [9], Bukhsh et al. [10], and Hanif and Shams [11] all document awareness gaps among Pakistani students. None reports an experimental intervention. The present study is, to the author's knowledge, the first controlled experimental evaluation of a cybersecurity workshop on Pakistani CS students. It adds a new category of evidence to the Pakistani literature.

The third concerns the knowledge-to-behaviour gradient. Bada, Sasse, and Nurse [51] documented that campaigns produce stronger knowledge gains than behaviour changes. The present study replicates that finding under controlled conditions and extends it with

retention data showing that the immediate gradient inverts at follow-up. This extension has not been reported in the prior literature to the author's knowledge.

5.6 Social desirability and the retention pattern

Self-report instruments are vulnerable to social-desirability bias [49]. In the context of this study, the concern was that a cybersecurity workshop would prime participants to give favourable self-assessments at the immediate post-test. If such priming occurred, the gains on self-report items should be inflated at T2 and should decay more rapidly at T3 than gains on items that test applied judgement.

The anticipated pattern was that the behaviour subscale, which asks participants what they do, would show the largest post-training inflation and the steepest decay. The applied-scenarios subscale, which tests what participants can judge under realistic conditions, was expected to be more resistant to this effect.

The data does not follow the predicted pattern in its specific form. The subscale that decays most is not behaviour (59% retention) but knowledge and attitude (39% retention). Behaviour retains a larger proportion of its gain than K/A. Scenarios retain the most (77%).

The reversal has three plausible explanations that are not mutually exclusive. First, K/A items reach 94.3% of their subscale maximum at T2. The steep decay is partly ceiling regression rather than a loss of genuine knowledge. Second, K/A items test declarative content that is acquired quickly but is also subject to the rapid forgetting documented by Murre and Dros [50]. Third, behaviour items measure self-reported practices that involve habitual routines. Once a participant reports adopting a new practice, the self-report tends to persist even if the practice itself is inconsistent.

The social-desirability concern is not eliminated by this reversal. It is redirected. The data suggests that the most transient component of the measured gain is declarative knowledge, not self-reported behaviour. This is the opposite of what a simple priming account predicts. It is, however, consistent with an established finding in memory research: declarative recall fades faster than procedural or habitual responses [50].

The practical conclusion is that the T2 effect size of g = +2.24 on the total score overstates the durable effect. The T3 effect size of g = +0.98 is a more conservative and more defensible estimate of what the workshop produces beyond the immediate session. The subscale decomposition provides further diagnostic information: the retained effect is concentrated in behaviour and applied scenarios rather than in declarative K/A content.

5.7 Implications and limitations

This section merges the practical implications with the study's limitations. The two are presented together because the limitations determine how far each implication can be pushed.

5.7.1 Implications for curriculum design

A single 90-minute workshop produces a large measured gain on the combined total score (g = +2.24 at T2, g = +0.98 at T3). This gain is detectable at both the immediate post-test and the two-week follow-up. For curriculum designers deciding whether to invest in any formal intervention, the intervention bar is low and the measured return is substantial.

The subscale gradient qualifies this implication. K/A gains are the largest but also the least durable (39% retention). Applied-scenario gains are small but the most durable (77% retention). A curriculum built on single-session workshops alone would capture the K/A gains but is unlikely to shift applied judgement to the same degree.

A layered approach is more consistent with the data. A short introductory workshop establishes the knowledge base. Follow-on sessions with repetition and applied practice build procedural skills that a single session does not install. The present data supports the first layer directly. The second layer is supported by the training design literature [24], [25], [32].

5.7.2 Implications for institutional decisions

The study provides controlled evidence that a short workshop produces measurable gains in a Pakistani university population with no prior formal awareness training. The intervention is deliverable in a standard classroom slot and does not require specialised infrastructure.

The T3 effect (g = +0.98) is the more conservative basis for institutional decisions than the T2 effect (g = +2.24). Roughly half the immediate gain persists at two weeks. Whether this retention rate is sufficient to justify programme investment is a decision for institutional stakeholders, not a conclusion this thesis can draw from the data.

The differential retention pattern (K/A 39%, Behaviour 59%, Scenarios 77%) indicates that the retained effect concentrates in applied and behavioural components rather than in declarative knowledge. From an institutional perspective, these are the components that matter most for reducing actual vulnerability.

5.7.3 Implications for instructors

The item-level analysis in Section 4.8 identified two scenario items with McNemar-significant gains: Q17 on URL analysis (68% to 96% at T2, 83% at T3) and Q18 on phone vishing (76% to 100% at T2, 96% at T3). Both correspond to workshop modules that used concrete examples and required active participant responses.

Q20 on public computer session management showed negligible movement across all three occasions (64% to 68% to 70%). This item received brief and abstract treatment in the workshop. The contrast between Q17/Q18 and Q20 indicates that items receiving direct and concrete treatment produce larger and more durable gains than items covered in passing.

5.7.4 Sample limitations

The sample comprises 50 participants drawn from seven universities in Islamabad. Forty-three completed all three measurement occasions. This is adequate for the 2 x 3 mixed ANOVA but narrow in geographic and institutional terms. Students from Khyber Pakhtunkhwa, Sindh, and Balochistan are not represented.

The sample is restricted to early-year CS students. This was deliberate: early-year students provide a clean pre-intervention baseline because they have not yet completed formal security coursework. The cost is that the findings do not automatically extend to later-year students or non-CS populations.

5.7.5 Self-report limitations

The K/A and behaviour subscales use self-report Likert items. Self-report is the standard measurement mode in the HAIS-Q literature [2], [34], [37] but is vulnerable to impression management [49]. Section 5.6 examined this concern and found that the retention pattern does not follow the simple priming prediction. The social-desirability concern is not eliminated but is partially addressed by the subscale decomposition.

The applied-scenarios subscale requires participants to judge realistic situations. It is less exposed to impression management than the other two subscales. The non-significant interaction on this subscale ($\eta^2p = 0.020$) is the finding in the study least contaminated by self-report bias. It also shows the highest retention rate (77%), which suggests that what little scenario skill was acquired is relatively durable.

A stronger design would supplement self-report with direct behavioural measures such as click-through rates on simulated phishing messages. This was not feasible within the constraints of the present field study. Section 5.8 discusses the extension.

5.7.6 Design limitations

Allocation was at the cohort level rather than the individual level. Whole classes were assigned to experimental or control conditions. This was necessary for field logistics but introduces a residual risk of cohort-level confounding that individual randomisation would have eliminated [42]. The baseline equivalence checks in Section 4.3 showed no significant pre-test differences, which mitigates but does not eliminate this concern.

The parallel-form equivalence between Forms A, B, and C was established at the total-score level but is imperfect at the item level. Some control-group items drifted between forms. The total-score interaction survives this imperfection comfortably, but item-level conclusions should be read with the form-drift pattern in mind.

Form C was administered remotely. Participants completed it in their own time via an emailed link, whereas Forms A and B were completed in a supervised classroom setting. This difference in administration conditions introduces a potential source of measurement inconsistency between the T1/T2 and T3 occasions.

Attrition at T3 was 14% overall (7 of 50 participants). The rate was 8% in the experimental group (2 of 25) and 20% in the control group (5 of 25). The differential attrition is a potential source of bias. With only seven dropouts, a formal comparison of completers and non-completers lacks statistical power. The complete-case analysis retains 43 participants, which is above the minimum of 36 required at 90% power in the a priori power analysis [43].

The two-week follow-up interval captures short-term retention. It does not test long-term habit formation. Habits develop over months or years, not days. The retention wave tells whether the knowledge base survives beyond the immediate post-session window. It does not tell whether behaviour changes persist.

5.7.7 Generalisability

The generalisability claim is deliberately narrow. The findings apply to early-year CS students at Islamabad-based Pakistani universities who participated voluntarily in a 90-minute classroom workshop. They do not automatically extend to non-CS students, to other regions, to later-year students, or to different intervention formats.

The narrow scope does not make the study uninformative beyond its immediate setting. Early-year CS students are a benchmark population that is expected to be among the most receptive to cybersecurity training. If this group requires a formal intervention to produce the gains reported here, broader populations are likely to require at least as much support.

5.8 Future work

Seven directions for future research emerge from the findings and limitations of this study.

First, replication with larger and geographically broader Pakistani samples. A study drawing participants from Khyber Pakhtunkhwa, Sindh, Balochistan, and smaller cities in Punjab would test whether the workshop effect generalises beyond the Islamabad catchment. A sample of 150 to 200 participants across four provinces would provide adequate power for region-level comparisons [43].

Second, longer follow-up intervals. A design with measurement at T1, T2, 3 months, 6 months, and 12 months would distinguish session-level priming from durable learning. The retention literature indicates that the forgetting curve is steepest in the first days and then flattens [50]. A multi-point design would capture this curvature and quantify the long-run effect.

Third, direct behavioural measures. Simulated phishing messages sent to participants at controlled intervals would test whether knowledge gains translate into reduced click-through rates. Ho et al. [26] demonstrated that this is feasible at scale. Adding behavioural observation to the self-report instrument would address the limitation identified in Section 5.7.5.

Fourth, comparison of workshop duration and spacing. A design comparing one 90-minute session with three 30-minute sessions spaced over a week would test whether distributed practice produces larger or more durable effects. The distributed-practice literature predicts that the spaced version would produce stronger retention on applied items.

Fifth, comparison of workshop content emphasis. The item-level analysis suggests that items receiving direct treatment produce larger gains. A direct test would compare two workshop versions that allocate their time differently across the same topic set and measure whether the pattern replicates.

Sixth, cross-cultural replication in other South Asian contexts. India, Bangladesh, Nepal, and Sri Lanka have comparable CS student populations with different institutional backgrounds. A multi-country replication would identify which components of the effect are specific to the Pakistani context and which generalise.

Seventh, institutional dissemination and adoption. The findings of this thesis provide an empirical basis that can be presented to Pakistani higher-education institutions and policy bodies. The measured effect sizes, the subscale decomposition, and the retention data offer concrete evidence for decisions about integrating cybersecurity awareness content into CS curricula. A structured dissemination effort, presenting the results alongside the workshop materials, would test whether the evidence is sufficient to influence institutional adoption.

5.9 Closing remarks

This chapter has interpreted the findings from Chapter 4 against the research questions, the hypotheses, and the literature reviewed in Chapter 2.

All three null hypotheses are rejected. The workshop produces a large Group by Time interaction ($\eta^2p = 0.430$). The experimental group outperforms the control group at both T2 ($g = +2.24$) and T3 ($g = +0.98$). Roughly half the immediate gain persists at the two-week follow-up.

The subscale gradient shows that knowledge and attitude gains are the largest but least durable, while applied-scenario gains are the smallest but most durable. The retention inversion is consistent with the distinction between declarative and procedural knowledge and adds a dimension that prior studies in this field have not reported.

The social-desirability prediction was not confirmed in its specific form. The subscale that decays most is K/A, not self-reported behaviour. The T3 effect size ($g = +0.98$) is the more defensible estimate of the workshop's impact beyond the immediate session.

Chapter 6 draws the conclusions.

6 Conclusion

6.1 Summary of the study

This thesis investigated whether a targeted cybersecurity awareness workshop could improve threat recognition among early-year Pakistani computer science students. The study used a quasi-experimental pretest-posttest design with a non-equivalent control group and a two-week follow-up. Fifty participants were recruited from seven universities in Islamabad and allocated at cohort level to an experimental group (n = 25) or a control group (n = 25). Forty-three participants completed all three measurement occasions.

The experimental group received a 90-minute interactive workshop covering phishing, social engineering, and disinformation. The control group received a matched, content-neutral session on digital infrastructure. Both groups completed parallel forms of an instrument adapted from the HAIS-Q [2] at baseline, immediately after the session, and two weeks later. The instrument measured knowledge, attitude, self-reported behaviour, and applied threat-recognition through eight scenario items.

6.2 Principal findings

All three null hypotheses are rejected. The workshop produced a large Group by Time interaction on the combined total score ($\eta^2p = 0.430$). H0 is rejected and H1 is supported. The experimental group rose from 47.72 at baseline to 62.52 at the immediate post-test and declined to 54.74 at the two-week follow-up. The control group remained flat across all three occasions.

The between-group comparison at the immediate post-test returned Hedges' g = +2.24. H0 is rejected and H2 is supported. At the two-week follow-up, the between-group comparison returned Hedges' g = +0.98. H0 is rejected and H3 is supported. The experimental group retains a large advantage over the control group two weeks after the intervention. Approximately 47% of the immediate gain persisted at follow-up.

The effect was not uniform across subscales. Knowledge and attitude gains were largest ($\eta^2p = 0.480$). Behaviour gains were intermediate (0.246). Applied scenario gains were

smallest (0.020) and did not reach statistical significance. This gradient indicates that the workshop shifted declarative knowledge more than applied threat-recognition skill.

The retention pattern across subscales inverted the magnitude pattern. Knowledge and attitude gains retained 39% of the immediate effect at follow-up. Behaviour gains retained 59%. Scenario gains retained 77%. The components that gained the most in the short term lost the most over two weeks. The components that gained the least retained the most. This retention inversion is consistent with the distinction between declarative and procedural knowledge.

Item-level analysis showed that gains concentrated on items the workshop covered most directly. URL analysis rose from 68% to 96% at the post-test and remained at 83% at follow-up. Phone vishing rose from 76% to 100% and remained at 96%. Items that received less direct coverage showed smaller and less durable gains.

6.3 Contribution and closing remarks

The study makes three contributions. First, it provides an empirical baseline of cybersecurity awareness among early-year Pakistani computer science students drawn from multiple institutions. Second, it reports a controlled experimental test of a short training intervention in that setting, with effect sizes at both post-test and follow-up. Third, it identifies a retention inversion across subscales that prior studies in this population had not documented.

The findings are subject to the limitations discussed in Section 5.7. The sample is drawn from a single city. Allocation was at cohort level rather than individual randomisation. Self-report items are susceptible to social desirability bias, particularly at the immediate post-test. The two-week follow-up tests short-term retention rather than long-term behavioural change.

The workshop produced a large and statistically significant effect on overall awareness that partially persisted at follow-up. The subscale gradient and the retention inversion indicate that declarative gains are easier to produce but harder to sustain, while applied skill gains are harder to produce but more durable. These patterns are relevant to the design of future cybersecurity training in Pakistani higher education.

References

[1] Verizon, “2024 Data Breach Investigations Report,” 2024. [Online]. Available: https://www.verizon.com/business/resources/reports/dbir/

[2] K. Parsons, A. McCormac, M. Butavicius, M. Pattinson, and C. Jerram, “Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q),” Computers & Security, vol. 42, pp. 165–176, 2014. https://doi.org/10.1016/j.cose.2013.12.003

[3] S. Nifakos, K. Chandramouli, C. K. Nikolaou, P. Papachristou, S. Koch, E. Panaousis, and S. Bonacina, “Influence of human factors on cyber security within healthcare organisations: A systematic review,” Sensors, vol. 21, no. 15, p. 5119, 2021. https://doi.org/10.3390/s21155119

[4] DataReportal, “Digital 2025: Pakistan,” Jan. 2025. [Online]. Available: https://datareportal.com/reports/digital-2025-pakistan

[5] A. U. A. Khan, “Cyber Crime in Pakistan: Trends, Challenges, and Legal Responses,” Advance Social Science Archive Journal, vol. 4, no. 1, pp. 1358–1366, 2025. https://doi.org/10.55966/assaj.2025.4.1.080

[6] “Cyber threats increased by 17% in 2023,” The Express Tribune, Feb. 2024. [Online]. Available: https://tribune.com.pk/story/2457021/cyber-threats-increased-by-17-in-2023

[7] “Cybercrime statistics 2023: trends and threats,” The News International (Pakistan), 2023. [Online]. Available: https://www.thenews.com.pk/tns/detail/1058269-cybercrime-statistics-2023-trends-and-threats

[8] S. Kausar and A. R. Laghari, “Social Engineering Attacks in Pakistan: Analyzing the Weakest Link in Cyber Security,” Journal of Development and Social Sciences, vol. 6, no. 1, pp. 364–374, 2025. https://doi.org/10.47205/jdss.2025(6-I)32

[9] N. F. Khan, N. Ikram, and S. Saleem, “Effects of socioeconomic and digital inequalities on cybersecurity in a developing country,” Security Journal, vol. 37, pp. 214–244, 2024. https://doi.org/10.1057/s41284-023-00375-4

[10] J. A. Bukhsh, M. Daneva, and M. van Sinderen, “Exploring user risk factors and target groups for phishing victimization in Pakistan,” arXiv:2510.09249, Oct. 2025. https://doi.org/10.48550/arXiv.2510.09249

[11] F. Hanif and R. Shams, “Evaluation of Cybersecurity awareness level: A Case Study of Students of Sir Syed University of Engineering and Technology,” in Proc. 4th IEEE International Conference on AI in Cybersecurity (ICAIC), Houston, USA, Feb. 2025. doi: 10.1109/ICAIC63015.2025.10848797

[12] F. E. Catota, M. G. Morgan, and D. C. Sicker, “Cybersecurity education in a developing nation: the Ecuadorian environment,” Journal of Cybersecurity, vol. 5, no. 1, tyz001, 2019. https://doi.org/10.1093/cybsec/tyz001

[13] W. Crumpler and J. A. Lewis, “The Cybersecurity Workforce Gap,” Center for Strategic and International Studies (CSIS), Washington D.C., Jan. 2019. [Online]. Available: https://www.csis.org/analysis/cybersecurity-workforce-gap

[14] A. Abdulla, S. Salih, and H. Abdulrahman, “Analysis of social engineering awareness among students and lecturers,” IEEE Access, vol. 11, pp. 96701–96714, 2023. https://doi.org/10.1109/ACCESS.2023.3311708

[15] R. Chandarman and B. van Niekerk, “Students’ cybersecurity awareness at a private tertiary educational institution,” The African Journal of Information and Communication, no. 20, pp. 133–155, 2017. https://doi.org/10.23962/10539/23572

[16] B. Kitchenham and S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” EBSE Technical Report EBSE-2007-01, Keele University, 2007.

[17] M. J. Page et al., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” BMJ, vol. 372, n71, 2021.

[18] V. Zimmermann and K. Renaud, “Moving from a ‘human-as-problem’ to a ‘human-as-solution’ cybersecurity mindset,” International Journal of Human-Computer Studies, vol. 131, pp. 169–187, 2019.

[19] T. Rahman, R. Rohan, D. Pal, and P. Kanthamanon, “Human Factors in Cybersecurity: A Scoping Review,” in Proc. 12th AITCF, 2021. https://doi.org/10.1145/3468784.3468789

[20] A. Pollini et al., “Leveraging human factors in cybersecurity: an integrated methodological approach,” Cognition, Technology & Work, vol. 24, pp. 371–390, 2022. https://doi.org/10.1007/s10111-021-00683-y

[21] D. M. Sarno and J. Black, “Who gets caught in the web of lies?,” Applied Cognitive Psychology, vol. 38, no. 1, e4070, 2024. https://doi.org/10.1177/00187208231173263

[22] Teodor Sommestad, Henrik Karlzén, The unpredictability of phishing susceptibility: results from a repeated measures experiment, Journal of Cybersecurity, Volume 10, Issue 1, 2024, tyae021, https://doi.org/10.1093/cybsec/tyae021

[23] R. Yang et al., “Predicting User Susceptibility to Phishing Based on Multidimensional Features” Computational Intelligence and Neuroscience, vol. 2022, 7058972, 2022. https://doi.org/10.1155/2022/7058972

[24] D. Jampen, G. Gür, T. Sutter, and B. Tellenbach, “Don’t click: towards an effective anti-phishing training,” Human-centric Computing and Information Sciences, vol. 10, no. 33, 2020. https://doi.org/10.1186/s13673-020-00237-7

[25] O. Sarker, A. Jayatilaka, S. Haggag, C. Liu, and M. A. Babar, "A multi-vocal literature review on challenges and critical success factors of phishing education," Journal of Systems and Software, vol. 208, 111899, 2024. https://doi.org/10.1016/j.jss.2023.111899

[26] G. Ho et al., "Understanding the Efficacy of Phishing Training in Practice," in Proc. 2025 IEEE Symposium on Security and Privacy (SP), 2025.

[27] Huang, G., Jia, W., & Yu, W. (2024). Media literacy interventions improve resilience to misinformation: A meta-analytic investigation of overall effect and moderating factors. Communication Research. https://doi.org/10.1177/00936502241288103

[28] A. Guess et al., "A digital media literacy intervention increases discernment between mainstream and false news in the United States and India," PNAS, vol. 117, no. 27, pp. 15536–15545, 2020. https://doi.org/10.1073/pnas.1920498117

[29] Dame Adjin-Tettey, T. (2022). Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. *Cogent Arts & Humanities*, *9*(1). https://doi.org/10.1080/23311983.2022.2037229

[30] K. Khando, S. Gao, S. M. Islam, and A. Salman, "Enhancing employees information security awareness in private and public organisations: A systematic literature review," Computers & Security, vol. 106, 102267, 2021. https://doi.org/10.1016/j.cose.2021.102267

[31] O. Pahlavanpour and S. Gao, "A Systematic Mapping Study on Gamification within Information Security Awareness Programs," Heliyon, vol. 10, no. 19, e38474, 2024. https://doi.org/10.1016/j.heliyon.2024.e38474

[32] A.K. Gwenhure and F.S. Rahayu, "Gamification of Cybersecurity Awareness for Non-IT Professionals: A Systematic Literature Review" International Journal of Serious Games, vol. 11, no. 1, pp. 83–99, 2024. https://doi.org/10.17083/ijsg.v11i1.719

[33] Weeratunge, N.H., Hjelsvold, R. & Katt, B. A systematic scoping review on game-based cybersecurity awareness education. *Int. J. Inf. Secur.* **25**, 71 (2026). https://doi.org/10.1007/s10207-026-01239-9

[34] K. Parsons et al., "The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies," Computers & Security, vol. 66, pp. 40–51, 2017. https://doi.org/10.1016/j.cose.2017.01.004

[35] A. McCormac et al., "A Reliable Measure of Information Security Awareness and the Identification of Bias in Responses" Australasian Journal of Information Systems, vol. 21, 2017. https://doi.org/10.3127/ajis.v21i0.1697

[36] M. S. Robbins and C. Robbins, "Impact of Information Security Awareness Training on Knowledge, Attitude, and Behavior: A K-12 Case Study" Journal of Cybersecurity Education, Research and Practice, vol. 2025, no. 1, art. 20, 2025. Available here

[37] R. Rohan et al., “A systematic literature review of cybersecurity scales assessing information security awareness,” Heliyon, vol. 9, no. 3, e14234, 2023. https://doi.org/10.1016/j.heliyon.2023.e14234

[38] V. Švábenský, J. Vykopal, and P. Čeleda, “What Are Cybersecurity Education Papers About?,” in Proc. ACM SIGCSE ’20, 2020, pp. 2–8. https://doi.org/10.1145/3328778.3366816

[39] B. Ahamed et al., “Empowering Students for Cybersecurity Awareness Management in the Emerging Digital Era,” SAGE Open, vol. 14, no. 1, 2024. https://doi.org/10.1177/21582440241228920

[40] R. B. Cialdini, Influence: Science and Practice, 5th ed. Boston: Pearson, 2009.

[41] M. Sailer, J. U. Hense, S. K. Mayr, and H. Mandl, “How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction,” Computers in Human Behavior, vol. 69, pp. 371–380, 2017. https://doi.org/10.1016/j.chb.2016.12.033

[42] W. R. Shadish, T. D. Cook, and D. T. Campbell, Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA, USA: Houghton Mifflin, 2002.

[43] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, “G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” Behavior Research Methods, vol. 39, no. 2, pp. 175–191, 2007, doi: 10.3758/BF03193146.

[44] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988.

[45] A. Field, Discovering Statistics Using IBM SPSS Statistics, 5th ed. London, UK: SAGE Publications, 2018.

[46] Tallinn University of Technology, “Principles of Academic Ethics (Code of Academic Ethics),” TTU Council Resolution No 44, 17 October 2017. [Online]. Available: https://oigusaktid.taltech.ee/en/principles-of-academic-ethics-code-of-academic-ethics/

[47] Estonian Research Council, “Estonian Code of Conduct for Research Integrity,” Tartu, Estonia, 2017. [Online]. Available: https://www.eetika.ee/en/content/estonian-code-conduct-research-integrity

[48] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," International Journal of Medical Education, vol. 2, pp. 53-55, 2011. doi: 10.5116/ijme.4dfb.8dfd

[49] D. L. Paulhus, "Two-component models of socially desirable responding," Journal of Personality and Social Psychology, vol. 46, no. 3, pp. 598-609, Mar. 1984. doi: 10.1037/0022-3514.46.3.598

[50] J. M. J. Murre and J. Dros, "Replication and Analysis of Ebbinghaus' Forgetting Curve," PLoS ONE, vol. 10, no. 7, e0120644, Jul. 2015. doi: 10.1371/journal.pone.0120644

[51] M. Bada, A. M. Sasse, and J. R. C. Nurse, "Cyber Security Awareness Campaigns: Why do they fail to change behaviour?" arXiv preprint arXiv:1901.02672, 2019. https://doi.org/10.48550/arXiv.1901.02672

[52] C. L. Anderson and R. Agarwal, "Practicing Safe Computing: A Multimethod Empirical Examination of Home Computer User Security Behavioral Intentions," MIS Quarterly, vol. 34, no. 3, pp. 613-643, Sep. 2010. doi: 10.2307/25750694

[53] H. A. Kruger and W. D. Kearney, "A prototype for assessing information security awareness," Computers & Security, vol. 25, no. 4, pp. 289-296, Jun. 2006. doi: 10.1016/j.cose.2006.02.008

[54] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish," in Proc. 3rd Symposium on Usable Privacy and Security (SOUPS '07), Pittsburgh, PA, USA, Jul. 2007, pp. 88-99. doi: 10.1145/1280680.1280692

[55] Government of Pakistan, "Prevention of Electronic Crimes Act, 2016," Pakistan Code, 2016.

[56] International Telecommunication Union, "Global Cybersecurity Index 2024," ITU, Geneva, 2024.

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I, Haider Rehman

1 grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis The Importance and Effectiveness of Cybersecurity Awareness Training for Pakistani Students, supervised by Ricardo G. Lugo

1.1 to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;

1.2 to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.

2 I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.

3 I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

Haider Rehman

[1] The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.

Appendix 2: Form A Questionnaire (T1, Baseline)

Form A was the baseline (T1) instrument administered to both the experimental and control groups before the workshop. It contains a participant information sheet, demographic items (Section A), eight knowledge and attitude statements (Section B Part 1), four behaviour-frequency items (Section B Part 2), and eight applied scenarios (Section C). The instrument was delivered via a Google Form.

Participant Information Sheet

About the Study

You are invited to participate in a research study examining digital skills and technical competency among Computer Science students. The study is being conducted at Tallinn University of Technology, Estonia. Your participation involves completing three short questionnaires over a period of approximately two weeks, and attending one session on digital skills. Each questionnaire takes around 12 to 15 minutes to complete.

How Your Responses Will Be Used

All responses are completely confidential. Your individual answers will never be shared with anyone, published, or made available in any form. Results will be presented only in aggregate (for example, "X% of participants answered option Y"), and no individual response will be identifiable in any report or publication arising from this study.

There is no scoring, marking, or grading involved. There are no right or wrong answers in the opinion-based questions, and your performance on any part of this questionnaire will not be linked to your identity in any way. Please answer all questions as honestly as you can. The value of this study depends entirely on genuine responses.

Why an Email Address is Collected

An email address is collected solely to send you links to the second and third questionnaires and for any necessary communication related to this study. You may provide any email address you are comfortable with. It does not have to be your university email. Your email address will not appear in any report, will not be shared with anyone, and will be permanently deleted once data collection is complete.

What Other Information is Collected

Your university name, age range, and gender are collected for statistical purposes only. These details will be reported only in aggregate form (for example, "60% of participants were aged 18 to 20") and will never be presented in a way that could identify any individual participant.

Your Rights

Participation is entirely voluntary. You may withdraw at any time without giving a reason, and your data will be removed from the study. By proceeding to the next page, you confirm that you have read and understood the above information and agree to participate.

Section A: Participant Information

Instructions: Please provide the following details. This information is used for communication and statistical purposes only, as described above. No personal information will be shared or published. Items marked with an asterisk (*) are required.

1. Email Address (*): Please provide any email address you are comfortable with. This will only be used to send you follow-up questionnaires and for any communication related to this study. It will not be shared with anyone or published.

---

2. Which university are you currently enrolled at? (*)

---

3. What year of your Computer Science degree are you currently in? (*)

- ○ 1st Year
- ○ 2nd Year

4. What is your age? (*)

- ○ 17 or Under
- ○ 18 to 20
- ○ 21 to 23
- ○ 24 or Older

5. What is your gender? (*)

- ○ Male

○ Female

○ Prefer not to say

6. Have you ever attended a dedicated cybersecurity workshop, course, or training programme (outside of your regular university coursework)? (*)

○ Yes

○ No

○ Maybe / Do not remember

7. Does your current university coursework include any module or subject specifically focused on cybersecurity? (*)

○ Yes

○ No

○ Maybe / Not sure

Section B – Part 1: Knowledge and Attitudes

Please read each statement carefully and indicate how much you agree or disagree. There are no right or wrong answers. We are interested in your honest opinion.

Scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

1. Using the same password for my email, social media, and university portal would not put my accounts at additional risk.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

2. Creating a different password for each of my accounts is worth the extra effort.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

3. If an email comes from someone I recognise, it is safe to open any attachments or links in it.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

4. I feel comfortable clicking on links in emails without checking where they lead.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

5. Downloading software from unofficial websites carries no more risk than downloading from official sources.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

6. I see no problem with using free, unlicensed versions of paid software if they are available online.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

7. Sharing my daily routine, travel plans, or current location on social media is harmless.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

8. Connecting to a free, open wireless network in a cafe or airport is just as safe as using my home network.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

Section B – Part 2: Self-Reported Behaviour

Please read each statement and indicate how often this applies to you.

Scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always.

9. I use a different password for each of my important online accounts.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

10. I check the sender's email address carefully when I receive an unexpected message asking me to take action.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

11. I review who can see my personal information on my social media profiles.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

12. When I come across a surprising news story on social media, I look it up on a separate source before accepting it.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

Section C: Applied Scenarios

The following scenarios describe situations you might encounter online. For each one, select the answer you believe is most appropriate.

13. You receive the following email:

From: security-alert@paypa1.com

Subject: Unusual Login Detected — Immediate Verification Required

"Dear User, We detected an unusual sign-in to your account from an unrecognised device. To secure your account, please verify your identity immediately by clicking the link below: [Verify My Account Now]. If you do not verify within 24 hours, your account will be suspended. PayPal Security Team."

What is the MOST suspicious element in this email?

- ○ The email mentions an unusual login from an unrecognised device.
- ○ The sender's email address does not match the official domain of the company.
- ○ The email contains a clickable link.
- ○ The email threatens to suspend the account.

14. A classmate sends you a WhatsApp message: "Hey, I found this amazing website where you can get free premium Spotify accounts. Just enter your university email and password to claim yours: www.free-spotify-premium-accounts.tk".

What should you do?

- ○ Avoid entering any credentials as the site may not be trustworthy.
- ○ Enter your university credentials since a friend sent the link and it must be safe.
- ○ Check if the website has a padlock icon in the browser, and if it does, it is safe to proceed.
- ○ Try the link using a different password than your usual one.

15. You come across a news article shared on Facebook with the headline: "BREAKING: Government Announces All University Degrees Will Be Cancelled After 2026." The article is from a website called "pakistan-news-real.blogspot.com" and has 3,000 shares.

What is the best response?

- ○ Share it immediately with classmates so they are aware.

- Assume it must be true because it has thousands of shares.
- Check whether any established news outlet has reported the same story.
- Ignore it entirely without any concern since it does not affect you.

16. You are browsing the internet and a pop-up appears: "WARNING! Your device has been infected with 3 viruses. Click here to scan and remove them immediately!" The pop-up looks professional and includes a well-known antivirus company's logo.

What should you do?

- Click the button to run the scan since it displays a recognised antivirus brand.
- Close the pop-up without clicking on anything inside it.
- Download the suggested software as a precaution.
- Right-click the pop-up to check if it is a real system notification.

17. Which of the following URLs is MOST likely to be a legitimate link to a university's student portal?

- http://portal-university.edu.pk.account-verify.com/login
- https://portal.university.edu.pk/login
- https://university-portal.freesites.net/student-login
- https://www.university.edu.pk.portal-login.org/auth

18. You receive a phone call from someone claiming to be from your mobile network provider. They say there is a problem with your account and they need your CNIC number and account PIN to fix it. They sound professional and know your name.

What is the safest course of action?

- Provide the information since they already know your name, which means they are likely genuine.
- Give them your CNIC but not your PIN as a compromise.
- Ask them to verify their identity by telling you your account balance first.
- Hang up and contact your provider through the helpline or app.

19. You want to install a video editing app on your Android phone. You find it on a third-party website instead of the Google Play Store. The website asks you to enable "Install

from Unknown Sources" in your phone settings. The app has positive reviews on the website itself.

What is the risk?

- ○ There is no additional risk since the app has positive reviews on the website.
- ○ The risk only applies to older Android versions, not newer ones.
- ○ The app could contain harmful software.
- ○ The app might not receive automatic updates and malfunction.

20. You are using a public computer at a university library to check your email. After you finish, you close the browser window and walk away. A few minutes later, you realise you may have forgotten something.

What is the most likely security concern?

- ○ The computer may have recorded your keystrokes and stored your password.
- ○ Closing the window may not have ended your session.
- ○ The library’s network may have stored your credentials and shared them with administrators.
- ○ There is no concern since university library computers are secured.

Appendix 3: Form B Questionnaire (T2, Immediate Post-Test)

Form B was administered immediately after the workshop session for the experimental group, and immediately after the lecture-only control session for the control group. It is a parallel form of the HAIS-Q designed to measure the same constructs (knowledge or attitudes, behaviour, and applied judgement) using different item wording and different scenarios to reduce testing effects. The instrument was delivered via a Google Form.

As a reminder to participants, responses are completely confidential and will only be presented in aggregate form. No individual answers will be identifiable. There is no scoring or marking; participants are asked to answer honestly.

Section A: Participant Information

1. Email Address (*): Please enter the same email address you provided in the first questionnaire so that your responses can be matched across all questionnaires. This email will only be used for sending further questionnaires or any communication related to this study, and will be permanently deleted after data collection is complete.

Section B – Part 1: Knowledge and Attitudes

Please read each statement carefully and indicate how much you agree or disagree. There are no right or wrong answers. We are interested in your honest opinion.

Scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

1. A password made of a common word followed by a number is difficult for automated tools to guess.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

2. I would rather keep my passwords simple and easy to remember than make them complex and hard to recall.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

3. It is possible for someone to send an email that appears to come from a trusted person or organisation but is actually from a stranger.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

4. If I receive an email with an attachment labelled with something relevant to me, such as exam results, I feel comfortable opening it without verifying who sent it.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

5. A padlock icon in the browser address bar guarantees that the website is trustworthy and its content is safe.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

6. I do not think or review a lot about granting permissions when installing a new app or browser extension.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

7. Online quizzes or games (such as horoscope-style challenges) that ask personal questions such as my mother's name or school are just harmless entertainment.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

8. An application that requests access to my contacts, messages, and camera is always just using all of that information to function properly.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

Section B – Part 2: Self-Reported Behaviour

Please read each statement and indicate how often this applies to you.

Scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always.

9. I update my passwords when I hear that a service I use has been compromised.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

10. When I receive an unexpected email asking me to download a file, I confirm with the sender through a separate channel before opening it.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

11. I accept connection or friend requests from people I do not personally know.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

12. When I see a screenshot of a news headline shared on social media without a link to the full article, I search for the original story independently.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

Section C: Applied Scenarios

The following scenarios describe situations you might encounter online. For each one, select the answer you believe is most appropriate.

13. You receive the following email:

From: noreply@github-security.net

Subject: Suspicious Activity on Your GitHub Account

"Dear Developer, We have detected a login from a new IP address. If this was not you, please secure your account immediately using the link below: [Secure My Account]. Failure to act within 12 hours may result in permanent account restriction. GitHub Support Team."

You decide this email might not be genuine. What is the BEST first step to verify?

- ○ Search online to check if other people have received similar emails recently.
- ○ Reply to the email and ask for more information.
- ○ Log into your GitHub account by typing github.com directly into your browser.
- ○ Forward the email to a friend who knows about technology and ask their opinion.

14. You receive a direct message on LinkedIn from someone claiming to be a recruiter from a well-known tech company you are aware of. They say they found your profile impressive and would like to offer you an internship. They ask you to fill out a Google Form with your personal details, CNIC, and bank account number for salary in order to begin processing the offer.

What should you do?

- ○ Fill out the form since LinkedIn is a legitimate and secure platform.
- ○ Fill it out but provide an alternate bank account you do not use frequently.
- ○ Do not share personal or financial details and visit the official careers site to verify.

- Ask the recruiter to send the form from an official company email first, then decide.

15. You see a viral post on X (Twitter) with an image that appears to show an official government document announcing emergency changes to student loan policy. The document has an official-looking letterhead and stamp. The account sharing it is only 2 months old and has 400 followers.

What is the best approach?

- Trust it because the document has an official letterhead and government stamp.
- Search for the announcement on official government channels.
- Share it with a disclaimer like "not sure if this is real" to let others decide for themselves.
- Report the account immediately without checking whether the information is accurate.

16. You search for a popular paid software (for example, Adobe Photoshop) and find a website offering it for free. The site looks professional and claims the download is "100% safe and virus-free, verified by our team." The download button is large and prominent, but you notice several other smaller buttons on the page that also say "Download" in different places.

What is the most likely risk?

- The software is probably legitimate, and ads are the reason for small buttons since it is free for users.
- There is no risk as long as you click the largest and most prominent download button.
- The website has verified the file is safe, so any of the download buttons should be fine.
- The actual file may be bundled with unwanted or harmful software.

17. Which of the following URLs is MOST likely to lead to a fake website?

- https://www.hec.gov.pk/scholarships
- https://hec-gov-pk-scholarships.web.app/apply
- https://hec.gov.pk/english/services/students

- https://www.hec.gov.pk/students/login

18. Someone calls you claiming to be from JazzCash customer support. They say your account has been flagged for unusual activity and they need your PIN and account number to "unblock" it immediately. When you hesitate, they say you can verify by checking your last transaction, and they read out a recent transaction amount correctly.

What is the safest response?

- Provide the information since they correctly identified your recent transaction.
- Give them only the account number but not the PIN as a precaution.
- Hang up and contact JazzCash directly through their helpline or app.
- Ask them more questions related to your account in order to verify.

19. You want to use a free VPN app to access restricted content. You download one from the Play Store that has 4.5 stars, 10 million downloads, and many positive reviews. The privacy policy states that the app may collect and share browsing data with third-party advertisers.

What is the primary concern?

- Free apps with high ratings and millions of downloads are verified safe by Google.
- The VPN may be collecting and sharing your browsing activity with third parties, which undermines the purpose of using it.
- It might slow down your internet speed slightly.
- The privacy policy warning is standard legal language that all apps include and does not indicate any real risk.

20. A classmate asks to borrow your laptop for a few minutes to "print a document." You agree and step away briefly. When you return, you notice they plugged in a USB drive. Everything looks normal, but you see a folder on your desktop that you did not create.

What is the most appropriate course of action?

- Open the folder to check what your classmate copied or left behind.

- Delete the folder as your classmate might have forgotten to delete the documents after printing.
- Run a full security scan on your laptop and delete the folder.
- Keep the folder as your classmate might need it later and it might contain important files.

Appendix 4: Form C Questionnaire (T3, Two-Week Follow-Up)

Form C was administered approximately two weeks after the workshop to assess retention of training effects. Like Form B, it is a parallel form of the HAIS-Q measuring the same three constructs through different item wording and different scenarios. Participants are asked, for statements about habits and practices, to respond based on what they currently do in their daily life. The instrument was delivered via a Google Form.

As a reminder to participants, responses are completely confidential and will only be presented in aggregate form. No individual answers will be identifiable. There is no scoring or marking; participants are asked to answer honestly.

Section A: Participant Information

1. Email Address (*): Please enter the same email address you provided in the first questionnaire so that your responses can be matched across all questionnaires. This email will only be used for sending further questionnaires or any communication related to this study, and will be permanently deleted after data collection is complete.

Section B – Part 1: Knowledge and Attitudes

Please read each statement carefully and indicate how much you agree or disagree. There are no right or wrong answers. We are interested in your honest opinion.

Scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

1. If one of my online accounts is broken into, my other accounts are not affected as long as they are on different websites.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

2. Writing my passwords in a notebook or a note on my phone is a reasonable way to keep track of them.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

3. An email that pressures me to act immediately is more likely to be genuine than one that gives me time to think.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

4. I would rather click the link in the email than open a new tab and type my bank's address myself.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

5. Once I have installed a browser extension and it works fine, I do not need to worry about it becoming unsafe later.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

6. Using private or incognito mode hides my browsing from my internet provider and the sites I visit.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

7. Shortened links shared by acquaintances on social media are safe to click.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

8. As long as my phone still works smoothly, postponing system updates for a few weeks is fine.

Strongly Disagree ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Strongly Agree*

Section B – Part 2: Self-Reported Behaviour

Please read each statement and indicate how often this applies to you.

Scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always.

9. I use passwords that are at least eight characters long and include a mix of letters, numbers, and symbols.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

10. When I get a link in an email, I hover over it or long-press on mobile to see the address before clicking.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

11. I think carefully about what personal details I am revealing before I post something online.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

12. Before forming an opinion on a controversial topic in the news, I look for coverage of the same story from more than one source.

Never ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 *Always*

Section C: Applied Scenarios

The following scenarios describe situations you might encounter online. For each one, select the answer you believe is most appropriate.

13. You receive the following email:

From: support@microsoft-account-verify.com

Subject: Your Microsoft 365 Subscription Is About to Expire

"Dear Student, Your university Microsoft 365 subscription will expire within 48 hours. To avoid losing access to your email and OneDrive files, please renew your subscription by confirming your university login credentials here: [Confirm My Account]. Microsoft Support."

What is the MOST suspicious element in this email?

- ○ Microsoft 365 subscriptions can expire, so the content seems routine.
- ○ The email is addressed to "Dear Student" instead of using your name.
- ○ The email has a clickable link.
- ○ The sender’s email address is not from an official domain.

14. A classmate sends you a WhatsApp message: "Bro I found a working trick. You can get ChatGPT Plus free for a year through this student promo. You just sign in with your university email on this site and it activates automatically. I already did it, works perfectly: CLICK LINK".

What should you do?

- ○ Use the link but sign in with a throwaway email you created just for this.
- ○ Do not use the link and ignore the message.

- Ask your classmate to send you screenshots of their activated account first to confirm it works.
- Use an incognito window to open the link so you can test it safely.

15. You see a WhatsApp forward claiming that a popular food brand's product has been found to contain a dangerous chemical. The message includes a photo of a leaked lab report and says "FORWARD THIS TO EVERYONE YOU KNOW." Several family members have already forwarded it to you.

What is the most appropriate response?

- Forward it to your family group chat immediately since it concerns health and safety.
- Trust it because the lab report image looks official and multiple family members sent it.
- Check whether food safety authorities or established media have reported the claim.
- Reply to your family members confirming you received it and will stop buying the product.

16. You find an online store selling a popular smartphone at 30% below the market price. The website looks professional and has product photos and customer reviews. However, the only payment option is a bank transfer.

What should you do?

- Proceed with the bank transfer since the website looks professional and has customer reviews.
- Avoid purchasing from this site.
- Buy a small item from the store first to check its legitimacy before investing a lot of money.
- Ask the store to ship the product first, against a minimal advance, with proof of delivery, and then send the money.

17. You received four emails related to different issues at university on your university email. Which of the following email addresses is MOST likely to be someone impersonating a university department?

- it.support@university-edu.pk
- services-helpdesk@university.edu.pk
- itservices@university.edu.pk
- noreply.it@university.edu.pk

18. You receive a phone call. The caller speaks politely in Urdu and says: "Assalam o Alaikum, I am calling from NADRA's Islamabad office regarding a discrepancy in your CNIC record. To protect your identity from misuse, I need to verify your CNIC number, date of birth, and the mobile number registered against your CNIC. This is a routine verification and will take less than two minutes."

What is the safest course of action?

- Provide the details since NADRA is a legitimate government organisation and identity verification is reasonable.
- Provide only your CNIC number as it is a publicly available detail but refuse to share any other details.
- Call NADRA's helpline yourself to verify the issue, and end the call.
- Ask the caller for their employee ID and office extension so you can confirm they are genuine before sharing anything.

19. You are connected to the free Wi-Fi at a coffee shop and decide to log into your online banking app. The Wi-Fi network does not require a password to connect. The banking app opens normally and looks the same as always.

What is the most likely risk in this situation?

- There is no risk since the banking app has its own security and encryption.
- Other users on the same network could potentially intercept data being sent between your device and the internet.
- The risk only exists if you are using a laptop (such as the bank website), not a smartphone app.
- The coffee shop owner can see your bank balance but not your password.

20. A classmate recommends a new AI study assistant called StudyBuddy AI at studybuddy-ai.com. You visit the site and click "Sign in with Google." You are

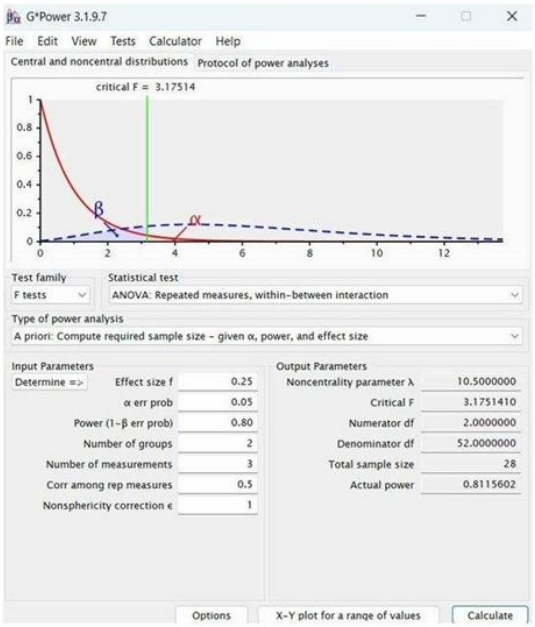
redirected to Google’s genuine login page at accounts.google.com and successfully enter your credentials. Google then shows you the following consent screen:

StudyBuddy AI wants to access your Google Account. This will allow StudyBuddy AI to: see your primary Google Account email address; see your personal info, including any personal info you have made publicly available; read, compose, and send email from Gmail; see, edit, and create Google Drive files. [Cancel] [Allow].

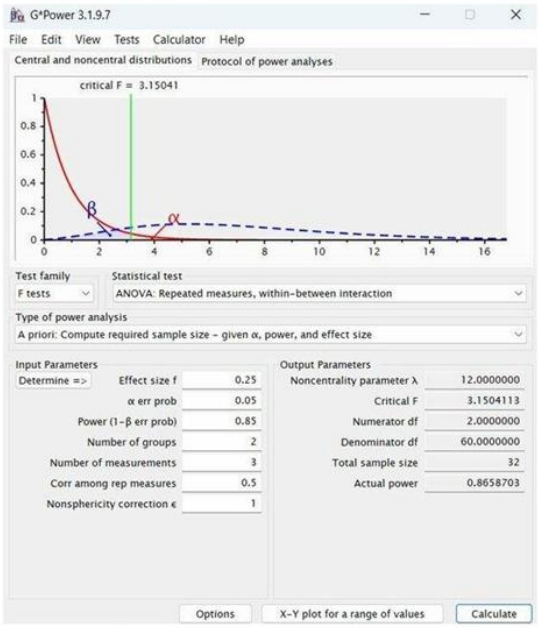
You click "Allow" because you logged in through Google’s real page and you want to try the tool. What is the most likely risk?

- StudyBuddy AI may have silently captured your Google password while you typed it on the login page.
- There is no significant risk, because you signed in through Google’s real login page rather than a fake one, and Google will block anything suspicious the app tries to do.
- StudyBuddy AI will only be able to access your data while the browser tab is open, and the permission ends automatically once you close the tab or log out of the site.
- StudyBuddy AI now has ongoing permission to read and modify your email and Drive files, until you revoke it.

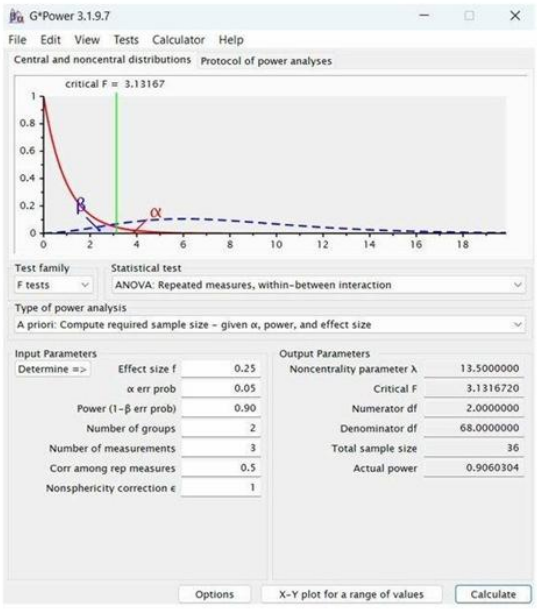
Appendix 5: G*Power Output



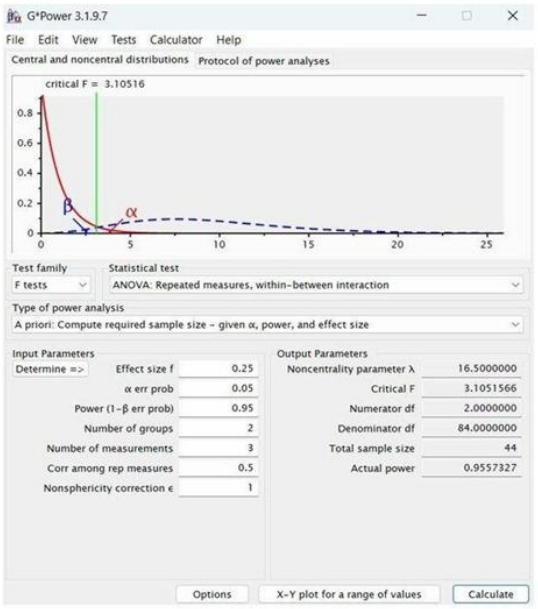
(a) Power = 0.80, required N = 28



(b) Power = 0.85, required N = 32



(c) Power = 0.90, required N = 36



(d) Power = 0.95, required N = 44

G*Power a-priori sample size output at four power levels (80%, 85%, 90%, 95%). Parameters: f = 0.25, α = 0.05, two groups, three measurements, ρ = 0.5, ε = 1.