UTRECHT UNIVERSITY

MASTER THESIS

A Preliminary Framework for Assessing Information Security Mindsets in An Organisational Setting

Author:
Tjerk de Haan

1st Supervisor:
dr. Katsiaryna Labunets

2nd Supervisor:
dr. Slinger Jansen

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN

Business Informatics
Department of Information and Computer Sciences

May 20, 2025

Abstract

In the context of increasing cyberthreats and the widespread use of shadow IT, organisations increasingly recognise the need to address the human dimension of cybersecurity. Despite the growing investment in Security Education Training and Awareness (SETA) programmes, many initiatives do not achieve meaningful behavioural change among employees. This study investigates how organisations can assess and categorise prevalent employee Mental Models (MMs) toward Shadow IT usage, to enhance cybersecurity culture and tailor awareness efforts more effectively.

Drawing on the concept of MMs, internal representations that shape an individual's reasoning and decisions, this research develops a novel framework that combines theoretical taxonomies with practical assessment strategies. A narrative literature review and semi-structured interviews with cybersecurity experts and researchers informed the framework. The study focuses mainly on extending traditional security taxonomies by obtaining metrics related to employee Knowledge, Attitudes, and Behaviour (KAB) dimensions through various elicitation approaches.

The findings highlight that inaccurate or incomplete MMs can lead to risky behaviour, such as the unintentional use of unauthorised technologies. However, when identified and categorised, these models offer valuable insight into how and why individuals perceive security threats. The proposed framework enables organisations to systematically measure subjective and objective proxy indicators, providing a more holistic and data-informed approach to cybersecurity awareness.

This thesis contributes to the academic discourse on human-centred cybersecurity by proposing a model for assessing MM. In addition, it offers practical implications for organisations seeking to prioritise security training efforts and evaluate the effectiveness of SETA programmes based on measurable behavioural and cognitive outcomes.

Keywords: Shadow IT, Mental Model, security behaviour, proxy metrics, information security, SETA, KAB, ISA

Contents

1 Introduction

Digital transformation has become a strategic necessity for modern organisations, driven by advances in information and communication technologies and the demand for operational agility and scalability (Ahmad, 2014; Kraus et al., 2019). As businesses increasingly rely on complex digital infrastructures to deliver value and maintain competitiveness, the risks associated with cyberthreats have also increased (Rafaiani et al., 2021).

Cybersecurity, broadly defined as the coordinated application of technologies, policies, and human behaviours to protect cyberspace and cyber-enabled systems, has emerged as a critical organisational concern (Craigen et al., 2014). Unlike traditional information security, cybersecurity extends beyond technical controls to include human factors, recognising that individuals can be both targets and non-knowing participants in cyberattacks (Safa et al., 2015; Von Solms and Van Niekerk, 2013).

The increasing sophistication and frequency of cyberattacks have made the issue a primary topic on global risk agendas. This increase is a direct result of the development of intelligent technology, digitisation, and globalisation (Sharif and Mohammed, 2022). In its 2024 Global Risk Report, the World Economic Forum lists cyberattacks among the top five global risks by likelihood and severity (WEF, 2024). Although cyberattacks are recognised by many as significant risks, their sheer extent might be underestimated (Bruijn and Janssen, 2017). Projections estimate that cybercrime will cause annual damages of up to $10.5 trillion annually by 2025, a 75% increase from 2021 figures (Morgan, 2022). These figures emphasise not only the scale of the threat, but also its growing economic and societal impact.

In 2021, the number of organisational security breaches increased by 27.4% compared to the previous year (Sharif and Mohammed, 2022). Both small businesses and multinational corporations have experienced major breaches, showing that no organisation is safe, regardless of size or resources (He and Z. Zhang, 2019; Sharif and Mohammed, 2022). Increasingly, organisations are targeted not just because of their technical vulnerabilities, but because of their perceived value and exposure. The consequences of such breaches include reputational damage, financial loss, legal liability, and operational disruption (Aldawood and Skinner, 2019; Farahbod et al., 2020; Sheehan et al., 2019).

In response, global cybersecurity spending is expected to exceed $215 billion in 2024, with organisations prioritising technical measures such as firewalls, access controls, and monitoring systems (Kapko, 2023; Ifinedo, 2014). However, these defences alone are insufficient. A consistent finding across cybersecurity research is that human error remains the most persistent and least predictable vulnerability (Bruijn and Janssen, 2017; Li et al., 2019).

Recognising this, organisations have begun to move toward more holistic security strategies. Such strategies incorporate Security Education, Training, and Awareness (SETA) programmes aimed at improving employee behaviour and reducing human-related risks (He and Z. Zhang, 2019; Van Steen et al., 2020). These efforts seek to foster a stronger culture of Information Security Awareness (ISA) and personal responsibility for secure behaviour.

1.1 Problem statement

Despite growing awareness of cybersecurity risks, reflected in the fact that 80% of organisations now identify it as a strategic priority, breach rates remain high (Department for Digital, 2020). Approximately 50% of businesses report incidents annually, with the figure increasing to 68% and 75% among medium and large companies, respectively. Although organisations often recognise the direct and immediate consequences of such breaches, many still underestimate or overlook the longer-term and indirect costs associated with internal vulnerability and behavioural risks (Furnell, Heyburn, et al., 2020). Such a limited vision could result in fewer investments in cybersecurity technical controls and ISA and skill training for employees.

A substantial proportion of cyber incidents originate not from external attackers, but from within the organisation. IBM's 2015 Cyber Security Intelligence Index reported that 55% of incidents had internal origins, with the majority attributed to inadvertent or negligent behaviour rather than malicious intent (IBM, 2015). Research indicates that employees do not deliberately engage in risky behaviour out of malice, but rather because they believe that their actions are justified or secure (Wash and Rader, 2011). These decisions are often based on incomplete or flawed MMs. Without a clear understanding of how these cognitive models form, evolve, and influence behaviour, organisations lack the tools to improve general ISA through interventions (Bada and Nurse, 2019; Yan et al., 2018). These findings emphasise the importance of human factors in cybersecurity, particularly among employees.

Many organisations have implemented SETA programmes to address these risks. However, such initiatives are often poorly received by employees and may have negligible or even counterproductive outcomes (Reeves et al., 2021). Consequently, human vulnerability remains one of the most persistent weak points in cybersecurity defences, especially when insider threats can be exploited by external parties (Alsharif et al., 2022). Organisations must more actively motivate employees to demonstrate sound security behaviour for this behaviour to disappear. Unfortunately, it is not well understood how to optimally encourage such behaviour (Bada and Nurse, 2019).

Understanding the reasoning behind poor security decisions requires insight into the psychological judgments that determine them. However, these judgements are internal, context-dependent, and cannot be measured directly. This presents a methodological challenge of how these MMs can be identified, assessed, and ultimately improved. Addressing this question is crucial to designing more targeted and impactful awareness and training strategies that more efficiently alter employee cybersecurity behaviour and ISA.

1.2 Research Aim

This research aims to develop a framework for identifying and classifying prevalent employee MMs about secure cybersecurity behaviour in organisational settings. By making these models observable and comparable, the framework aims to enhance the effectiveness of awareness and training initiatives, such as SETA programmes. This is done by aligning them with the cognitive profiles and vulnerabilities of different MM groups.

The framework draws conceptual inspiration from the work of Bitton et al. (2018), who proposed a hierarchical taxonomy to guide mobile security awareness strategies. Their study demonstrated that ISA interventions are more impactful when tailored to the threat categories most relevant to an individual's risk profile. Adapting this approach, the present study integrates existing cybersecurity taxonomies with the cognitive science concept of MMs.

By identifying recurring MMs in a structured way, this study contributes to a novel approach to assessing human factors in cybersecurity and supports evidence-based improvement of ISA in organisations.

1.3 Study Context

With these MMs representing conceptualised cybersecurity behaviour in a general sense, the limits on the scope of this research are established. Due to direct access to expertise and ongoing research on Shadow IT, this area was chosen as the focus of this study. Shadow IT is the autonomous deployment of unapproved IT by actors within a business (Klotz et al., 2019). This deployment involves the covert use of IT services that are not approved and do not align with the IT department's policies.

For example, when looking at cloud applications, the typical firm has 15 to 22 times more cloud applications running than have been authorised by the IT department (Corbin, 2015). This also cuts across sectors, with even highly regulated industries such as healthcare and financial services having substantially more applications running than estimated by the IT department. Despite this, CIOs still vastly underestimate the use of Shadow IT within their organisation (Corbin, 2015). Although understanding the potential risks and benefits associated with Shadow IT is becoming increasingly important, there has been a lack of substantial research in this field. This lack of research is especially true compared to phishing or ISA in general (Fürstenau and Rothe, 2014).

1.4 Main Research Question

With the problem of low ISA within organisations, which results in unintentional cyber vulnerabilities, it is essential to examine the underlying reasons. Understanding the behaviour and motivation of individuals is a successful basis to more effectively address internal security threats Zaytsev et al., 2017. Therefore, the Main Research Question that this study aims to answer is the following:

MRQ: "How can Employee Mental Models toward Shadow IT Usage be Identified and Assessed in an Organisational Context?"

Organisations need such identification and assessment methods to effectively identify security risks and increase awareness and education (Rafaiani et al., 2021). This human risk perspective is an essential component of an organisation's cybersecurity strategy (Bada and Nurse, 2019). With knowledge of the potential MMs of employees, an organisation can develop a stronger security culture. This culture enables better addressing human factors or behavioural issues that result in cyberrisks (Santos-Olmo et al., 2016). As it is less understood how best to encourage good security behaviour among employees, this area is of great importance (Bada and Nurse, 2019).

1.5 Contributions of the Research

This research aims to contribute to the scientific understanding and practical implementation of cybersecurity awareness by exploring how MMs of employees towards Shadow IT usage can be systematically interpreted and inferred. The research is located at the intersection of behavioural information security, organisational risk management, and human-centred cybersecurity. The contributions are both scientific and practical in nature.

1.5.1 Scientific

From a scientific perspective, this research advances the application of MMs in cybersecurity research by examining how these cognitive representations influence employee reasoning and behaviour in relation to security practices. It builds upon existing theory by offering a structured approach for eliciting and interpreting these models in organisational contexts. A key contribution is developing a novel framework to infer MMs through observable metrics, combining subjective inputs such as self-reported knowledge and attitudes with observed behavioural proxies. This structured, yet interpretive, approach avoids claims of direct measurability, instead focusing on the inference of MMs from multidimensional factors. The research also contributes to refining established security behaviour taxonomies by integrating cognitive and motivational dimensions. This allows for more holistic categorisation of user behaviour, not only based on what people do, but also on why they do it. This supports the development of more personalised and psychologically grounded intervention strategies. Furthermore, the study contributes to the field of human-centric risk modelling by demonstrating how MMs can be used to evaluate the potential effectiveness of behavioural interventions. It becomes possible to better understand how security MMs evolve and how targeted awareness efforts can influence their development over time.

1.5.2 Practical

From a practical standpoint, the framework presented in this research offers several valuable applications for organisations aiming to strengthen their cybersecurity awareness programmes. By enabling a more detailed understanding and categorisation of employee MMs, organisations can more effectively tailor SETA initiatives to the needs and risk profiles of different employee groups. This enhances the relevance and engagement of ISA programmes, allowing for a more efficient allocation of training resources towards areas of highest behavioural risk and individuals. Furthermore, the ability to interpret MMs offers insight into compliance and adoption challenges, helping to identify cognitive barriers that can contribute to employee resistance to security policies or technologies. The framework also provides a method for evaluating the impact of ISA strategies over time, allowing organisations to assess whether employees' reasoning about cybersecurity is shifting in line with educational goals. Ultimately, by deepening the understanding of how employees think about security, not just how they behave, this research supports a more effective cybersecurity culture across organisations.

1.6 Thesis Layout

This thesis follows a mixed-methods structured approach. Chapter 2 reviews related work on Shadow IT and MMs, establishing the theoretical foundation. Chapter 3 outlines the research design, including the research questions, methodology, and potential threats to validity.

Chapter 4 examines the prevalent challenges associated with ISA. It draws on existing literature to identify organisational and contextual factors that influence security behaviour and the effectiveness of awareness efforts. Chapter 5 examines the aspects related to employees that must be considered in MM assessments, focusing on individual-level factors that influence the formation and variation of MMs, such as knowledge, experience, and role. Chapter 6 examines how MMs on ISA can be assessed in organisational settings. It proposes a structured approach to elicit and interpret MMs through measurable metrics, primarily grounded in the Knowledge, Attitude, and Behaviour (KAB) dimensions. To demonstrate the practical applicability of the framework, an illustrative example is provided using the context of Shadow IT. Although this example draws on existing taxonomies and realistic scenarios, it does not represent a formal empirical case study and is instead intended to illustrate the framework's conceptual application.

Chapter 7 presents the general analysis and interpretation of the results of the previous chapters. Next, in Chapter 8, contributions to both the academic field and practical applications are discussed, along with the limitations of the study and its ethical concerns. Last, Chapter 9 answers the MRQ, followed by suggestions for future work.

2 Related Work

This chapter reviews the existing literature relevant to the core topics of this research: Shadow IT and MMs. Due to the limited academic focus on the intersection of these domains, this review adopts an exploratory approach to identify theoretical gaps and inform the study's framework.

2.1 Shadow IT

The rise of Shadow IT, defined as hardware, software, or services used without the knowledge or approval of the IT department, reflects an ongoing tension between organisational control and user-driven flexibility (Haag and Eckhardt, 2017). Shadow IT can introduce innovation and efficiency gains, but also poses substantial challenges in terms of cybersecurity, compliance, and governance (Kopper and Westner, 2016; Kretzer and Maedche, 2014).

This tension has become more apparent following large-scale shifts to remote work during the COVID-19 pandemic. More than 300 million employees globally transitioned to working from home, leading to an increase in the use of personal devices and unmonitored applications (L.-W. Wong et al., 2022). These practices, often intended to maintain productivity, increased vulnerability to attacks and led to a measurable increase in organisational cybersecurity incidents (Hijji and Alam, 2021; Lallie et al., 2021; Wash, 2010).

Shadow IT behaviours manifest in a variety of forms. Mallmann et al. (2019) categorise these into the sub-types unapproved cloud services, self-installed applications, self-made solutions, and employee-acquired devices. See Table 2.1 for an overview. Although some of these tools may support legitimate work needs, they also represent pathways for unauthorised access, data leakage, or policy workaround.

Table 2.1: Different sub-types of Shadow IT (adapted from Mallmann et al. (2019))

Shadow IT Type	Description	Example
Unapproved cloud services	Use of Internet-based software or SaaS not approved or known by the IT department. If accessed outside the workplace, it is also known as Mobile Shadow IT.	WhatsApp, Facebook, Skype, Dropbox, Google Apps
Self-made solutions	Solutions developed by employees on company systems to perform work tasks.	Excel spreadsheets, custom applications
Self-installed applications	Software installed by employees on company systems to support their tasks.	Freely available internet software
Self-acquired devices	Devices purchased by employees outside of IT procurement, including personal device use at work.	Smartphones, tablets, notebooks

The use of Shadow IT is closely associated with user non-compliance, often framed as intentional or negligent deviation from IT policies (Györy et al., 2012). However, the motivations behind non-compliance remain insufficiently understood. Some studies attribute it to perceived inefficiencies in approved systems, while others highlight a lack of awareness about associated risks (Silic and Back, 2014; Kirlappos et al., 2015). The underlying issue is not purely behavioural, but also cognitive. It is essential to look at how employees conceptualise risk, responsibility, and appropriate workarounds.

To address this, organisations require more than technical controls or rigid policy enforcement. They must understand how users internally rationalise their choices, particularly in security-sensitive contexts. This requires a deeper understanding of the MMs of employees, which will be discussed next.

2.2 Mental Models

As cyberthreats increasingly target the human element of organisations, understanding how people perceive threats has become increasingly important. MMs provide a foundational concept in human-computer interaction and cognitive psychology for analysing internal representations (Staggers and Norcio, 1993). These models are inherently context dependent, simplified, and often incomplete, yet shape user decisions, risk perception, and behaviour in significant ways (Johnson-Laird et al., 1998; Staggers and Norcio, 1993; Wash, 2010).

The relevance of MMs has grown along with the shift in attack strategies from exploiting technical flaws to manipulating human behaviour (Alladi et al., 2020; Bendovschi, 2015; Greitzer et al., 2014). Although organisations invest heavily in technical defences, behavioural risk remains difficult to quantify and address (Wiley et al., 2020). Unlike traditional risk assessments, MMs offer a lens through which these risks can be assessed. They help explain how users interpret threats, evaluate consequences, and justify their choices. By uncovering how employees reason about cybersecurity, organisations can develop more tailored and effective SETA programmes. These insights enable more accurate assessments of behavioural risk (Rafaiani et al., 2021).

Traditional risk assessment methods, while helpful, often fail to capture the cognitive and behavioural subtleties that influence cybersecurity-related decisions (Rafaiani et al., 2021). This challenge underlines the need to conceptualise and interpret these internal reasoning processes, an area where MMs can be particularly valuable.

2.2.1 Awareness Influence on Mental Models

The formation and refinement of MMs are directly influenced by an individual's knowledge and the level of ISA. Users with low ISA tend to construct fragmented or inaccurate models based on wrong assumptions, for example, believing that Apple products are immune to malware or that cybercriminals only target large corporations (Wash, 2010). These folk models reflect intuitive, experience-based reasoning, instead of structured security knowledge.

In contrast, users with higher ISA, whether developed through professional experience, training, or exposure to incidents, construct stronger MMs that align with the expertise perspectives (Jones et al., 2011; Tsohou et al., 2015). These models reflect a deeper understanding of threat factors, motivations, and mitigation strategies and are indicators of compliant security behaviour (L.-W. Wong et al., 2022).

However, ISA alone is insufficient, as research shows that generic awareness programmes often do not produce meaningful behavioural change. Chang and Coppel (2020) and Safa et al. (2015) emphasise that while organisations may have sufficient resources for standard solutions, successful awareness initiatives require careful design, professional planning, contextual relevance, and individual-level focus. Changing cybersecurity behaviour is complex and cannot be achieved through superficial awareness programmes alone, as these programmes often do not produce lasting behavioural change (Reeves et al., 2021).

2.2.2 Accuracy of Mental Models

As MMs are created based on one's ISA and knowledge, they are functional rather than complete or accurate (Wash, 2010; Wash and Rader, 2011). These folk models are abstractions of reality and contain inaccuracies about what they represent. Although technically inaccurate, these models still influence user decisions, sometimes in ways that diverge from organisational policies or best practices.

This insight has important implications for both system design and risk communication. Technologies or training programmes built on the assumption that users possess entirely correct MMs often fail to change behaviour. Instead, research suggests that it is possible to use inaccurate models to encourage positive security outcomes (Wash and Rader, 2011). Incorrect or incomplete MMs on cybersecurity can be just as important as accurate MMs, as they provide insight into how non-expert users perceive cybersecurity. By understanding which misconceptions are the most common and how users justify their behaviour, interventions can be designed to better align, correct, or adapt to existing MMs.

2.2.3 Mental Models on Shadow IT

An illustrative application of MM theory is found in the Shadow IT domain. Employees who adopt unapproved tools do not necessarily lack awareness; instead, they often act on the basis of an MM that frames these tools as beneficial, low risk, or even necessary. A study by Jansen et al. (2023) identified ten distinct MMs related to Shadow IT behaviour, classified into risk-averse and risk-taking profiles. An overview of the ten MMs identified is shown in Table 2.2. These models capture how users justify bypassing IT controls, whether due to a perceived need for autonomy, misjudged invincibility, or misguided trust in personal safeguards.

Although Jansen et al. (2023) offer a valuable taxonomy of attitudes, their study does not explore how MMs could be used as predictive indicators of secure behaviour or how they could inform the development of tailored interventions. This gap further supports the need for research focused on eliciting and evaluating MMs as practical tools for security management.

Table 2.2: Mental Models reflecting internal drivers influencing attitudes toward Shadow IT, adapted from Jansen et al. (2023)

Risk-Averse Mental Models	Risk-Taking Mental Models
Consequence-Avoidance Orientation	Common Sense Fallacy
Knowledge-Based Conservatism	Illusion of Self-Sufficiency
Risk Transfer Mindset	Misguided Sense of Protection
Cautious Seasoned Judgement	Performance-Driven Rule Bending
	Longevity-Based Invincibility
	Cost-Driven Compromise

2.3 Cybersecurity Taxonomies

Cybersecurity taxonomies are structured classification systems designed to interpret, categorise, and quantify various elements of cybersecurity, including technical vulnerabilities, behavioural risks, and human decision-making processes (Derbyshire et al., 2018; Rea-Guaman et al., 2018). As cybersecurity threats grow in complexity, taxonomies serve as foundational tools to build predictive models that anticipate risk behaviour and support proactive intervention strategies (Ferreira et al., 2023; Kure et al., 2022).

2.3.1 Mental Models in Cybersecurity

Recent efforts in predictive cybersecurity modelling have focused increasingly on human factors, such as behavioural tendencies and cognitive biases. Farahbod et al. (2020) argue that the predictive power of these models depends on their ability to integrate subjective human input, such as internal constructs that influence decision-making. MMs provide this foundation, as they reflect how individuals conceptualise threats, assess risk, and justify their behaviour (Wash and Rader, 2011).

However, the challenge lies in translating subjective mental constructs into structured, quantifiable data. One possibility to address this is the development of taxonomies grounded in behavioural and MM theory. These models capture not only what users know or do, but also how they think, through measurable elicitation methods.

2.3.2 Hierarchical Taxonomies

One such example is the hierarchical taxonomy proposed by Bitton et al. (2018), which aimed to assess security awareness among mobile users by categorising behaviours and attitudes into technology-focused areas and psychological dimensions. Bitton argues that one of the key challenges in security awareness research is defining clear, measurable criteria for identifying a "security aware" individual. Their taxonomy is based on the prototype developed by Kruger and Kearney (2006). These models assess ISA through the KAB framework (what a person knows, thinks, and does), where each dimension is broken down into subcomponents linked to specific security domains.

This hierarchical structure allows for the systematic evaluation of user awareness across psychological and technical dimensions. It begins with broad goals, such as increasing ISA, and narrows to actionable indicators at the individual level. When implemented with surveys, simulations, or monitoring systems, this approach supports the assessment of awareness gaps and behavioural risks in a structured and replicable manner.

2.3.3 Taxonomies as Predictive Tools

Taxonomies are not static classification tools; they are helpful predictive instruments when paired with behavioural data, system-generated logs, or observational metrics (Kruger, Drevin, et al., 2006; Parsons, McCormac, et al., 2014). For example, a low behavioural score in a particular KAB subdomain, when combined with known system vulnerabilities, can signal increased vulnerability and justify targeted and proactive awareness and intervention strategies.

More recent studies extend this predictive potential. Jansen et al. (2023) introduced a typology of MMs that categorises users as risk-averse or risk-taking in the context of Shadow IT. However, while these models categorise behavioural logic, they do not progress into predictive frameworks capable of anticipating risk. This gap presents an opportunity, as integrating MM-based classifications into behaviour-focused taxonomies could enable more adaptive and user-specific awareness strategies.

Furthermore, studies show that security behaviour is highly contextual, influenced by personal experiences, perceived self-efficacy, and organisational culture (Safa et al., 2015; Wash, 2010). To remain valid and valuable, future taxonomies must therefore support continuous and contextualised assessment. This would allow organisations to update their understanding of employee risk perception over time.

2.4 Summary

The review of the existing literature highlights three key insights central to this research. First, Shadow IT presents a significant behavioural risk in organisations, driven not merely by policy violations but by employees' internal reasoning and perceived needs. Second, MMs offer a compelling lens through which to understand such behaviour. However, they remain underutilised in predictive risk frameworks. Third, while existing cybersecurity taxonomies, especially those grounded in the KAB dimensions, provide a basis for structured assessment, they rarely incorporate the cognitive dimensions represented by MMs or translate them into measurable, user-specific insights.

These gaps emphasise the need for a methodological approach that can elicit, interpret, and categorise employee MMs within a cybersecurity context. To address this, the next chapter presents the design and execution of a mixed-method research strategy, integrating an NLR and expert SSIs.

3 Research Method

This chapter outlines the methodological approach used to address the main research question:

"How can Employee Mental Models toward Shadow IT Usage be Identified and Assessed in an Organisational Context?"

It presents the sub-research questions (SRQs) that guide the study, describes the methods used to address them, and explains the rationale for their selection. The chapter concludes with an analysis of potential threats to validity and the measures taken to mitigate them.

3.1 Research Sub-Questions

Three SRQs were formulated to support a structured investigation of the MRQ. Each SRQ targets a distinct aspect of the theoretical and empirical landscape surrounding MMs.

- *SRQ 1: What are the predominant challenges related to organisational information security awareness (ISA)?*

As discussed in Chapter 2, organisational ISA and employee MMs are closely interlinked. Low ISA often fosters inaccurate or incomplete MMs, which in turn influence risky or non-compliant security behaviour, reinforcing low awareness. This feedback loop presents a core challenge for cybersecurity education and training programmes, such as SETA. This SRQ is addressed through an NLR to identify theoretical and empirical challenges, and SSIs to explore how these challenges manifest in practice.

- *SRQ 2: What employee-related aspects need to be considered in the assessments of Mental Models (MMs)?*

MMs are inherently individual and shaped by previous experiences, knowledge, and context. Consequently, no two individuals' MMs are identical, especially given their diverse past experiences and knowledge levels. For example, distinctions between expert users and folk employees, as seen in Chapter 2, reflect substantial variations in security perception and behaviour. This raises the question of what other aspects of individuals may influence their mental representation of reality. This SRQ is first explored through an NLR to identify relevant personal and contextual factors. These insights are then examined and validated through SSIs to determine their influence on MM formation and assessment.

- *SRQ 3: How can Mental Models of cybersecurity be elicited and classified in the context of organisations?*

Although MMs help explain user behaviour and attitudes, they are subjective and not directly observable. Structured and quantifiable metrics are required to operationalise them. Existing research on cybersecurity behaviour and MMs suggests that MMs may be decomposed into standardised components or taxonomies, which can be used to compare and classify mental representations. This SRQ is approached through an NLR to examine previous attempts at modelling or quantifying MMs. The emerging metrics and taxonomies are then refined and evaluated through SSIs with cybersecurity experts and practitioners. This results in the creation of a preliminary framework.

3.2 Literature Review Strategy

A literature review is a critical component of secondary research, offering structured information on the current state of knowledge on a topic without generating new empirical data (Jahan et al., 2016; Snyder, 2019). It synthesises existing findings, identifies gaps, and supports the formulation of research questions and conceptual frameworks (Baumeister and Leary, 1997).

Several types of literature review approaches are available, including systematic, narrative, and integrative reviews (Snyder, 2019). Systematic reviews follow a strict protocol and are highly replicable, making them suitable for tightly defined research questions (Ferrari, 2015). However, their rigidity makes them less suitable for interdisciplinary exploratory inquiries, such as in the present study, which involves concepts that span multiple domains.

In contrast, narrative literature reviews (NLRs) offer a flexible framework for integrating diverse theoretical perspectives and empirical insights. They are particularly well suited to emerging cross-disciplinary topics studied by different academic communities using heterogeneous methods (Snyder, 2019; G. Wong et al., 2013). Although NLRs lack rigid inclusion protocols for systematic reviews, academic quality can be improved through transparency in source selection, structured synthesis, and incorporation of systematic elements (Ferrari, 2015).

3.3 Narrative Literature Review

Given the interdisciplinary nature of this research, an NLR was conducted to obtain information on ISA, employee behaviour, MMs, and Shadow IT. This approach was chosen for its flexibility in integrating heterogeneous findings from multiple academic domains.

Although the NLR does not follow a rigid protocol for systematic reviews, best practices for transparency and search traceability have been adopted to improve methodological rigour (Brereton et al., 2007; Ferrari, 2015). The review was structured using the IMRAD model of (Ferrari, 2015) to maintain clarity and consistency.

3.3.1 Search Strategy

A structured search strategy was employed across multiple academic databases to address the SRQs and inform the development of the framework. Although not a full systematic review, this strategy followed best practices for collecting evidence-based literature (Brereton et al., 2007). The selection of the database was guided by evaluations from Brereton et al. (2007) and Gusenbauer and Haddaway (2020), ensuring the coverage of both technical and behavioural sciences. The following databases were consulted:

- ACM Digital Library
- IEEE Xplore
- Scopus
- Web of Science
- Wiley Online Library
- WorldCat

Search terms were grouped into thematic clusters relevant to the research domains ISA, cybersecurity behaviour, MMs, and Shadow IT. These were used in various combinations depending on the platform syntax. An overview of the keyword groups is shown in Table 3.1

To ensure quality and academic credibility, sources were screened using journal and conference ranking systems. The Scimago Journal Rank (SJR) was used to assess journal prestige, while the CORE Rankings Portal was consulted for conference proceedings (González-Pereira et al., 2010). Where appropriate, credible grey literature was selectively included based on relevance and conceptual value.

Table 3.1: Grouped search terms used in the literature review

Core Theme	Sample Keywords and Search Phrases
Information Security Awareness	"Information security awareness", "Information security policy"
Mental Models	"Mental models", "Mental model assessments", "Quantify mental models", "Employee mental models", "Security mental models"
Cybersecurity Behaviour	"Security behaviour taxonomy", "Cybersecurity behaviour"
Shadow IT	"Shadow security"

3.3.2 Inclusion- and Exclusion Criteria

Following the initial search, a set of predefined inclusion and exclusion criteria was applied to ensure the relevance and reliability of the sources. Manual screening was conducted in several stages, beginning with a review of the title, abstract, and keywords. When necessary, this was followed by the introduction and conclusion.

Inclusion Criteria

- The presence of one or more keywords (as outlined in Table 3.1) within the title, abstract, or introduction
- A focus on peer-reviewed journals and high-quality conference proceedings
- Selective inclusion of credible white or grey literature, if deemed academically relevant
- Relevance to fields such as cybersecurity, psychology, human-computer interaction, or organisational behaviour

Exclusion Criteria

- Articles written in any language other than English
- Publications that lacked academic credibility or were not indexed in recognised ranking systems, such as the SJR or CORE
- Literature deemed outdated or no longer reflective of current developments in cybersecurity or MM research

The time window for inclusion prioritised papers published since 2010, reflecting technological shifts affecting cybersecurity behaviours, such as smartphone adoption. However, older foundational works were retained if conceptually essential and frequently cited.

3.3.3 Literature Screening

A prioritised screening approach was adopted to further structure the review process, balancing the breadth of coverage with practical relevance to address the SRQs. The literature screening process followed an approach consistent with the NLR methodology, with specific adaptations to optimise relevance and focus. Rather than conducting exhaustive filtering before addressing each SRQ, an iterative citation-based prioritisation strategy was applied. The articles with the highest citations were first reviewed to establish a strong conceptual foundation, followed by a sequential review of articles in descending citation order.

For all selected sources, relevance was assessed based on alignment with the SRQ themes and the conceptual focus of the study. The screening continued until thematic saturation was reached or when it was determined that additional sources no longer contributed meaningfully to the research objective.

3.4 Interviews

Although the NLR established a theoretical foundation for understanding ISA and MMS, the literature alone cannot fully capture the practical, contextual insights often held by domain experts. To address this gap, SSIs were conducted as a complementary method.

SSIs were selected for their ability to uncover nuanced perspectives that are typically inaccessible through secondary data sources. These include evolving taxonomies, applied assessment practices, and situated understanding of ISA and MM challenges within organisations. The interview topics focused on two core areas aligned with the SRQs: the nature and challenges of ISA in real-world settings and the interpretation, evaluation, and elicitation of MMs. The method allowed for both consistency across participants and flexibility in exploring emerging themes. This balance supported the collection of rich, context-specific data while ensuring alignment with the study's framework.

Interviews offer several methodological advantages, including the ability to probe complex reasoning and to control the flow of questions. However, they also have limitations, such as time intensity, limited scalability, and the potential influence of social desirability or interviewer bias (Alshenqeeti, 2014). A summary of the main strengths and limitations is presented in Table 3.2.

Table 3.2: Interview Advantages and Disadvantages in Research, adapted from Alshenqeeti (2014)

Advantages	**Disadvantages**
High response rate	Time-intensive to conduct and analyse
Low rate of incomplete responses	Limited scalability due to small sample sizes
Ability to capture real-world perspectives	Limited anonymity for participants
Controlled sequence and flow of questions	Potential for unconscious interviewer bias
Flexible and adaptive structure	Risk of inconsistencies across interviews

Other data collection instruments, such as surveys and questionnaires, were considered. Although these tools offer logistical benefits, such as efficiency and broad reach, they were not chosen as the primary method for this study. The key limitation lies in their inflexibility as questionnaires often fail to capture the depth and complexity of subjective reasoning (Patten, 2016; Nayak and Narayan, 2019). Questionnaires also risk low engagement, socially desirable responses, and a lack of contextual nuance. However, they were still used as a supporting role, as participants completed a brief demographic and background questionnaire prior to the interview. This enabled more informed preparation and interpretation of responses.

3.4.1 Interview Participants

To ensure the breadth and relevance of the perspectives, the participants were purposively selected from two key expert groups: organisational practitioners and academic researchers. All interviewees had cybersecurity expertise, but their roles, experience, and professional settings varied, allowing the integration of applied and conceptual insights.

Organisational experts included professionals in technical roles, governance functions, and consulting. These participants contributed practice-oriented knowledge on implementing and evaluating ISA, as well as challenges related to employee behaviour and MM assessment.

Academic researchers specialised in human-centred cybersecurity, MM theory, and behaviour-based security research. Their contributions provided theoretical grounding and domain-specific understanding of MM assessment and classification.

All participants were familiar with human factors in cybersecurity, and each was considered an expert in the subject matter. This dual source strategy enabled an efficient combination of theory and practice. The demographics, qualifications, and experience of the participants are summarised in Table 3.3.

Table 3.3: Interview participant demographics

ID	Rank	Highest Degree	Expert Grouping	Experience	Age group
P1	Senior	Master of Science	Organisational Expert	0-4	18-30
P2	Senior	Master of Science	Organisational Expert	8-12	31-45
P3	Senior	Post Master	Organisational Expert	12-16	31-45
P4	Manager	Master of Science	Organisational Expert	24-28	46-60
P5	Management	Master of Science	Organisational Expert	4-8	31-45
P6	Medior	CISSP	Organisational Expert	0-4	18-30
P7	Researcher	PhD	Academic Researcher	8-12	31-45
P8	Senior Management	PhD	Academic Researcher	28-32	46-60

3.4.2 Recruitment Strategy

Participants were recruited according to predefined eligibility criteria, including role, academic or professional credentials, and demonstrated experience in cybersecurity or human behaviour research. Participants were approached via email and LinkedIn with an invitation that included the study description and a request for referrals. This snowball sampling strategy aimed to capture diverse yet relevant individuals.

Diversity in role, background, and organisational affiliation within the sample was a deliberate goal, as this study aimed to develop a generalisable approach to assess MMs. This heterogeneity increases the transferability of findings and thereby supports the development of a generalisable MM assessment approach.

To structure data collection, this study followed the SSI protocol outlined by (Krombholz et al., 2019). Participants completed a short pre-interview questionnaire to collect demographic information and self-assessed cybersecurity experience. However, unlike (Krombholz et al., 2019), this study did not incorporate a drawing task to directly elicit MMs. This choice reflects the study's focus on evaluating existing MM assessment methods, rather than creating new MMs from scratch.

A target sample size of approximately 12 participants was established. According to Baker and Edwards (2012), six to twelve interviews are often sufficient to yield meaningful insights in qualitative research, particularly when thematic saturation is reached. Given that this study involved two distinct stakeholder groups, the sample size was doubled to ensure thematic coverage across both domains.

Due to time constraints and limited availability of qualified academic researchers, the final sample consisted of eight expert interviews. Although this number was sufficient to surface recurring themes and cross-role comparisons, some subgroup-specific nuances may not have been fully captured. This is especially true for the academic researcher expert group.

3.4.3 Interviewing strategy

This study employed an SSI approach to elicit expert perspectives on employee MMs and ISA. SSIs are particularly appropriate when the objective is to explore individual experiences, perceptions, and the reasoning process, rather than to derive statistically generalisable findings (McGrath et al., 2019).

The format offers a structured approach while allowing flexibility to pursue unanticipated insights during the interview (Adams, 2015; Adeoye-Olatunde and Olenik, 2021). Prepared questions were designed to ensure coverage of each SRQ, while follow-up questions enabled the interviewer to probe specific topics in greater depth, particularly those related to the reasoning behind participants' security and MM perception.

The interview protocol followed the principles outlined by Turner III and Hagstrom-Schmidt (2022) for qualitative design. Each session began with introductory questions, such as the participant's role within the organisation, to establish rapport and clarify the focus of the interview (McGrath et al., 2019). The interview then progressed to the participants' general knowledge of ISA, their previous experiences, and perspectives on organisational practices.

The main body of the interview was organised into thematic blocks, aligned with the study's three SRQs:

- **SRQ1:** Questions explored participants' familiarity with cybersecurity and ISA practices within their organisations
- **SRQ2:** Questions focused on how individual characteristics and roles influenced participants' perceptions of cybersecurity
- **SRQ3:** Questions addressed how MMs might be assessed or operationalised in practice

An informal pilot interview was conducted to evaluate the clarity, coverage, and timing of the interview protocol. Feedback from this pilot informed minor revisions to the question phrasing. The complete interview guide, partially adapted from Silic and Back (2014), is shown in Appendix A.1.

3.4.4 Data Collection

The interviews were conducted in English whenever possible, although Dutch was used when preferred by the participant. Each session lasted around an hour, depending on the participant's availability and level of expertise. All interviews were conducted online, at the participant's preference, and recorded with their informed consent.

The interviews were conducted online through Microsoft Teams [1], in accordance with the guidelines of Utrecht University. Although MS Teams offers built-in transcription, the final transcripts were generated using Amberscript [2] for greater precision.

To protect participant privacy, the transcripts were pseudonymised. This involved replacing identifying information with pseudonyms, enabling partial reversibility only through a separate key accessible solely by the research team. Pseudonymisation, distinct from complete anonymisation, was considered appropriate given the non-sensitive nature of most responses and the need for traceability during analysis (Mourby et al., 2018). All data processing complied with the GDPR and Utrecht University's ethical guidelines.

3.4.5 Data Analysis

The interview data was analysed using the framework analysis method, as described by Ritchie and Spencer (2002). This structured approach comprises five interconnected stages: familiarisation, thematic framework development, indexing, charting, and mapping/interpretation.

Analysis began with a thorough familiarisation, involving repeated listening and reading of the transcripts and associated field notes. During this phase, initial ideas and recurring patterns were noted. These informed the construction of a thematic framework aligned with the study's SRQS.

[1] https://www.microsoft.com/en-us/microsoft-teams/group-chat-software
[2] https://www.amberscript.com/en/

All transcripts, after pseudonymisation, were systematically coded using this framework. The codes were grouped into broader categories, creating a descriptive index. Both practitioner and academic participants were indexed under a shared coding scheme to ensure analytical consistency across groups. To enhance transparency and reduce interpretive bias, coding annotations were made directly within the transcripts. This allowed supervisors to trace analytical decisions and verify theme construction.

The indexed content was then reorganised through charting, with data grouped by theme rather than by individual participant. This enabled cross-case comparison and facilitated the identification of overarching patterns, differences, and shared perspectives.

In the final mapping and interpretation phase, key themes were further synthesised to uncover relationships, categories, and explanatory concepts. Charts and memos were revisited to examine connections, contrasts, and theoretical implications. The SRQs guided the interpretation process, ensuring analytic alignment with the research objectives.

The result of this process was a thematically structured account of key cybersecurity awareness risks, relevant employee-related variables, and MM assessment strategies, grounded in empirical data from interviews.

3.4.6 Ethical considerations

Prior to participation, all interviewees received a clear explanation of the purpose, scope, and procedures of the study. Informed consent was obtained using a standardised form based on the Utrecht University guidelines (Utrecht University, 2024), see Appendix A.2. This form detailed the study objectives, data handling practices, participant rights, and privacy measures. The interviews started only after written consent was granted.

To mitigate the risk of exposing personally or professionally sensitive information, all recordings and transcripts were securely stored. Furthermore, the transcripts were pseudonymised to reduce the risk of direct identification, and participants were offered the opportunity to review their transcript and request the redaction of specific content if desired.

Participation in this study was entirely voluntary. The consent form emphasised that participants could withdraw at any stage without providing a reason and without consequence. Contact details were provided for both the researcher and the academic supervisor, allowing participants to ask questions or raise concerns at any point before or after the interview.

3.5 Framework Development

This study proposes a framework to support the identification and assessment of cybersecurity-related employee MMs, based on insights from both NLR and expert SSIs. The framework aims to translate abstract MM concepts into structured and measurable proxy components.

The design draws conceptual inspiration from the framework by Kruger, Drevin, et al. (2006) to assess awareness of ICT security. Their model provides a foundational structure for defining focus areas, assessing awareness, and deriving metrics. However, the framework developed here extends beyond general awareness to focus specifically on the identification and interpretation of MMs.

The first stage of the framework involves identifying relevant focus areas and key domains in which employee MMs will be assessed. These domains include, for example, cybersecurity awareness, Shadow IT usage, and mobile security behaviour. A value-focused approach (VFA) is applied to define these areas, integrating stakeholder concerns, organisational values, and observed risk dimensions Drevin et al. (2007).

Following this, a mixed data collection strategy is applied. Employees are surveyed or interviewed on their knowledge, attitudes, and behaviours within the focus areas. These subjective insights are supplemented with objective data, such as system logs, to enhance the objectivity of the assessment (Kruger, Drevin, et al., 2006).

The combination of these data sources supports the construction of individual MM profiles. These profiles are structured using elements of value trees and scorecards, which represent the logical structure and relative weighting of the MM components. Aspects of the taxonomy identified from the existing literature are integrated with emergent concepts from the SSIs.

The development process itself was guided by the Baustein framework for design-oriented research (Schoormann et al., 2024). Specifically, the theory-driven configuration of Baustein was applied. This configuration supports translating theoretical knowledge into artefacts by deriving meta-requirements, formulating design principles, and instantiating these into a tailored framework.

An illustrative example was developed to demonstrate the practical application of the proposed framework. This example, discussed in Chapter 6, applies the framework to a scenario involving Shadow IT-related MMs. It draws on realistic input data derived from the existing literature and expert insight, but does not constitute a formal empirical case study. Instead, it serves as a proof-of-concept illustration to help interpret the logic and usability of the framework.

In summary, the framework provides a structured path from conceptual foundations to practical assessment mechanisms, aimed at improving the diagnosis and understanding of employee MMs in the context of cybersecurity risks.

3.6 Threats to Validity

As with all empirical research, this study faces several potential threats to validity. Validity refers to the trustworthiness of inference drawn from research results, specifically whether conclusions reflect reality and whether they could be compromised by methodological flaws (Feldt and Magazinius, 2010; Zhou et al., 2016). A crucial part of empirical research studies is to analyse and mitigate threats to its validity (Ampatzoglou et al., 2019).

Following standard empirical frameworks, validity threats are organised into four categories: Construct Validity, Internal Validity, External Validity, and Conclusion Validity Wohlin et al. (2012) and Zhou et al. (2016). Table 3.4 provides an overview of these domains.

Table 3.4: Overview of Common Validity Threats, adopted from Zhou et al. (2016)

Category	Definition
Construct Validity	Identify correct operational measures for the concepts being studied
Internal Validity	Seek to establish a causal relationship, whereby certain conditions are believed to lead to other conditions, as distinguished from spurious relationships.
External Validity	Define the domain to which a study's findings can be generalised.
Conclusion Validity	Demonstrate that the operations of a study, such as the data collection procedure, can be repeated with the same results.

3.6.1 Construct Validity

Construct validity concerns whether the chosen methods appropriately operationalise the abstract concepts under investigation. In this study, the key construct, employee MMs of cybersecurity, is inherently abstract and subjective. This presents a challenge for precise measurement.

To address this, MMs were elicited, and not directly measured, using both theory-informed interview protocols and insights from the literature. The framework analysis method supported the development of structured themes, and the questions were aligned with predefined SRQs to ensure consistency and traceability.

3.6.2 Internal Validity

Internal validity refers to the extent to which the observed findings can be confidently attributed to the research design rather than external or confounding variables. In the context of qualitative interviews, a key risk is social desirability bias, as participants might provide answers they believe are expected or favourable to the organisation.

To mitigate this, participants were explicitly informed of the study's academic independence and the use of pseudonymisation. An emphasis was placed on the voluntary and confidential nature of participation to reduce pressure to conform or self-censor.

3.6.3 External Validity

External validity refers to the generalisability of findings beyond the specific research setting. Although the NLR draws from a broad interdisciplinary base, which enhances generalisability at the theoretical level, the interview component poses more risk.

A potential threat arises from sampling within a single organisation, which could produce context-specific findings. To mitigate this, participants were selected from various roles and departments to ensure variation in perspectives. Furthermore, ideally, data collection continued until thematic saturation was observed, further supporting the analytical robustness.

3.6.4 Conclusion Validity

Conclusion validity assesses the degree to which the study results are credible and replicable, given the procedures followed. In this study, the validity concerns whether the conclusions, particularly the proposed framework, are adequately supported by the collected data.

To ensure validity, the interview protocols were tested in a pilot and reviewed for clarity and alignment with the research objectives. The use of a structured analysis method, framework analysis, ensured systematic coding and interpretation. The sources of literature for NLR were selected using a reproducible and clearly documented protocol, reducing selection bias.

Although qualitative research does not aim for statistical generalisability, the combination of transparent data handling, saturation-driven sampling, and method combination strengthens the applicability and credibility of the study's conclusions. Therefore, conclusion validity is considered adequately supported in the context of the objectives and design of the study.

3.7 Summary

This chapter details the methodological approach adopted in this study, including the research strategy, data collection, and analysis procedures. A combination of NLR and expert SSIs was used to explore the assessment of MMs in the context of cybersecurity. Ethical safeguards were implemented to ensure compliance with institutional and legal standards, and potential threats to validity were critically examined and mitigated.

With the methodological foundation established, the following chapters present the empirical findings derived from NLR and SSIs. These results will be organised around the study's SRQs and will inform the development of the proposed assessment framework.

4 Prevalent Challenges to Organisational ISA

We conduct an initial NLR, followed by semi-structured interviews to answer the first sub-research question: *What are the predominant challenges related to organisational information security awareness?*. This chapter will first focus on the results of the NLR, discussing several challenges found prevalent within the literature. Then, interviews are held to discuss these challenges with domain experts. The results of those interviews are discussed separately, after which a general conclusion is made.

4.1 NLR Process

The first step of the literature review is the digital search for relevant literature. In this study, a similar approach to that seen in the work of Jansen et al. (2023) is used.

Search: Different digital libraries were scanned in April 2024 using the following search string:

- *"(Information security awareness" OR SETA) AND (challeng* OR issue* OR challenge* OR probelem*) AND organisation"*

This search string returned 112 Scopus documents, 37 from Web of Science, three from ACM Digital Library, 35 from IEEE Xplore, three from Wiley Online Library, and six from WorldCat. To limit the number of weakly related results, it was decided to scan only the title, keywords, and abstract. A total of 196 documents were collected.

Filtering: The total number of 196 documents had varying levels of quality and were therefore filtered, resulting in a final list of highly relevant and high-quality documents. The exact filtering criteria as described in Section 3.3.2 were used.

Due to multiple digital libraries, there is an inherent risk of creating duplicate references. Duplicates are first identified by EndNote, a reference manager [1]. While this process is automatically executed, all duplicate results are manually checked before removal. EndNote flagged one duplicate from Scopus, 25 from Web of Science, 12 from IEEE Xplore, two from Wiley, and one from WorldCat. Furthermore, the remaining list was manually checked to ensure that EndNote did not miss anything. This manual check removed 15 more duplicates, resulting in 140 unique documents.

[1]https://endnote.com/

4.2 NLR Results

The literature screening process resulted in 28 articles relevant to addressing SRQ1. Figure 4.1 presents the distribution of publication years, revealing a notable peak around 2010. Nine of the selected papers were published prior to 2010; many of these were incorporated through snowball sampling from more recent, highly cited, papers. An overview of all selected articles, including their research objectives and primary focus areas, is provided in Table 4.1.

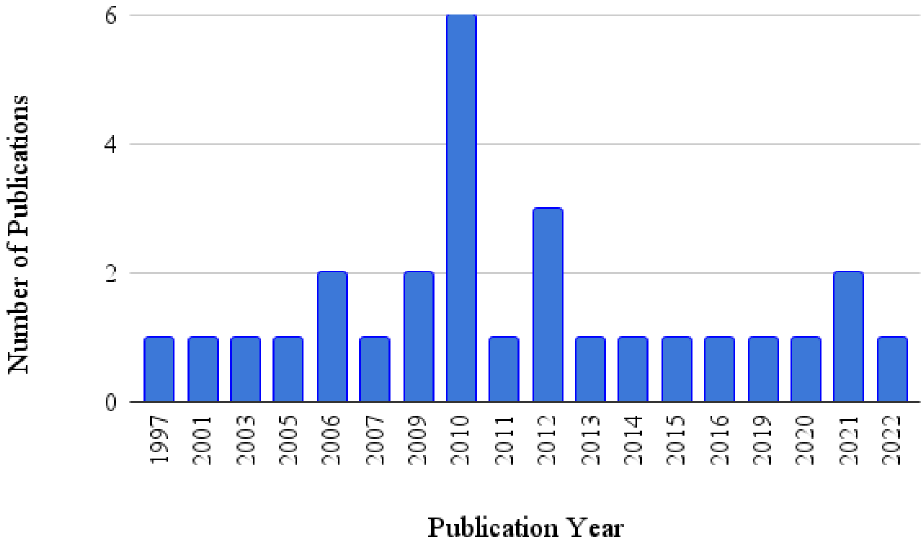

Figure 4.1: Publication overview of used literature

4.2.1 General ISA Findings

Analysis of the 28 selected studies revealed consistent themes regarding the human dimension of cybersecurity. Although nine papers explicitly targeted ISA as their primary focus, the remainder addressed adjacent domains such as security behaviour, Information Security Policy (ISP) compliance, and decision-making in organisational contexts.

In general, there is a broad consensus that technical security controls alone are insufficient to guarantee IS (Furnell and Clarke, 2012; Spears and Barki, 2010). Even when strong technical defences are in place, overall cybersecurity remains vulnerable to human error. Minor errors by employees can create major vulnerabilities (Shaw et al., 2009; Öğütçü et al., 2016). Therefore, organisations must invest not only in technology but also in the socio-organisational aspects of cybersecurity (Bulgurcu et al., 2010; Safa et al., 2015; Shaw et al., 2009).

An important socio-organisational aspect is the compliance of employees with ISPs (Kirsch and Boss, 2007). ISPs provide guidelines to employees on their roles and responsibilities in protecting organisational information assets. However, the mere presence of policies does not automatically ensure that employees act accordingly (Bulgurcu et al., 2010; Stanton et al., 2005; Whitman et al., 2001). Some employees may not be sufficiently motivated or lack the awareness needed to follow policies effectively.

Table 4.1: Objectives and Focus Areas of Studies Included in the NLR

Reference	Objective	Focus Area
Abawajy (2014)	Discuss delivery methods for improving ISA and security behaviour	ISA
Aldawood and Skinner (2019)	Highlight ongoing issues and pitfalls of social engineering attacks	SE
Bulgurcu et al. (2010)	Identify rationality-based factors influencing employee compliance with ISPs	ISPs
Conteh and Schmick (2021)	Evaluate IT infrastructure vulnerabilities with focus on social engineering	SE
Crossler and Bélanger (2009)	Investigate how characteristics and SETA programmes affect secure behaviour	Behaviour
Furnell and Clarke (2012)	Explore the human aspects of computer security and their integration into security strategy	ISA
Khando et al. (2021)	Provide methods and factors for enhancing employee ISA	ISA
Kirsch and Boss (2007)	Develop a model to explain information security behaviour	Behaviour
Kruger and Kearney (2006)	Report on the development of a measurement model for ISA	ISA
Kumaraguru et al. (2010)	Educate users about phishing and trust decision-making	SE
Lacey (2010)	Investigate practical issues and lessons in ISA and organisational change programmes	ISA
Leach (2003)	Identify factors that influence secure behaviour and ways to improve them	Behaviour
Martocchio and Judge (1997)	Investigate how self-deception and self-efficacy influence conscientiousness and learning	Behaviour
Öğütçü et al. (2016)	Investigate risky behaviours threatening information security	Behaviour
Pattinson, Jerram, et al. (2012)	Investigate behavioural responses to phishing emails	SE
Reicher (2010)	Discuss the role of social and emotional learning in cybersecurity	Stress
Safa et al. (2015)	Investigate motivational factors affecting user security behaviour	Behaviour
Shaw et al. (2009)	Examine the impact of hypermedia, multimedia, and hypertext on ISA	ISA
Sharif and Mohammed (2022)	Analyse financial losses and future trends in cybersecurity risks	FC
Silic and Lowry (2020)	Explore gamification as a method to improve security training	ST
Son (2011)	Develop an intrinsic motivation model to explain employee security rule compliance	ISPs
Spears and Barki (2010)	Examine organisational user participation in security and its impact on IS security controls	IS
Stanton et al. (2005)	Examine secure behaviour using a two-dimensional map based on expertise and intention	Behaviour
Stewart and Lacey (2012)	Examine the failure of mainstream ISA techniques and suggest improvements based on psychology	ISA
Talib et al. (2010)	Understand the knowledge-practice relationship between workplace and home environments	ISA
Valentine (2006)	Explore efficient approaches for employee security awareness programmes	ISA
Whitman et al. (2001)	Address the importance of policies for organisational threat protection	ISPs
Willison and Warkentin (2013)	Analyse insider threats and extend threat taxonomies for greater granularity	Behaviour

SE = Social Engineering, ST = Security Training, FC= Financial Costs

In this context, Bulgurcu et al. (2010) defined ISA as an employee's understanding of general security risks (General ISA) and their specific knowledge of the organisation's policies (ISP awareness). Employees with higher levels of ISA tend to show more favourable attitudes towards ISP compliance and demonstrate more secure behaviour overall (Safa et al., 2015; Shaw et al., 2009). It was also found that a higher level of ISA strengthens beliefs about the benefits of compliance and the risk of non-compliance, while reducing the perceived cost of compliance. Furthermore, attitudes, normative beliefs, and self-efficacy to comply were identified as important factors that, together with ISA, influence the employee's intentions to comply with security policies.

Although increasing ISA through training is often proposed as a recommended solution, research shows that changing behaviour through training alone is complex and frequently ineffective if not properly designed and contextualised (Aldawood and Skinner, 2019; Kumaraguru et al., 2010; Safa et al., 2015). Based on the reviewed literature, 11 prevalent challenges were identified that hinder the effectiveness of ISA programmes and secure behaviour.

4.2.2 Prevalent Challenges

Improving employee awareness and behaviour requires well-designed information security training programmes (Safa et al., 2015). Although awareness plays a crucial role in ensuring information security, it is evident that it is not meaningful on its own. One could decide not to act protectively, even when aware of security misuse. As important for information security are behaviour, previous experiences, exposure to threats, and perception of danger (Öğütçü et al., 2016; Safa et al., 2015).

As no system is perfectly secure, information security risk cannot be entirely mitigated by improving ISA and behaviour. Instead, the objective of ISA programmes is to train employees to reduce risks to an acceptable level (Aldawood and Skinner, 2019; Öğütçü et al., 2016). The reviewed literature highlights several key challenges that organisations must address to improve the effectiveness of ISA initiatives and foster secure employee behaviour, starting with having a clear methodology.

Clear Methodology: Although awareness programmes, such as SETA, are considered the key to success in ISA, and are designed to provide employees with the knowledge and motivation to comply with ISPs, many initiatives lack a structured methodology to develop effective and engaging materials (Crossler and Bélanger, 2009). As a result, initial impacts often decrease rapidly, producing little lasting change in employee behaviour (Khando et al., 2021; Lacey, 2010; Shaw et al., 2009).

Constant Change: One of the critical success factors for ISA programmes is the relevance, timeliness, and consistency of security information (Kruger and Kearney, 2006). One of the challenges that organisations face with these programmes is the constant change in security threats. Online attackers continuously evolve new risks and attacks to elude widely accepted security technologies. As a result, ISA is always a few steps behind, and training must be constantly updated to remain relevant and reduce security breaches (Aldawood and Skinner, 2019; Bulgurcu et al., 2010; Safa et al., 2015; Shaw et al., 2009; Öğütçü et al., 2016).

Employee Motivation: Weak training design often leads employees to view ISA programmes as tedious, ineffective, or cumbersome. They find sessions to be lengthy and challenging to focus on, especially when there are other work-related tasks they are concerned about (Aldawood and Skinner, 2019; Leach, 2003; Silic and Lowry, 2020). Even after completing sessions, employees may lack intrinsic or extrinsic motivation to apply what they learned (Safa et al., 2015). This disengagement reinforces non-compliance with ISPs and remains a barrier to developing a strong security culture (Willison and Warkentin, 2013).

Medium Richness: The method of delivery of ISA materials strongly affects their effectiveness (Abawajy, 2014). It is important to be able to deliver information in different ways to convey messages more effectively to various types of employees. A combination of media, such as workshops, games, emails, and online modules, is advised to improve training results (Safa et al., 2015; Shaw et al., 2009). However, this selection is a difficult task, as too much media richness could have a harmful effect on learning performance(Shaw et al., 2009).

Organisational Culture: Embedding security awareness into organisational culture is essential to increase the overall secure environment within an organisation. However, changing employee perceptions and behaviours is complex and gradual (Safa et al., 2015; Shaw et al., 2009; Son, 2011). Larger, data-intensive organisations, in particular, must prioritise ISP compliance efforts (Bulgurcu et al., 2010). However, even organisations that recognise the importance of culture often struggle to define or operationalise it effectively (Lacey, 2010; Aldawood and Skinner, 2019)

Pressure and Stress: External pressures and time constraints often reduce the participation of employees in ISA programmes (Lacey, 2010; Talib et al., 2010). Compliance requirements may be perceived as barriers to productivity, leading employees to prioritise operational tasks over training (Bulgurcu et al., 2010; Conteh and Schmick, 2021). Offering training outside working hours is generally also ineffective, as it disrupts work-life balance and reduces willingness to participate (Reicher, 2010; Aldawood and Skinner, 2019)

Programme Personalisation: ISA initiatives must be tailored to individual employee needs and organisational roles. The challenge is to make these programmes reflect the needs of a variety of employee groups while reflecting the needs of the business (Aldawood and Skinner, 2019; Shaw et al., 2009; Valentine, 2006). One-size-fits-all approaches have repeatedly underperformed. Instead, when conveying information, it is important to look at what is taught, when it is taught, and how it is taught. ISA training must be adapted according to the level of awareness of employees and the organisational hierarchy (Talib et al., 2010; Sharif and Mohammed, 2022). However, budget constraints can limit the degree of personalisation feasible.

Resource Availability: Resource limitations remain one of the main barriers to effective ISA programmes. Organisations often prioritise investments in technical solutions over human-focused initiatives. This results in most campaigns being little more than token gestures with a small budget and limited content (Aldawood and Skinner, 2019; Lacey, 2010). Nevertheless, careful targeted, lower-cost campaigns can yield significant security improvements and cost savings if strategically planned (Lacey, 2010). As not all employees have the same level of awareness, training costs could be significantly reduced by detecting and focusing on at-risk employees and creating purpose-focused sessions for those groups. Furthermore, the cost of risk containment is believed to be much higher than that of adequate training in the first place (Aldawood and Skinner, 2019)

Risk Communication: Effective communication of risk severity and frequency is essential to foster realistic perceptions of threats (Shaw et al., 2009). Keeping employees informed about vulnerabilities and potential breaches improves their engagement and proactive behaviours, as they feel more invested in the safety of the organisation (Safa et al., 2015; Aldawood and

Skinner, 2019). Studies show that informed users perform significantly better in resisting threats such as phishing attempts (Pattinson, Jerram, et al., 2012).

Technical Standpoint: A technical standpoint refers to designing ISA programmes primarily from the perspective of technical experts, focusing on what should be taught rather than what employees actually need to learn. Many ISA programmes are developed based on this approach, often judged by the forcefulness of their delivery or the novelty of interventions, rather than their practical suitability and effectiveness. Instead, effective awareness campaigns must be grounded in the realities, motivations, and behaviours of employees to produce lasting change (Lacey, 2010; Stewart and Lacey, 2012).

Threat Appraisal and Self-Efficacy: The perception of the severity of a threat, and the belief of an employee in their ability to securely respond, influence compliance behaviours (Öğütçü et al., 2016; Safa et al., 2015). Wrong threat assessment, together with low self-efficacy, reinforced by inadequate or poorly delivered training, can cause employees to overestimate their abilities or underestimate risks. Even when employees think they are informed about security and how to protect themselves, this is often not the case (Martocchio and Judge, 1997; Talib et al., 2010; Aldawood and Skinner, 2019).

4.3 Interview Results

The challenges identified through the NLR were validated and contextualised through SSIs with experts in the cybersecurity domain. These interviews served to test the findings against practical experiences, allowing the identification of potential discrepancies and the emergence of additional challenges not covered in the literature. The grounding and coverage of topics are provided and indicated as the number of times a topic was mentioned and the number of participants who discussed it, respectively. A complete overview of the resulting qualitative codes is provided in Appendix A.3.

4.3.1 Prevalent Challenges

Although the experts confirmed many of the challenges derived from the NLR, the perceived relevance of each varied. In particular, two interviewees explicitly stated that all 11 challenges had an *equal prevalence (2/2)*. However, this view was not widely shared. Most of the participants identified specific challenges as more pressing than others, often offering additional nuance or framing them differently than described in the literature.

Organisational Culture (12/7): The literature highlights both the importance and difficulty of fostering a strong organisational culture around cybersecurity, an observation shared by nearly all interviewees: *"Organisational culture is crucial because it sets the norm. If everyone follows security measures, it becomes the default behaviour. People are encouraged to act securely if they see examples around them."* [P7]. Several experts stressed that employees must feel supported and encouraged to participate in security, and that culture is significantly shaped by leadership: *"[...] they lack a secure culture where people feel their questions are welcome and answered. Staff should be supported by the security team, not hindered."* [P3]

The role of top management was consistently cited as a determining factor in shaping security culture: *"[...] if the leaders do not think it is important, neither will the employees. So leaders must set an example"* [P1]. The importance of starting from the top down was shared amongst several participants: *"[...] there's a lot to do with organisational culture and motivation, from the top down. It starts at the top"* [P4]

Programme Personalisation (8/5) & Medium Richness (2/2): Five participants highlighted the need for personalised awareness programmes and media: *"So, personalisation and medium richness are two big issues. People here are highly motivated, but they are also very particular about how they want to learn."* [P3]. The literature suggests that no single format suits all employees, a view widely supported in the interviews: *"There's also often a lack of effective and versatile training. Not everyone learns in the same way [...]"* [P2]. It was also noted that generic one-size-fits-all training can reduce engagement: *"Currently, most awareness programmes are very generic, targeting the entire organisation. This leads to demotivation."* [P7]

Medium richness, how well the training medium matches the content and learning need, is strongly related to *programme personalisation*. Both the literature and the participants see this as a key factor in programme effectiveness: *"Many specialists feel that a general awareness approach isn't suitable for their roles. [...]. So, personalisation and medium richness are two big issues. People here are highly motivated, but they are also very particular about how they want to learn."* [P3]

Resource Availability (7/5): The challenge of limited resources was widely acknowledged, both in the literature and by participants. Although much of the academic focus has been on budget constraints, experts noted that the problem extends beyond money: *"Yes, limited resources are not just financial. It also comes down to having a solid business plan to show management how to track progress. Most investments go into technical measures, while only about 5% go into the human side, which is insufficient given that most successful attacks have a human element"* [P6]

However, some interviewees pushed back on this narrative. They argued that budgets are generally sufficient, but the ability to justify awareness efforts is lacking: *"But security budgets are actually pretty decent in most companies right now. So, I don't fully agree with people who claim to have no money. Most of the time, the issue is that they cannot present a strong business case for a specific solution"* [P7]

In addition, beyond financial resources, many participants pointed to the shortage of qualified personnel: *"[...] there's the issue of limited resources. This can be because the organisation itself has limited resources, but it is also due to the shortage of cybersecurity experts, especially those focused on awareness"* [P1] and *"Then there's the issue of manpower: there are simply not enough people or hours dedicated to security awareness."* [P2]

Despite these issues, some experts maintained that resource constraints should not be used as an excuse: *"But even with limited resources, or a shortage of security staff, or a smaller budget, you can still run an effective security awareness programme. It doesn't have to be expensive." [P4]*

Technical Standpoint (6/5): Several participants criticised the narrow technical lens through which many ISA efforts are implemented. Security policies often lack usability considerations, and programmes are developed without meaningful input from other departments: *"[...] to have an effective programme, you need to collaborate with other departments like communications and marketing. They know how to influence behaviour, even if they don't specialise in security."* [P6]

The participants expressed concern that technical experts often do not think behaviourally: *"I don't think we need more experts just telling us what to do, we need more focus on actual behaviour change"* [P1]. Other participants also pointed to the persistent trade-off between usability and security:*"These are the security controls being developed, but the end user is often not considered. There is always this tension between security and usability."* [P8]

Employee's Motivation (5/5): Motivation was viewed as closely related to compliance. Some participants noted that training often fails to engage employees who are willing to act securely: *"[...] I found many employees highly motivated to become champions of security in their own domains. However, training can be dull for some, or below their level, which demotivates them"* [P2]. Motivation as a critical driver is instead cast in the shadow of knowledge: *"Behaviour arises from capability, opportunity, and motivation, a simple behavioural model. However, we often focus too much on knowledge and forget that motivation is a critical driver of behaviour. That is where things can go wrong."* [P8].

Others believed that motivation is generally sufficient but not always activated: *"[...] most people want to be secure. They may not know how to do it correctly, and they might make mistakes. Still, there is a motivation to be careful with company assets and to understand the importance of cybersecurity." [7]*

Clear Methodology (4/4): A recurring challenge was the lack of a structured and clear methodology in awareness initiatives: *"From the perspective of someone aiming to improve information security awareness, I think the biggest risk is the lack of a clear methodology."* [P1]. Participants observed that many programmes feel arbitrary or reactive: *"The lack of a clear methodology is also an issue. Often, there is no clear introduction or set goals for training, just random phishing tests or memos, without explaining why we are doing it."* [P2]. This view confirms the literature's critique of the fragmented, non-systematic nature of many ISA efforts.

Time and Stress (4/4): Participants agreed that external pressures, such as stress and tight deadlines, negatively influence secure behaviour. Security is often deprioritised when efficiency is rewarded: *"Everyone is under pressure and stress constantly in their regular jobs, not just in security-related roles. [...] The more time pressure there is, the more security is seen as an obstacle or roadblock"* [P7]. The time constraint in those situations is a leading factor in secure behaviour: *"If efficiency and speed are key measures in your organisation, secure behaviour, which usually takes time, becomes a challenge."* [P3]

Constant Change (3/2): Constant change in tools, policies, and threats was seen by some interviewees as a major challenge. In particular, it was linked to a loss of confidence and continuity in secure behaviour: *"It's difficult for people to learn secure behaviour, and then when things change, they lose confidence."* [P3]

Others took a more nuanced view, pointing out that while the environment evolves, the core behavioural expectations remain stable: *"It's always the same message: don't click links, lock your screen, tidy up your workspace, close the door, wear your badge, or report anything suspicious. Did I mention not to click the link? It is the same information over and over. So, it is not as if the core message is changing constantly."* [P4] and *"There's a constant change in threats, information, and techniques. [...] and when you look at what employees need to know or report, I think that has not really changed." [P5]*

Risk Communication (3/2): Participants mentioned that communication is often one-way. Without feedback or follow-up, employees lose trust in the process: *"[...] you never hear back after reporting. So, why would I report next time? Probably, it was not an issue because the organisation did not mention it."* [P5]. Others proposed integrating security messaging into existing routines to improve visibility and relevance: *"One of the best things is to integrate security into existing rituals. For example, many companies have regular Monday meetings, and you can include security updates during those." [P6]*

Self-Efficacy and Threat Appraisal (1/1): Although this challenge is often repeated in the literature, only two participants recognised it. Their views were stark opposites. One saw it as central to ISA: *"So, I would name three key factors: organisational culture, motivation, and threat appraisal, along with self-efficacy. These are the biggest problems organisations face and the areas they can improve upon the most."* [P4]. However, the other strongly disagreed: *"To be honest, we use that as well in our team, but I think it's total nonsense because most people are fully capable; they're just lazy."* [P5]. This illustrates the mixed reception of psychological models in real-world security practice.

4.3.2 Additional Identified Challenges

In addition to the eleven challenges derived from the literature, several new challenges emerged from the interviews. These findings highlight further areas of concern that are not adequately covered in existing research. A summary of these additional challenges is provided in Appendix A.3.

Focus on the Human Aspect of Security (4/3): A recurring concern was the limited attention paid to the human side of information security. Participants noted that organisations continue to prioritise technical controls over behavioural and cultural strategies, despite widespread acknowledgement that most successful attacks exploit human weaknesses: *"Most investments go into technical measures, while only about 5% go into the human side, which is insufficient given that most successful attacks have a human element"* [P6]

Even when human-centred efforts are made, they are often reduced to educational initiatives without deeper engagement with behaviour, motivation, or organisational dynamics: *"If you look at awareness solely from an educational perspective, I think you are missing quite a lot. [...] they focus on education, education, education. But you need to look at the people"* [P4]. This suggests a conceptual gap in how organisations interpret ISA, often treating it as a one-directional knowledge transfer rather than a process of enabling secure behaviour.

Difference in Expected and Actual Behaviour (2/2): Two participants pointed out that there is often a mismatch between how organisations assume employees will behave and how they actually behave in day-to-day operations. This gap is rarely addressed systematically: *"People behave in a certain way, and to say that they make mistakes is unjust. The fault often lies with the system design, which does not account for the fact that people operate within the system"* [P8]. This challenge underscores the need for more realistic behavioural modelling within ISA strategies and highlights the risks of designing programmes based solely on formal policy assumptions.

Focus on High-Risk Individuals (2/2: The importance of targeted awareness strategies for high-risk individuals, such as executives, administrators, or departments handling sensitive data, was also highlighted by several participants. Despite the elevated risk profile of these groups, they are often included in broad, generic training that fails to account for their specific threat exposure: *"[...] a system administrator, who plays a key role in security and is a prime target for threats, should have a completely different awareness programme than a janitor. [...] Customising awareness programmes based on job functions makes them much more applicable and effective."* [P7]. This finding aligns with broader calls for risk-based personalisation in ISA programme delivery, but highlights that such approaches remain underutilised.

Measuring ISA Performance Results (2/2): Another underexplored challenge was the lack of robust methods to measure the effectiveness of ISA efforts. Although organisations attempt to track performance through phishing simulations or training completion rates, these metrics rarely reflect actual behavioural change: *"One challenge is getting the budget for security awareness efforts and tracking results. [...] but it is hard to quantify the results of security awareness efforts."* [P6]. This limits the ability of organisations to justify investments in ISA and may contribute to the perception that awareness initiatives are less valuable than technical controls.

These additional findings reinforce a broader theme in the interviews. Although organisations increasingly acknowledge the role of human behaviour in security, practical implementation often falls short. Underestimation of behaviour differences, lack of performance metrics, and inadequate adaptation to specific organisational roles continue to hinder effective ISA programme outcomes.

4.3.2.1 Facilitation Factors

Although the primary focus of the SRQ was on ISA challenges, several participants also briefly mentioned factors that can support or strengthen ISA and behaviour within organisations. These facilitation factors were not explored in depth, and no follow-up was conducted on them. Therefore, they are only reported here as observed during the interviews. For further elaboration, the readers are referred to the full transcripts.

Facilitation factors mentioned included interactive and social methods for engagement, such as interactive security events, the use of internal social platforms, and direct conversation engagement. Several participants emphasised the value of clear guidance, specifically providing explicit instructions on what is allowed and where to go with questions. In addition, the embedding of security in existing organisational rituals, such as weekly team meetings, was also discussed. Lastly, phishing simulations and actor-led social engineering tests were suggested as practical tools for both training and risk identification.

4.4 Summary

To answer SRQ1, *What are the predominant challenges related to organisational information security awareness?*, the research explored both theoretical and practical perspectives on ISA. Although ISA is widely recognised as essential to ensure information security, the findings make clear that awareness alone is insufficient. Behavioural factors, perceptions of risk, and motivational dynamics also play a critical role. What these factors share is their dependence on effective training, communication, and a broader organisational context.

The NLR identified 11 key challenges that contribute to the difficulty of improving ISA and behaviour: Organisational culture, programme personalisation, medium richness, resource availability, technical standpoint, employee motivation, clear methodology, time and stress, constant change, risk communication, and self-efficacy and threat assessment.

Although specific themes, such as personalisation and motivation, were discussed more frequently, all 11 were recognised as relevant to understanding why ISA efforts often fall short. Expert interviews largely supported the findings of the literature, though with some variation in emphasis. Although two participants viewed all the challenges listed as equally relevant, most experts highlighted only a subset of especially critical factors.

All but two of the challenges were identified by more than one participant as prevalent. However, interviews also revealed significant variation in the way these challenges were prioritised, with specific challenges considered critical by some and negligible by others. This variation reflects the complex, context-specific nature of human behaviour and highlights the difficulty of establishing a general understanding of what influences ISA and behaviour in organisations.

4.4.1 Visualising Challenge Interdependencies

Figure 4.2 presents an overview of the key challenges that impact ISA and secure behaviour. In addition to illustrating how each factor affects ISA, the figure also visualises the interdependencies between several of these challenges.

The relationship between *Resource Availability* and *Clear Methodology* stems from the resource-intensive nature of the ISA initiatives. Many awareness programmes fail due to poor design, and developing structured and effective methodologies typically requires financial and human investment. *Constant Change* also places ongoing demands on available resources, as awareness content must be regularly updated to remain relevant. Another bidirectional relation exists between *Resource Availability* and *Programme Personalisation*. On the one hand, tailoring ISA efforts to individuals increases development costs, while better-targeted, more effective programmes may reduce long-term costs by preventing incidents. Similarly, *Resource Availability* is influenced by *Risk Communication*. Inadequate communication, such as failing

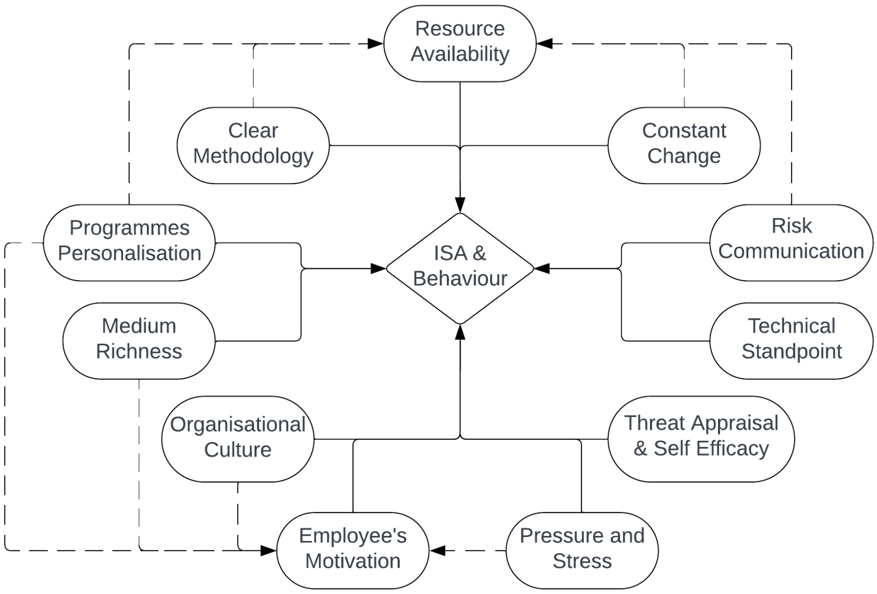

Figure 4.2: Overview of prevalent challenges related to ISA and behaviour

to notify employees about current threats or response protocols, can increase the likelihood and severity of incidents, placing additional strain on organisational resources.

A second cluster of interdependencies centres around *Employee Motivation*. The absence of *Programme Personalisation* and *Medium Richness* was frequently mentioned as leading to disengagement, as employees struggle to connect with content that does not align with their needs or preferences. *Organisational Culture* also plays a crucial motivational role. When leadership models secure behaviour and foster open dialogue, employees are more inclined to adopt similar practices. Finally, *Pressure and Stress* were seen to reduce motivation by limiting the willingness of employees to prioritise secure behaviour in high-pressure environments.

These findings highlight that ISA challenges do not exist in isolation. Instead, they interact and form a complex system in which addressing a single challenge may be insufficient without considering its broader organisational context. This includes the analysis of SRQ1 and offers a general view of the factors that influence ISA and secure behaviour. In the next chapter, attention turns to SRQ2, focusing on which employee-related aspects should be considered when assessing MMs in the context of IS.

5 Employee-related Aspects of Mental Model Assessments

The previous chapter outlined a range of organisational-level challenges that influence the success of ISA initiatives. Key factors such as "Lack of Clear Methodology", "Organisational Culture", "Constant Change", and "Expert Governance" offered a high-level understanding of why ISA programmes often fall short. However, these explanations mainly reflect organisational dynamics rather than individual variation.

In contrast, factors like "Programme Personalisation" and "Medium Richness" highlighted the need to consider individual differences in designing effective awareness programmes. These factors imply that what works for one person may not necessarily work for another. However, the root cause behind this variability is unexplored.

A potential explanation lies in the difference between individuals' MMs. Since MMs shape how people interpret and respond to security-related information, it is critical to understand what influences their development. This chapter explores the employee-related aspects that contribute to such variation in MMs and, by extension, in ISA and security behaviour.

To investigate this, a second NLR was conducted, specifically focused on identifying the individual-level factors of MM variation, complemented by expert SSIs. The combined insights aim to address the second SRQ: *What employee-related aspects need to be considered in the assessments of Mental Models?*

5.1 NLR Process

The same NLR methodology described in 4.1 was applied. A comprehensive search was carried out across multiple databases using the following search string:

- ("Information security awareness" OR SETA OR "information awareness" OR "security awareness" OR "information security") AND (intrinsic OR personality) AND (demographic OR individual* OR culture OR education OR location)

The query yielded 122 results from Scopus, 68 from Web of Science, 99 from IEEE Xplore, 29 from ACM Digital Library, 20 from Wiley Online Library, and 19 from WorldCat, resulting in a total of 357 publications.

After automatically filtering duplicates using EndNote and performing a manual review, 92 duplicates were removed. The remaining documents were then sorted by citation count, with the 25 most cited articles individually reviewed for direct relevance. Of these, 11 were deemed directly suitable to address SRQ2. The others were used for snowball sampling to identify additional relevant studies, particularly those of conceptual or historical significance.

5.2 NLR Results

A total of 37 articles were ultimately included in this analysis. 11 of these were identified through the direct NLR process, five had previously been collected, and the remaining 21 were obtained by backward and forward snowballing from key references. The distribution of the years of publication for these articles is shown in Figure 5.1. Furthermore, the papers used, along with their objectives and main focus area for this chapter, are given in Table 5.1

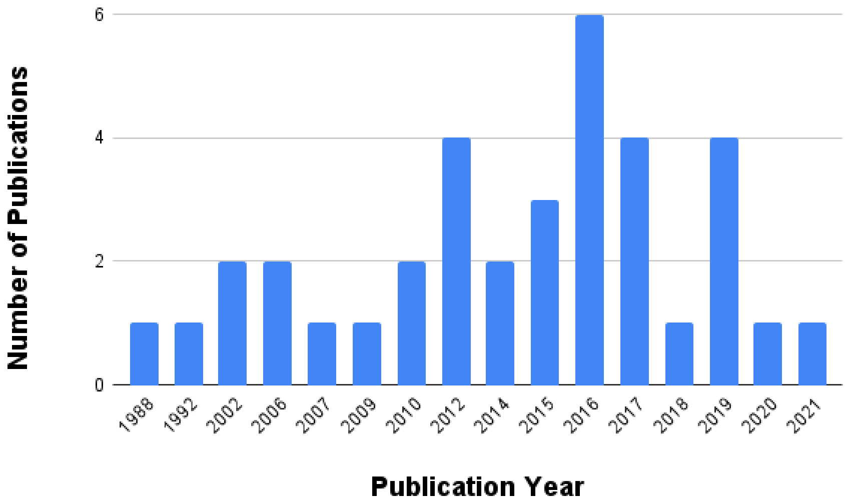

Figure 5.1: Publication Overview of used research

5.2.1 General Findings

To effectively tailor intervention programmes at the individual level, it is essential to account for how ISA interacts with individual differences (McCormac et al., 2017). Alotaibi et al. (2016) identifies several human factors from the prior literature that influence user behaviour in relation to compliance with ISPs, including *personality traits*, *cultural background*, and *gender*. This list is further expanded by the findings of Darwish et al. (2012) and Parker and Flowerday (2020), both of whom investigate the characteristics associated with susceptibility to phishing. Their research highlights additional relevant factors such as *age*, *level of education*, *Internet usage patterns*, and *ICT knowledge*.

5.2.2 Personality Traits

Personality is a widely recognised individual difference that influences ISA. As a relatively stable and inherent psychological construct, personality can affect the way individuals perceive and respond to security threats, thereby influencing the effectiveness of SETA programmes (Aldawood and Skinner, 2019). Kajzer et al. (2014) found a significant correlation between personality traits and the efficacy of security awareness messaging, suggesting that tailoring awareness initiatives to user personality profiles and thinking styles could lead to more impactful behavioural change. Similarly, Khando et al. (2021) argue that a deeper understanding of employee personality traits can help organisations proactively mitigate insider threats and strengthen ISA results.

Table 5.1: Objectives and Focus Areas of Studies Included in the NLR

Reference	Objective	FA
Albladi and Weir (2017)	Model indirect effects of Big Five traits on cyberattack vulnerability	PT
Aldawood and Skinner (2019)	Highlight ongoing issues and pitfalls of social engineering attacks	PT
Alotaibi et al. (2016)	Identify organisational and individual-level factors influencing compliance behaviour	DG
Alseadoon et al. (2015)	Examine how personality influences susceptibility to phishing	PT
Anderson and Agarwal (2010)	Explore cognitive and psychological determinants of safe computing intentions among home users	WL
Bulgurcu et al. (2010)	Assess how cost–benefit beliefs and ISA influence compliance	DG
Cho et al. (2016)	Examine the effect on personality traits on secure behaviour	PT
Chua et al. (2018)	Analyse how demographics shape ISP awareness and compliance	DG
Darwish et al. (2012)	Profile phishing victims using demographic & behavioural indicators	DG
Farooq et al. (2015)	Investigate influence of demographics and training on ISA levels	DG
Halevi et al. (2016)	Examine how culture, personality traits, and demographics affect cybersecurity behaviour	PT & DG
Hadlington et al. (2019)	Examine how work locus of control and work identity influence ISA	WL
Henshel et al. (2016)	Propose an ontology to integrate cultural attributes into models of human factors in cyberattacks	DG
Hovav and D'Arcy (2012)	Analyse deterrence-based security countermeasures across cultures	DG
Jagatic et al. (2007)	Assess effectiveness of phishing using social network-based spoofing	DG
Jeong et al. (2019)	Identify and consolidate human factors influencing cybersecurity behaviour	DG
Junglas and Spitzmuller (2006)	Explore how personality traits affect privacy concerns	PT
Kajzer et al. (2014)	Explore influence of personality traits on the effectiveness of ISA themes	PT
Khando et al. (2021)	Systematically review methods and factors for enhancing ISA in private and public organisations	PT
Lawson et al. (2017)	Explore interactions between personality traits and persuasion tactics in phishing emails	PT
Li et al. (2019)	Test how cybersecurity policy awareness and organisational cues affect current employee behaviour	DG
McCormac et al. (2017)	Examine the relationship between ISA and individual difference variables	PT & DG
McCrae and John (1992)	Present the Five-Factor Model for describing personality traits	PT
Moody et al. (2017)	Identify situational and personality factors to explain the susceptibility of individuals to phishing	PT & DG
Öğütçü et al. (2016)	Investigate risky behaviours threatening information security	DG
Parker and Flowerday (2020)	Identify factors that contribute to an increased phishing susceptibility	PT
Parrish Jr et al. (2009)	Develop a personality-based model for predicting phishing susceptibility	PT
Pattinson, Jerram, et al. (2012)	Investigate behaviour response of computer users when exposed to phishing	PT
Pattinson, Butavicius, et al. (2015)	Identify factors contributing to individuals' secure behaviour and awareness	PT & DG
Salgado (2002)	Conduct a meta-analysis on how Big Five traits predict counterproductive workplace behaviours	PT
Schmidt et al. (2016)	Identify cultural differences across national, organisational, and individual levels	DG
Sheng et al. (2010)	Analyse how age and gender influence phishing susceptibility	DG
Shropshire et al. (2006)	Apply the Five-Factor Model to understand its predictive power for IT security behaviour	PT
Smahel et al. (2012)	Examine the relationship between online friendships, internet communication styles, and internet addiction	DG
Spector (1988)	Develop and validate the Work Locus of Control Scale for workplace control belief assessment	WL
Talib et al. (2010)	Examine how users acquire and apply information security knowledge in both organisational and home settings	WL
Uebelacker and Quiel (2014)	The influence of personality traits on social-engineering susceptibility	PT
L.-f. Zhang (2002)	Investigate the distinctiveness and overlap between thinking styles and personality traits	PT

FA = Focus Area; PT = Personality Traits, DG = Demographics, WL = Work Location

The most prevalent framework for measuring and interpreting personality traits in psychological and behavioural research is the Five-Factor Model, commonly referred to as the Big Five (Kajzer et al., 2014; Shropshire et al., 2006). This model is widely regarded as both comprehensive and generalisable across disciplines, including ICT security research (Uebelacker and Quiel, 2014). It includes five core dimensions (Khando et al., 2021; McCrae and John, 1992; Uebelacker and Quiel, 2014):

- *Openness to Experience* - curious, adventurous, creative, and flexible;
- *Conscientiousness* - self-disciplined, organised, dependable, and rule-abiding
- *Extraversion* - sociable, energetic, outgoing, and ambitious
- *Agreeableness* - kind, likeable, compassionate, trustworthiness, and compliance
- *Neuroticism* - worrisome, angry, insecure, impulse-ridden, and anxiety

Openness to Experience: Individuals high in openness tend to be imaginative and actively seek new experiences. Junglas and Spitzmuller (2006) found that this openness is related to how individuals assess risk, with more open individuals potentially underestimating security threats. Alseadoon et al. (2015) observed that such individuals may be more likely to comply in unsafe situations, possibly due to a higher threshold for perceived risk or an underestimation of the risk of becoming a potential target. However, these findings are not universally supported, as Pattinson, Butavicius, et al. (2015) reported that higher openness was correlated with better security behaviours. Lastly, McCormac et al. (2017) found no significant relationship between openness and ISA, either positive or negative, highlighting the ongoing debate in the literature.

Conscientiousness: This trait is consistently associated with more secure behaviour. Conscientious individuals are diligent, cautious, and rule-abiding, traits that reduce risk in security contexts (Darwish et al., 2012; Halevi et al., 2016). In particular, Halevi et al. (2016) identified conscientiousness as the most predictive trait of secure behaviour, a finding repeated in the literature on workplace behaviour (Salgado, 2002). Furthermore, Pattinson, Butavicius, et al. (2015) found that conscientious individuals, in addition to agreeable and less impulsive individuals, displayed more cautious behaviour, especially when unfamiliar with digital systems.

Extraversion: Extraverted individuals, who are socially expressive and open to interaction, have shown a mixed relationship with ISA. Studies by Lawson et al. (2017) and Albladi and Weir (2017) revealed that high levels of extraversion were linked to increased susceptibility to phishing, potentially due to the more trusting nature of extroverted people. However, other research has found no significant association between extraversion and ISA, and in some cases, even a modest positive correlation with performance in secure behaviours (Halevi et al., 2016; McCormac et al., 2017; Pattinson, Jerram, et al., 2012). These discrepancies suggest that extraversion's influence may be context-dependent.

Agreeableness: Agreeable individuals prioritise social harmony and may be more susceptible to deception in social engineering contexts (Uebelacker and Quiel, 2014). Their tendency to trust others can impair their ability to perceive malicious intent. Cho et al. (2016) and Alseadoon et al. (2015) demonstrated that high agreeableness significantly increased susceptibility to phishing, especially when combined with openness and extraversion. Likewise, Darwish et al. (2012) found a general trend of higher risk among agreeable individuals. Although some studies suggest that agreeableness can contribute to secure behaviour when combined with conscientiousness, this is often in contexts that do not directly assess susceptibility (Parker and Flowerday, 2020).

Neuroticism: Emotional instability and increased stress sensitivity are reflected in neuroticism and have been primarily associated with increased security-conscious behaviour. Individuals high in neuroticism may act more cautiously due to anxiety, which reduces their vulnerability to threats such as social engineering (Cho et al., 2016; Uebelacker and Quiel, 2014). Moody et al. (2017) and Parrish Jr et al. (2009) suggest that this caution is related to lower self-efficacy and a preference for strict rules, fostering compliance. Similarly, L.-f. Zhang (2002) found that neurotic individuals prefer structured environments and are more likely to adhere to established protocols. This shows how neuroticism is negatively associated with self-efficacy (Halevi et al., 2016). In the context of social networks, Albladi and Weir (2017) reported that neuroticism, along with conscientiousness and agreeability, significantly reduces the susceptibility to phishing.

5.2.2.1 Summary of Personality Trait Findings

The influence of personality traits on information security behaviour reveals a nuanced and, at times, inconsistent picture. Among the Big Five, conscientiousness emerges as the most consistently positive predictor of secure behaviour, with high levels associated with diligence, rule-following, and reduced susceptibility to threats. Neuroticism, often associated with anxiety and lower self-efficacy, appears to raise cautiousness and compliance, leading to generally more secure actions in various contexts.

In contrast, agreeableness and extraversion are more commonly associated with increased risk, particularly in the context of social engineering and phishing, where trust and sociability may impair threat detection. However, some findings suggest that these traits may support security when coupled with conscientiousness.

The role of openness to experience is the most contested. Although some studies associate it with an increased risk due to curiosity and risk tolerance, others report positive correlations with secure behaviour or no significant association at all. These inconsistencies suggest that the impact of openness may depend heavily on situational factors or the specific nature of the security behaviour assessed.

Overall, the literature emphasises the value of incorporating personality assessments into ISA interventions. Tailoring awareness efforts to individual personality profiles, particularly by reinforcing behaviours among conscientious or neurotic individuals and mitigating risks among highly agreeable or extraverted users, could improve programme effectiveness and reduce security vulnerabilities.

5.2.3 Demographics

The second employee-related aspect is demographics. Demographic factors represent a broad and diverse set of variables that can influence ISA and behaviour. To improve clarity and interoperability, the most prominent demographic dimensions identified in the literature are discussed in categories.

Gender and Age: The literature offers mixed findings on the impact of gender on ISA. Farooq et al. (2015) and Jagatic et al. (2007) suggest that men tend to have greater security knowledge, while women are more frequently targeted or fall victim to cyberthreats. In contrast, McCormac et al. (2017) report that women achieve higher ISA scores than men, challenging prior assumptions and indicating the need for context-sensitive interpretation.

With respect to age, there is more consensus. Multiple studies have found that individuals between the ages of 18 and 25 are particularly vulnerable to cyberthreats, including phishing (Sheng et al., 2010). A near-linear relationship has been observed between increased age and higher levels of ISA (McCormac et al., 2017; Pattinson, Butavicius, et al., 2015). An explanation offered for this trend is the frequency of internet usage. Younger individuals are typically more active online, increasing their exposure to potential threats (Smahel et al., 2012). Another possible factor is experience, with studies suggesting that exposure to previous incidents increases awareness and enables individuals to more effectively mitigate future risks (Li et al., 2019; Moody et al., 2017). Furthermore, Darwish et al. (2012) found that personality traits such as agreeableness, which tend to be more pronounced in younger individuals, particularly women, can contribute to their increased susceptibility.

However, not all studies support the importance of gender and age as predictive factors. Moody et al. (2017), for instance, found that neither demographic factor had a significant effect on secure behaviour. Such inconsistencies may be partially explained by cultural variability, as the impact of gender and age appears to differ between national contexts (Hovav and D'Arcy, 2012).

Culture: Beyond individual traits, national culture also plays a critical role in shaping security-related attitudes and behaviours. Although organisational culture was addressed in the previous chapter, national culture concerns values and norms shared within a geographic or ethnic group (Jeong et al., 2019; Schmidt et al., 2016). For example, individualistic cultures such as the U.S. may approach cybersecurity with greater emphasis on personal responsibility, while collectivist cultures, such as China, may prioritise conformity and group norms (Alotaibi et al., 2016).

Empirical evidence from cross-cultural studies reinforces this view. Hovav and D'Arcy (2012) demonstrated that the perceived effectiveness of security countermeasures differed significantly between U.S. and South Korean participants. Similarly, Henshel et al. (2016) argued that national cultural traits influence employees' security efficacy, performance, and compliance behaviour. They propose that understanding cultural variations could enable organisations to better predict insider threats and customise their cybersecurity strategies accordingly.

In contrast, a study by Halevi et al. (2016) involving participants from the U.S., India, the UAE, and Ghana found that although culture significantly predicted privacy attitudes, it had little influence on self-efficacy or actual security behaviour. This divergence in findings highlights the complexity of cultural variables and suggests that cultural effects may be more indirect or context-dependent.

In support of this, Chua et al. (2018) found that demographic and cultural factors influence awareness and compliance in various domains, although their impact on ISP-specific behaviours was less conclusive. In a Malaysian context, they did not observe significant differences based on gender or ethnicity. Still, they noted that age, industry type, and level of education were significantly associated with compliance levels, especially as individuals matured and gained professional experience.

Education and Technological Knowledge: Education and knowledge of ICT are frequently cited as enabling factors for secure behaviour. Bulgurcu et al. (2010) proposed that higher educational level and technical proficiency positively influence compliance intentions with ISPs. In addition, the industry in which an employee operates was also believed to influence compliance. Employees in industries more vulnerable to security-related crimes were expected to be more compliant. However, their empirical results did not reveal a significant predictive value for these control variables, suggesting that other factors may moderate the relationship.

In contrast, Öğütçü et al. (2016) found a clear positive correlation between education level and ISA. Participants with higher educational levels demonstrated reduced risk-taking behaviours and greater compliance with security protocols. These findings indicate that the level of education can serve as a protective factor, likely due to its association with increased cognitive ability, digital literacy, and risk awareness.

5.2.3.1 Summary of Demographic Factors

The influence of demographic factors on ISA and secure behaviour is complex and, in some cases, inconclusive. Age appears to be the most consistent predictor, with older individuals generally demonstrating higher ISA levels and more cautious behaviour. Accumulated experiences and reduced exposure to risk-prone online activities likely explain this trend. In contrast, gender presents mixed findings, with some studies suggesting higher vulnerability among women and others reporting higher ISA performance in female participants. These contradictions underscore the need for contextual and possibly culturally sensitive interpretations.

Cultural background also plays a nuanced role. Although some studies indicate that national culture significantly affects compliance and effectiveness of security programmes, others suggest that its impact may be limited to privacy attitudes rather than concrete behaviour. This variability implies that culture likely interacts with other factors, such as personality, policy context, and training design.

Finally, education and technological knowledge are generally associated with improved ISA and lower risk behaviour, although the findings are not uniform. While some studies did not find direct predictive value, others highlight education as a key enabler of risk-aware behaviour. In general, demographic variables should not be viewed as deterministic, but rather as contextual factors that, when combined with personality and organisational influences, shape individual responses to cybersecurity threats.

5.2.4 Work Location

The final employee-related aspect is the location of work, which can shape both the level of awareness and the behaviour related to security. A related psychological construct is *Work Locus of Control* (WLCS), which refers to the degree to which individuals believe that outcomes in their work environment are under their personal control (Spector, 1988). Importantly, WLCS should not be confused with *self-efficacy*. Where self-efficacy concerns an individual's belief in their ability to successfully perform a specific task, WLCS reflects a broader explanatory belief about whether outcomes at work are determined by one's actions (internal WLCS) or by external forces such as fate, luck, or management decisions (external WLCS).

Hadlington et al. (2019) found that individuals with an external WLCS showed lower levels of ISA and weaker engagement with security practices. This may stem from the belief that security results are beyond their personal influence. In contrast, employees with a more internal WLCS perceive themselves as active agents in shaping outcomes, which correlates with more proactive security behaviour.

Remote work amplifies these dynamics. Employees operating from home or other decentralised locations often lack direct oversight and are not immersed in the organisational culture of security. As Anderson and Agarwal (2010) explain, home computer users are typically excluded from structured training, lack access to immediate IT support, and may not receive timely updates or reminders about secure behaviour. Their study highlights that such individuals, of different ages, form a significant point of weakness within organisations and are less likely to follow good cybersecurity practices unless interventions specifically target their cognitive and motivational drivers.

Similarly, Talib et al. (2010) found that most ISA knowledge is acquired in formal organisational settings. When employees transition to home environments, the absence of institutional reinforcement often leads to a drop in secure behaviour, despite having previous training. However, when organisations encourage a strong sense of responsibility and ownership, through both messaging and cultural reinforcement, this knowledge can be transferred and maintained outside traditional settings.

These findings emphasise the importance of considering the location of work in ISA initiatives. Remote or decentralised employees may require targeted interventions that improve not only their knowledge but also their sense of control and accountability. Addressing this could help bridge the security gap between on-site and remote staff.

5.3 Interview Results

To complement the literature findings, experts were interviewed and presented with the three main factors believed to influence individual differences in MMs: *Demographics*, *Personality Traits*, and *Work Location*. They were asked to explain which factors they considered the most influential in shaping MMs relevant to ISA and behaviour, and to reflect on the reasons why.

Importantly, all three factors were discussed throughout the interviews, both from high-impact and low-impact perspectives, indicating a rich diversity of viewpoints. Some participants clearly prioritised one factor over another, while others believed that all three had equal *Shared Importance (2/2)*. A complete overview of frequency and thematic coding can be found in Appendix A.3.

Demographics (11/8): Demographics emerged as the influence factor most frequently discussed. Most of the participants expressed strong agreement that demographic characteristics meaningfully shape an individual's MM. Several pointed to the foundational role that demographic differences, such as age, education, gender, and background, play in shaping how individuals perceive and respond to cybersecurity.

P7 stated, *"I'd say demographics, personality traits, and work location, in that order."*, clearly positioning demographics as the dominant driver. Others elaborated on how changing core characteristics like age or education can meaningfully alter one's mental framing: *"[...] if you change their age, gender, industry, education, culture, or background, their mental models will likely be more diverse, even if they share similar personality traits." [P1]*.

Specific elements within demographics were also singled out. Education was found to be particularly impactful, especially in shaping overconfidence: *"I think demographics play a role, especially when the educational level is high. People tend to believe that their knowledge extends to other areas." [P3]*.

Age was the most frequently cited demographic subfactor. Participants linked it to risk perception and behavioural flexibility. For example, one participant explained, *"Younger people may have a different risk perception compared to someone older [...]" [P4]*, while another participant noted that *"[...] older people are less susceptible to change. They need more time and are more set in their old ways. Cybersecurity can feel intimidating for them [...]" [P5]*.

The interplay between age and accumulated experience was also recognised as a key mechanism shaping MMs over time. As one expert stated: *"As you grow older, for example, your experiences in life shape your mental model. Someone who has been in IT security for 50 years will have accumulated a lot of experience." [P6]*.

However, not all participants were equally convinced. One expressed a more critical stance: *"Yes, demographics play a role, but not to the extent that one might think. It's more about how we perceive people who learn differently." [P2]*. This reflected a view that demographic factors may be secondary to cognitive or learning styles.

Personality Traits (7/7): Personality traits were mentioned by all participants, except one. They were frequently described as a significant influence on how people perceive threats and behave in contexts relevant to security. These traits were found to be closely related to the internal cognitive processes that shape the development of MMs.

As summarised by one participant, *"Personality traits also affect how people perceive threats" [P4]*, while another elaborated, *"[...] certain personality traits may lead you to actions that fit your personality but make your work less secure, or might even create an incident unintentionally" [P3].*

Extroversion was the trait most commonly referenced when discussing the influence of MM, often cited by participants who self-identified as extroverted themselves. These individuals acknowledged that social tendencies occasionally lead to premature or insecure actions. For example, one shared, *"I'm a bit of an extrovert myself, so I like to write and share things with people. Sometimes I hit send on an email and later think, "Wait, should I have shared this information with those people? Should I have made it more anonymous or general?" [P3].* Another echoed this, *"I'm extroverted and enjoy interacting with people, which shapes how I view risks, compared to someone who is more introverted and prefers working alone." [P4].*

Openness to experience was also mentioned as a relevant trait. However, participants stressed that personality traits do not operate in isolation. As one participant explained, *"It's a combination of factors, though. It's not just about openness to experience. [...] I think it's a mix of experience and personality traits." [P6].*

Others reinforced this intersectional view. One noted that most of the Big Five traits play a role in secure behaviour, stating, *"I'd say conscientiousness, extroversion, agreeableness, and openness to experience are key factors. People who are curious and question why they're receiving a suspicious email are more likely to be cautious" [P5].*

However, the influence of personality was not accepted indiscriminately. One participant placed its importance below demographics, arguing that, *"[...] you can have two people working in the same company, who are the same age, gender, and education level, and they can still have very different personality traits but hold similar mental models." [P1].* This view suggests that environmental or contextual factors may outweigh individual traits in some cases.

Work Location (6/6): Work location was the least frequently mentioned factor in the literature. However, it was still discussed in more than half of the interviews. The participants were divided in their views, as some saw it as underappreciated, while others regarded it as less influential than the other two factors.

Several experts believed that individuals underestimate how location affects perceived risks and behaviour. For example, one participant commented, *"[...] I don't think people fully appreciate the different risks they face depending on their location" [P3]*, further elaborated with *"People feel more secure at home. In the office, they know they should log off, take their laptop with them, or put it in a locker." [P3].*

A recurring theme was the effect of work location on organisational connectedness. Physical presence was seen to reinforce both formal training and informal exposure to secure practices.

One participant noted that, *"[...] working from home can disconnect employees from the organisational culture and the company itself. I think this, along with mandatory training, influences security awareness significantly." [P6]*. Another added, *"Your work location matters, there's a difference between working from home and being in the office. You encounter different people and conversations in each setting." [P4]*.

However, some participants perceived work location as having a relatively minor role. One stated, *"Work location has the least impact, as long as the technical setup is secure" [P7]*. Interestingly, even sceptical participants acknowledge the benefits of office presence over working from home, *"I think there's not much difference between working from home or the office. However, in the office, you might get more organised and develop more sensitivity to the organisation, since you hear more about it and it's less transactional. You're in the building, and you care more." [P5]*.

5.3.1 Additional Factors

Following the discussion of the three main influence factors, the participants were invited to suggest any other variables they believed might significantly shape individual MMs related to information security. Although most suggestions could be broadly categorised under demographic or personality themes, one topic stood apart, *neurodiversity*

Neurodiversity (3/2): Neurodiversity, a term rooted in contemporary psychology, recognises that cognitive and neurological differences are a natural and valuable part of human variation. It opposes deficit-based perspectives and instead frames conditions such as autism, ADHD, and dyslexia as forms of diversity in how people think, learn, and interact with the world (Baumer N, 2021).

Two participants directly mentioned neurodiversity, highlighting it as an important but often overlooked factor in how people perceive and process security-related information. One expert emphasised the need to expand the discussion beyond conventional cognitive frameworks, noting, *"[...] I think it's about learning ability. In my experience, it's linked to personality and personal experience, including education. But we also need to consider neurodiversity." [P2]*.

Another participant pointed out that neurodiverse individuals may possess unique perceptual and cognitive tendencies that shape their MMs in unexpected ways, *"For example, if someone is neurodiverse, like being autistic, they may have certain perceptions that don't correlate with their knowledge level." [P6]*.

These observations suggest that traditional ISA frameworks may overlook critical differences in how individuals interpret risk signals or internalise behavioural norms. Although not yet well represented in the academic literature on MMs, neurodiversity may warrant dedicated exploration in future research and could form the basis for a more inclusive training design.

5.4 Summary

This chapter addressed SRQ2: *What employee-related aspects need to be considered in Mental Model assessments?* MMs offer a valuable lens to explain individual differences in ISA and secure behaviour. To identify relevant employee-related factors that influence MMs, an NLR was performed. The NLR findings were then presented to information security experts in a series of interviews, allowing them to validate, critique, and expand on the identified factors. The NLR revealed three core factors influencing MM variation, *Personality Traits, Demographics,* and *Work Location*.

In terms of **Personality Traits**, the Big Five framework was used. All five traits were found to influence ISA and secure behaviour. However, the nature of this influence varied. Traits such as *Openness to Experience, Extraversion,* and *Agreeableness* were associated with positive and negative outcomes depending on context. In contrast, *Conscientiousness* and *Neuroticism* were generally associated with more secure behaviour. Expert interviewees largely supported these findings. Extroversion and openness were explicitly mentioned as traits that could undermine secure behaviour. At the same time, several participants emphasised that personality should not be viewed in isolation and may have a lesser impact than commonly assumed.

Demographic factors were also found to be influential, although with more variability in the literature. Key demographic variables included *age, gender, education level,* and *cultural background*. Some studies demonstrated a strong correlation between these variables and ISA, particularly among older people and those with higher educational attainment, demonstrating greater awareness. However, other studies found limited or no significant effect, suggesting that the influence of demographics may be context- or culture-dependent. Experts also reflected this view. Although some participants considered demographics the most critical influence on MM variation, others focused more specifically on age or education, and a few questioned its overall impact.

The third factor, **Work Location**, received less attention in the literature but emerged as increasingly relevant in interviews. Research showed that remote or home workers tend to exhibit lower levels of ISA, often due to reduced exposure to training, support, and organisational culture. Experts echoed this concern, suggesting that people working from home may feel more secure than they actually are, and are more likely to develop habits that deviate from organisational best practices. Still, a minority of experts perceived work location as secondary, provided that technical setups were robust and policies consistent.

In addition to the three main factors, neurodiversity was introduced by multiple participants as a potentially important but underexplored influence. Cognitive and neurological differences, particularly in learning and risk perception, were suggested to shape how MMs are formed. This highlighted the need for more inclusive approaches to ISA.

A visual overview of the three employee-related factors, along with their relevant subfactors, is shown in Figure 5.2. This figure serves to summarise the findings and highlights the multifaceted nature of factors that contribute to individual differences in security-related MMs.

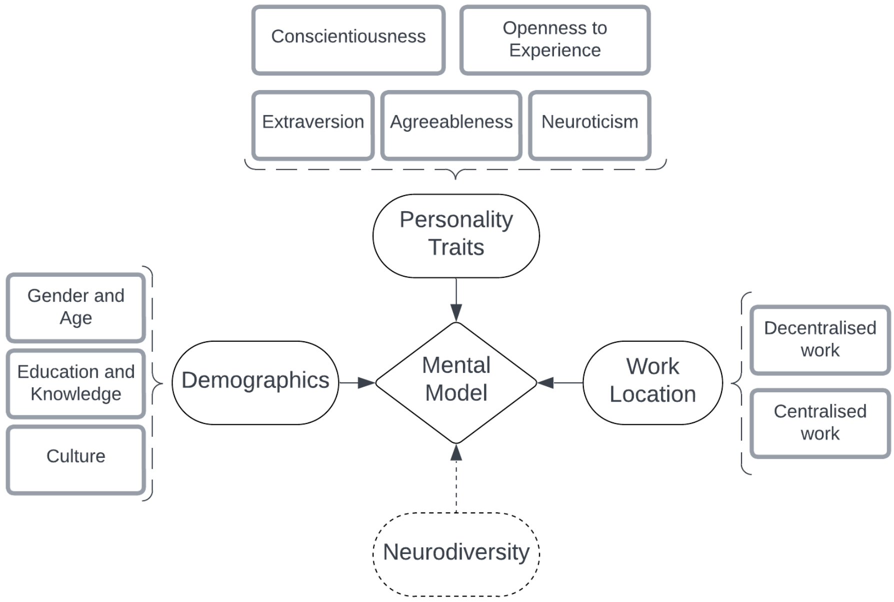

Figure 5.2: Overview of employee-related aspects influencing MM formation

Note. Neurodiversity is shown with a dashed outline to indicate that it was introduced solely during expert interviews and was not identified in the Narrative Literature Review (NLR).

In summary, while each factor, whether personality-based, demographic, or geographic, offers valuable insight, the chapter findings suggest a shared importance among them. Personality traits such as conscientiousness, demographic factors such as age and education, and environmental elements such as work location all contribute to shaping how individuals perceive and engage with security-related decisions. Together, these factors form the broader context in which MMs are developed. Based on the results of both the literature and expert interviews, all three primary factors were considered essential for inclusion in comprehensive MM assessments. In the next chapter, attention turns to SRQ3, focusing on how cybersecurity MMs can be elicited and classified in the context of organisations.

6 Eliciting Mental Models on Security Awareness

Chapters 4 and 5 explored the organisational and employee-level factors that influence ISA, and through it, the MM assessment. Building on this foundation, a critical question remains as to how these MMs can be effectively elected and classified within organisational settings.

To address this, a third NLR was conducted, focused specifically on identifying methods and tools to obtain the components that represent an MM. This literature review is complemented by expert interviews with practitioners and researchers in the field. Together, these insights enable us to address the third SRQ: *How can Mental Models of cybersecurity be elicited and classified in the context of organisations?*

6.1 Proposed Framework

Drawing from the literature and expert interviews, a framework was developed for the elicitation and classification of MMs in a structured and practical way. This framework is visualised in Figure 6.1, and follows the Business Process Model and Notation (BPMN) standards [1].

The process involves two primary organisational roles, *Senior Management* and *Experts*. Senior management is responsible for initiating the process by defining the core focus area and relevant topics. The remainder of the process, including designing assessments, gathering data, and analysing results, is carried out by the expert. Each step within the framework is explained in detail in the subsequent sections, offering practical guidance for implementation within various organisational contexts.

6.2 NLR Process

The NLR approach mirrors the method outlined in Section 4.1. However, a significant portion of the selected literature was obtained through the foundational work by Bitton et al. (2018). Therefore, only four new papers are introduced in this chapter, all focused on the elicitation of MMs. They are retrieved from digital libraries with the help of the following search query:

- "Mental model" OR "Mindset"
- AND measur* OR elicit*
- AND Aware* OR "Security Aware* OR SETA OR ISA

[1] https://leansixsigmagroep.nl/lean-agile-en-six-sigma/bpmn-business-processing-model-notation/

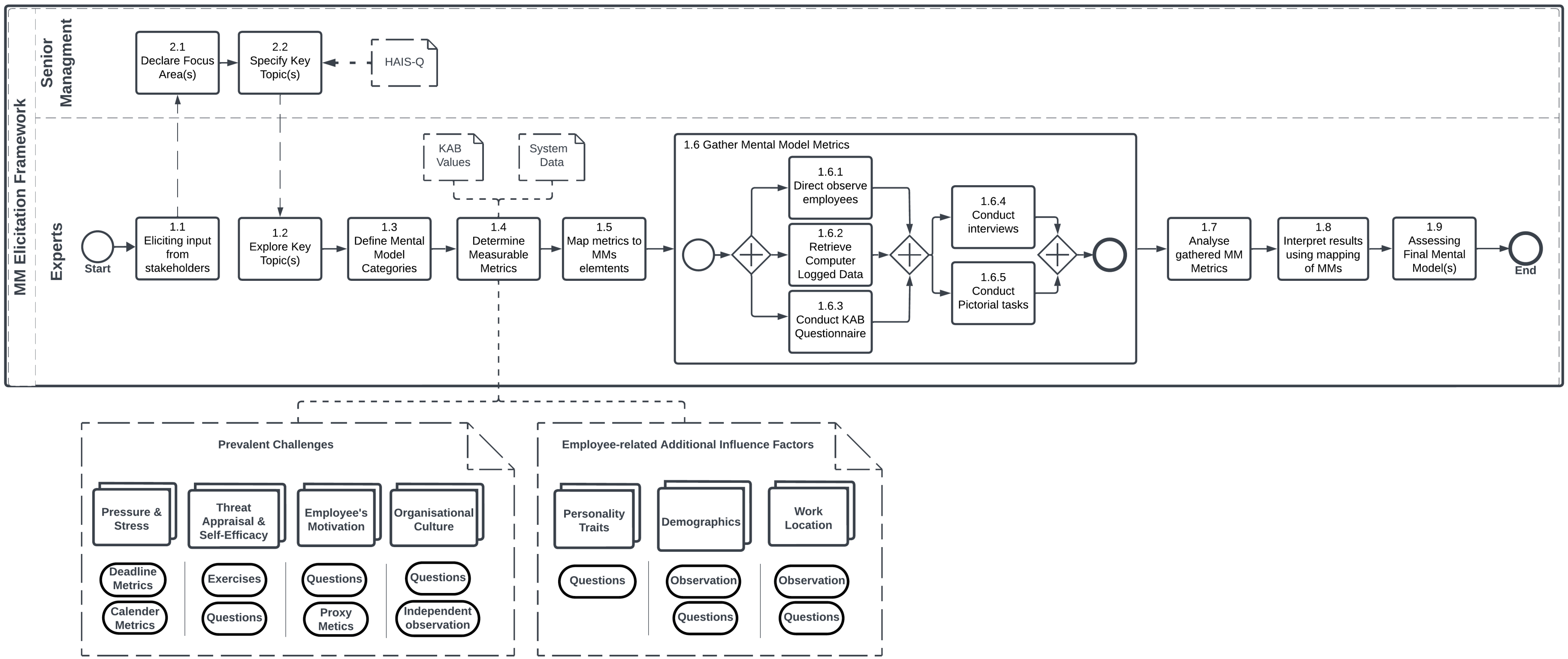

Figure 6.1: Framework for Mental Model Elicitation

6.3 NLR Results

The final selection included 13 articles; of these, seven were identified by snowballing the citations of Bitton et al. (2018), demonstrating the foundational nature of that study in the elicitation of MMs. One article emerged from earlier phases of this research, while the remaining four were newly retrieved during the NLR. The distribution of the years of publication for these articles is shown in Figure 6.2. Furthermore, a summary of the articles used, along with their objectives and main focus area for this chapter, is presented in Table 6.1.

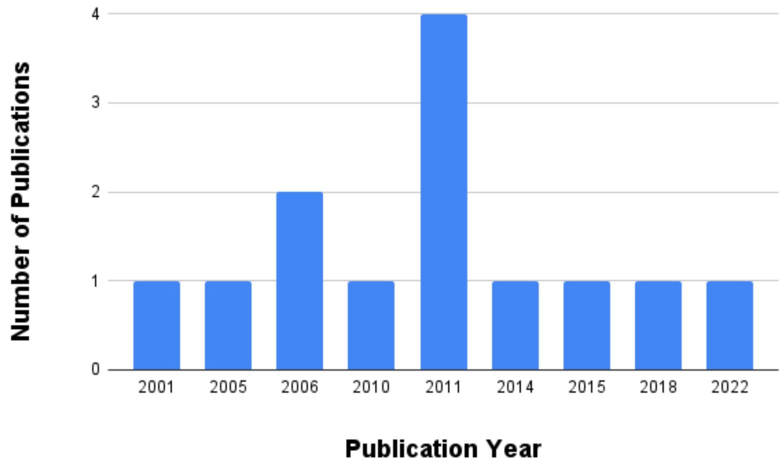

Figure 6.2: Publication Overview of used research

Table 6.1: Goal and Focus of SRQ3 NLR documents

Reference	Objectives	Focus area
Androulidakis and Kandus (2011a, 2011b, 2011c)	Survey on security awareness and practices	ISA Metrics
Bitton et al. (2018)	Develop a hierarchical taxonomy for measuring awareness	ISA Metrics
DeChurch and Mesmer-M (2010)	Meta-analyse methods for measuring shared MMs in teams	ET
E. E. Doyle et al. (2022)	Review methods for MM elicitation in risk communication	ET
Grenier and Dudzinska-P (2015)	Propose a multimethod model for MM elicitation	ET
Jones et al. (2011)	Synthesise MM theory and methods across disciplines"	ET
Kruger, Drevin, et al. (2006)	Develop prototype model for assessing ISA	ISA Scoring
Kruger and Kearney (2006)	Propose framework for ISA evaluation using system data	ISA Scoring
Parsons, McCormac, et al. (2014)	Develop and validate the HAIS-Q for employee ISA	HAIS-Q
Siponen (2001)	Define five dimensions of ISA	Assess ISA
Stanton et al. (2005)	Develop and test taxonomy of end-user security behaviours	Assess ISA

ET = Elicitation Techniques

6.3.1 General Findings

Eliciting MMs is a critical first step in understanding the cognitive structures that underlie individual decision-making and behaviour. Despite its importance, the academic literature provides limited guidance on how to select appropriate elicitation methods and metrics for different contexts or research goals (Grenier and Dudzinska-P, 2015).

This raises the fundamental question of why we should even elicit and classify MMs in the first place. One central reason is to deepen our understanding of the cognitive processes that drive individual decisions (Grenier and Dudzinska-P, 2015). Beyond this, Jones et al. (2011) summarise several broader objectives for MM elicitation:

- To explore similarities and differences between stakeholders' understanding of an issue to improve communication between stakeholders
- To integrate different perspectives, including expert and local, to improve overall understanding of a system
- To create a collective representation of a system to improve decision-making processes
- To support social learning processes
- To identify and overcome stakeholders' knowledge limitations and misconceptions associated with a given resource
- To develop more robust knowledge to support negotiations over unstructured problems in complex, multifunctional systems

Although these purposes are not explicitly related to cybersecurity, their relevance becomes clear when applied in the context of ISA. By eliciting employee MMs, organisations can uncover misalignments in understanding, identify conceptual gaps, and better tailor cybersecurity communications and training. The result is a more nuanced and targeted approach to ISA that fosters improvements in knowledge, attitudes, and behaviours, ultimately strengthening the organisation's security position.

6.3.2 Elicitation Techniques

In fields such as natural resource management, MM elicitation techniques have been widely used. Most of these approaches assume that MMs can be represented as conceptual networks, comprising nodes (concepts) and edges (relationships) (Jones et al., 2011). These techniques are typically classified into two broad categories, *direct elicitation* and *indirect elicitation*.

Direct Elicitation: Direct elicitation, also called *graphical elicitation*, involves asking participants to explicitly create a visual or structured representation of their MM. Standard methods include concept mapping, card sorting, and pictorial tasks (E. E. Doyle et al., 2022; Grenier and Dudzinska-P, 2015). Direct elicitation methods can be further categorised into two main approaches. The first provides respondents with predefined categories of cognitive content to rate, while the second approach induces respondents to provide their MM without using predefined categories (DeChurch and Mesmer-M, 2010).

One of the main strengths of direct elicitation is its potential for immediate verification, as participants can review and clarify their outputs, reducing the risk of misinterpretation by the researcher (Jones et al., 2011).

Within the proposed framework, Activity *1.6.5 Complete Pictorial Task* represents the direct elicitation of MMs. Within Activity *1.7 Analyse gathered MM metrics*, the pictorial output of this elicitation is evaluated by experts by comparing it to a predetermined baseline.

Indirect Elicitation: Indirect elicitation, also known as *verbal elicitation*, seeks to infer an individual's MM based on language, either spoken or written. Typically, it involves open-ended and semi-structured interviews along with case-based discussions (Jones et al., 2011).

These methods enable participants to describe their understanding of a system or issue in their own terms. However, a significant challenge with indirect elicitation is the potential for bias introduced by the interviewer's interpretation (Grenier and Dudzinska-P, 2015). To mitigate this, it is recommended to use a diverse range of question types and ensure methodological consistency between interviews.

Within the proposed framework, Activities *1.6.3 Conduct KAB Questionnaire* and *1.6.4 Conduct Interviews* represent the indirect elicitation of MMs.

Both direct and indirect methods play an essential role in constructing a comprehensive view of how people perceive, understand, and respond to cybersecurity issues. The following sections will examine how these techniques are applied and how the resulting information can be grouped and analysed.

6.3.2.1 Eliciting Mental Models

The MM elicitation process proposed by Grenier and Dudzinska-P (2015) serves as the primary foundation for the approach adopted in this study. Their framework outlines a three-step method designed to construct a holistic understanding of an individual's cognitive structure through both direct and indirect measures.

Modified Twenty Statements Test (TST): The first step involves administering one or more modified versions of the TST. Developed initially to explore self-concept, the TST asks participants to answer a series of open-ended questions. This format is particularly well-suited for capturing the subjective dimensions of MMs, such as knowledge, attitudes, and behaviour. In the context of cybersecurity MMs, the TST can be adapted to explore perceptions of risk, personal responsibility, and security practices. The open-ended responses from this step serve as a rich source of data and are used to inform the development of the second step, semi-structured interviews.

Within Figure 6.1 Activity *1.6.3 Conduct KAB Questionnaire* takes inspiration from this first step by Grenier and Dudzinska-P (2015).

Semi-Structured Interviews: The second step involves conducting SSIs, which are guided and enriched by insights derived from the responses of the TST. This design enables interviewers to more deeply explore participants' understanding of key topics while also exploring how they articulate the relationships between key topics.

The first round of interviews is typically used to explore general attitudes and perceptions around a defined topic. A potential second round of interviews, preferably using a think-aloud protocol, can provide additional insights into the participants' reasoning processes and decision strategies. To capture a higher level of cognitive complexity, this second round often integrates concept mapping as an elicitation tool.

Within the proposed framework, Activity *1.6.4 Conduct Interviews* represents this step. Although there is a difference in the objective and method between each round of interviews, conducting interviews will remain a singular activity for visual simplicity.

Concept Mapping: The third and final step involves the direct elicitation of concept mapping, allowing participants to visually represent key concepts and the perceived relationships between them that they associate with a given topic. This graphical format provides a valuable contrast to the more abstract verbal responses obtained in earlier steps.

The proposed framework represents the direct elicitation of MM data as Activity *1.6.5 Complete Pictorial Task*. Activity *1.7 Analyse gathered MM metrics* is part of the direct technique, as the pictorial output of this elicitation is evaluated by experts.

By combining questionnaire-based, interview-based, and graphical elicitation techniques, this multiangle approach increases the likelihood of capturing a robust and multifaceted representation of an individual's MM.

Specifying the Focus Area: One additional, but essential, step not explicitly addressed in the elicitation method by Grenier and Dudzinska-P (2015), is the prior specification of the focus area and key topics. Defining the thematic scope of the elicitation process is crucial to ensuring relevance and consistency across participants. It shapes the design of questionnaires, guides interview questions, and contextualises concept mapping exercises.

In this research, the overarching focus area is *IT security in organisational settings*, with the key topic being *Shadow IT*. These domains are not only relevant but also prone to variations in employee understanding, making them ideal for MM assessment.

The specification of focus areas and key topics is found in the proposed framework as Activities *1.1 Eliciting input from stakeholders, 1.2 Explore Key Topic(s), 2.1 Declare Focus Area(s),* and *2.2 Specify Key Topic(s).*

6.3.2.2 Measuring Mental Models Proxies

Having outlined the methods for eliciting MMs, the next step involves determining how these outputs, whether taken from questionnaires, interviews, or pictorial representations, can be meaningfully measured. While elicitation reveals the internal structure of thought, measurement translates the corresponding factors into actionable and comparable data points.

One main goal in MM assessments in the context of cybersecurity is the measurement of ISA. ISA can be broken down into five dimensions, with *organisational ISA* being the most relevant to this context (Siponen, 2001). This dimension provides a shared reference point for evaluating the outcomes of the MM elicitation.

To determine how ISA-related constructs such as knowledge, attitudes, and behaviours can be quantified, several established approaches were reviewed. These are summarised in Table 6.2, which provides an overview of the studies in the NLR used to address ISA measurement.

Table 6.2: Comparison of ISA Models and Evaluation Studies

Paper	Domain	Dimension K	A	B	Focus Area Approach	ISA Data Gathering	ISA Measuring Method	ISA Level Determination
Androulidakis and Kandus (2011a, 2011b, 2011c)	MSA	✓	✓	✓	Undisclosed	Questionnaire	Mean calculation	Linear Mean weight
Bitton et al. (2018)	MSA	✓	✓	✓	Reviewing attack cases, databases, and news	Interviews, Questionnaire, Observing	Scorecard approach	Weighted sum of values
Kruger, Drevin, et al. (2006)	OSA	✓	✓	✓	Value Focused Approach	Questionnaire	Scorecard approach	Weighted sum of values
Parsons, McCormac, et al. (2014)	OSA	✓	✓	✓	Interviews, IS survey and reviewing ISPs	Questionnaire	Mean calc. of KAB dimensions	Summed mean of KAB values
Siponen (2001)	GSA	✓	✓	✓	Undisclosed	Not applicable	Not applicable	Not applicable
Stanton et al. (2005)	OSA	✓	✓	✓	Interviews	Questionnaire	Categorical placement	Unweighted

MSA = Mobile Security Awareness, OSA = Organisational Security Awareness, GSA = General Security Awareness, K = Knowledge, A = Attitude, B = Behaviour

Assessing ISA Through KAB: ISA is commonly conceptualised through the KAB dimensions: Knowledge (what an individual knows), Attitude (how they feel or perceive risk), and Behaviour (how they act) (Siponen, 2001). This model fosters many established measurement strategies. One taxonomy, developed by Stanton et al. (2005), categorises end-user behaviour based on two dimensions: *technical expertise* and *intentionality*. This results in a behavioural taxonomy that classifies actions, from well-intentioned mistakes to malicious breaches. Expert-rated surveys assess behaviours on a scale of maliciousness (1-5) and required expertise (1-5). The results produce a modal map of the behavioural categories, providing insight into user motivations and capabilities.

A subset of behavioural metrics, focused on password practices, is shown in Table 6.3, alongside the original behaviour taxonomy in Figure 6.3. These metrics, obtained through different assessment techniques such as self-reported surveys, monitoring, and activity logging, can provide immediate indicators of user intent and competency.

Within the proposed framework, this process is reflected in several activities. First, *1.5 Map metrics to MM elements* reflects how the measured metrics will be categorised based on the different MMs. Furthermore, *1.6.1 Direct observe employees*, *1.6.2 Retrieve Computer Logged Data*, and *1.6.3 Conduct KAB Questionnaire* are all available options for collecting the MM metrics. Lastly, *1.8 Interpret results using mapping of MMs* deals with actually placing the gathered results into one of the predetermined MMs.

Table 6.3: Metrics for Password Behaviour, taken from Stanton et al. (2005)

Examples of Measurable Password Metrics
Frequency of changing passwords
Password quality
Writing down passwords
Password sharing

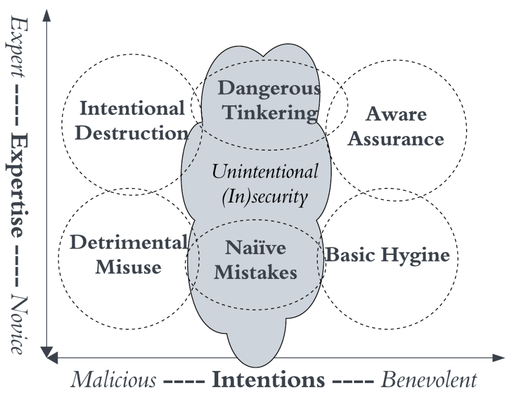

Figure 6.3: End-User Security Behaviour Taxonomy, adopted from Stanton et al. (2005)

Scoring ISA Through KAB: Early work by Kruger and Kearney (2006) introduced a structured prototype to measure ISA based on KAB dimensions; see Figure 6.4. In subsequent research, this model combines survey instruments with system-generated behaviour data, as shown in Figure 6.5. The advantages of system data are that they are easy to obtain and more reliable, as there is a more objective view of user awareness (Kruger, Drevin, et al., 2006). Metrics such as virus infection counts, password strength, and response to phishing attempts offer observable proxies for ISA, see Table 6.4.

To structure ISA measurements, Kruger proposes a *Value-Focused Approach (VFA)* that begins with stakeholder input to identify critical areas of concern. These focus areas are then broken down into subfactors within a hierarchical value tree. The goal is to structure complex ISA challenges into measurable elements that align with organisational priorities.

The VFA was taken as inspiration for the specification of focus areas and key topics. This starts with Activity *1.1 Eliciting input from stakeholders*, and continues with Activities *2.1 Declare Focus Area(s)*, *2.2 Specify Key Topic(s)* and *1.2 Explore Key Topic(s)*.

The measurement approach includes survey-based questions (true/false and 3-point Likert scales) and system monitoring data, culminating in a simple weighted scorecard approach, defined as: $V(a) = \sum_{i=1}^{n} v_i(a)w_i$. This approach includes the overall value, performance, and importance of a factor. As not all factors contribute equally, the Analytical Hierarchy Process (AHP) was used to determine factor weight based on expert judgement. Such scoring models are well-known management science techniques, with this additive model being one of the most widely used forms of a value function.

This scoring method, to measure the metrics obtained, is integrated in Activity *1.7 Analyse gathered MM Metrics* in the framework.

Table 6.4: System Data Metrics, taken from Kruger, Drevin, et al. (2006)

Examples of System Data Metrics	
Number of virus infections	Password strength
Requests to visit unauthorised websites	Backups on network drives
Number of IT security incidents	Downloading files from the Internet
Reactions on phishing e-mails	Incidents reported
Installation of anti-virus software	Compliance to security policies

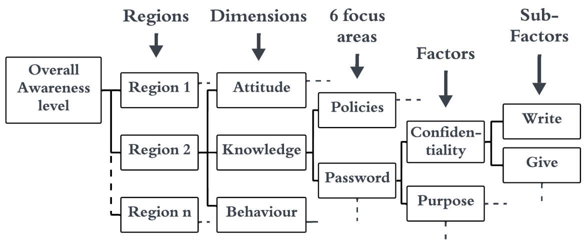

Figure 6.4: Tree structure of problems, adopted from Kruger and Kearney (2006)

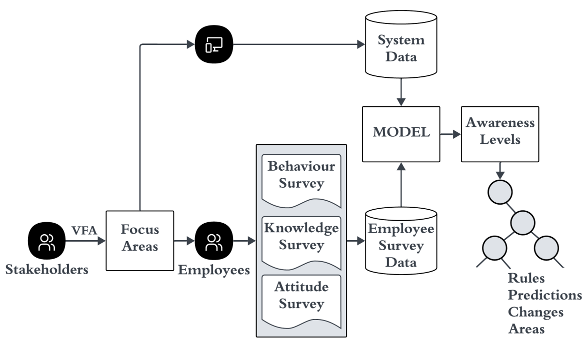

Figure 6.5: Proposed Framework, adopted from Kruger, Drevin, et al. (2006)

ISA Measurable Metrics: Bitton et al. (2018) proposes an expansive hierarchical taxonomy with more than 150 KAB-related criteria to measure ISA, particularly in mobile contexts. These criteria are mapped to specific subfocus areas, such as application handling, each tied to relevant psychological components such as risk perception or cognitive bias. The focus areas are defined based on previous research on security awareness, attack cases, attack databases, and security-related topics in the news.

Selected criteria used to assess knowledge and attitude regarding application usage are shown in Table 6.5, while the broader taxonomy structure appears in Figure 6.6. Knowledge and attitude data can be collected using interviews and questionnaires, while behavioural metrics are best collected through monitoring or system logs. Like Kruger, Bitton employs AHP to prioritise and weigh criteria.

Although Bitton's approach to defining focus areas was not adopted, their measurement strategy significantly influenced Activities *1.6.1 Direct observe employees, 1.6.2 Retrieve Computer Logged Data, 1.6.3 Conduct KAB Questionnaire* and *1.6.4 Conduct Interviews* of the framework.

A series of studies by Androulidakis and Kandus (2011a, 2011b, 2011c) further validate the assessment of mobile ISA through questionnaires. The five-point Likert questionnaire obtained both demographic data and KAB indicators. The calculated metrics, the Mean Security Awareness Value (MSAV) and the Mean Security Feeling Value (MSFV), are derived from Likert scale responses and mathematical averaging.

Table 6.5: Metrics for Application Installation and Handling, taken from Bitton et al. (2018)

Examples of Application Metrics	
Knowledge	**Attitude**
Knowledge of malicious applications	Perceiving installing from untrusted sources as dangerous
Knowledge on managing application permissions	Perceiving the existence of malicious applications plausible
Knowledge on how to configure privacy settings	Perceiving proper management of permissions as useful
Knowledge on validating the source of an application	Perceiving installing unknown applications as dangerous
Knowledge of application vulnerability causes	Perceiving using an outdated application as dangerous
Knowledge of application update importance	Perceiving that some updates potentially harm the device

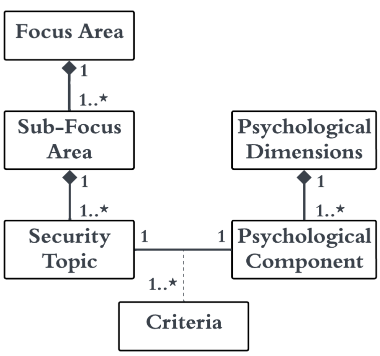

Figure 6.6: Class diagram of the main building blocks of the security awareness taxonomy, adopted from Bitton et al. (2018)

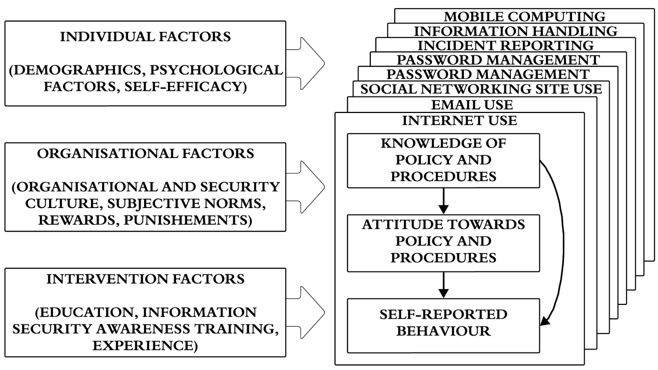

Figure 6.7: The HAIS model, adopted from Parsons, McCormac, et al. (2014)

Table 6.6: Metrics for Mobile Phone Security Awareness, taken from Androulidakis and Kandus (2011a)

Examples of Measurable Metrics		
Demographics	**Security**	
Gender	Knowledge about O/S	Backup behaviour
Age	Knowledge of IMEI	Antivirus behaviour
Study/Career	Password behaviour	Antivirus knowledge
No. of stolen phones	Bluetooth behaviour	Feeling secure
	Lending behaviour	Feeling informed
	Download behaviour	Feeling aware

Table 6.6 provides representative metrics, focused mainly on the dimensions of knowledge and behaviour. They include user practices such as the use of passwords, the frequency of backups, and the lending behaviour of mobile devices. These metrics also support the inclusion of demographics as an additional employee-related influence factor within the framework.

Activity *1.6.3 Conduct KAB Questionnaire* in the framework is further strengthened by these findings. In addition, their approach also reflects in the *Employee-related Additional Influence Factor* of *Demographics*.

HAIS-Q: The Human Aspects of Information Security Questionnaire (HAIS-Q) developed by Parsons, McCormac, et al. (2014) provides another validated approach to assessing ISA, and to examine the relationship between knowledge and attitude towards policy and procedures, and computer behaviour at work. Based on interviews with senior leadership and a review of ISPs, HAIS-Q identifies seven high-risk focus areas most prone to non-compliance, visible in Figure 6.7, and develops a KAB-based statement to assess each, see Table 6.7.

This tool is primarily focused on "neutral behaviours", behaviours associated with human errors not intended to harm the organisation or its resources. Stanton et al. (2005) defines this type of behaviour as "naïve mistakes". The structure of HAIS-Q includes 63 KAB statements and 10 additional personality measures, which contribute to a robust awareness profile.

Similarly to the work by Androulidakis and Kandus, the study of Parsons is used to further exemplify the importance of Activity *1.6.3 Conduct KAB Questionnaire*. Also similar is how a *Employee-related Additional Influence Factor* from the study is incorporated into the framework, being *Personality Traits*.

Table 6.7: Behaviour Metrics for HAIS-Q Areas, taken from Parsons, Calic, et al. (2017)

Examples HAIS-Q Metrics		
Password Management	**Internet Use**	**Incident Reporting**
Using same password	Downloading files	Reporting suspicious behaviour
Sharing passwords	Accessing dubious websites	Reporting all incidents
Using strong passwords	Entering information online	Ignoring poor security behaviour by colleagues

6.4 Interview Results

To complement the findings of the literature review, a series of expert interviews was conducted to explore practical considerations in eliciting and classifying MMs. Participants were asked about the validity, utility, and limitations of the various measurement techniques. Their responses provide helpful insight into how the theoretical methods translate into real-world assessment contexts.

KAB Suitability (7/6): Before evaluating specific elicitation techniques, participants were asked about the usefulness of the KAB dimensions in assessing MMs. This discussion is rooted in the broader literature linking KAB to ISA, such as in the work of Siponen (2001), and extended by models such as the KAB-based ISA measurement approach proposed by Kruger and Kearney (2006). Since ISA itself has been linked to MMs, particularly in shaping employee decision-making and secure behaviour, understanding whether KAB is suitable for MM assessment is critical.

The opinions of the participants reflected a mix of practical support and theoretical caution. Some supported the notion that MMs could be interpreted through KAB-aligned constructs, particularly in the context of designing interventions: *"I think they could. Yes, I think so. We use the Knowledge, Motivation, and Opportunity (KMO) model because it's easier to map interventions to it. But if you can align interventions with the mental models and their outcomes, then yes, you could also use those KAB dimensions. No issue with that." [P5].*

Others were more sceptical about the conceptual compatibility of KAB, especially attitude and behaviour, with MMs, suggesting that MMs are inherently more cognitive and knowledge-oriented: *"If you consider a mental model that only represents a knowledge point from an individual, then I don't think you can really measure attitudes in that context as well. I believe mental models can mainly be used to measure knowledge. I don't think mental models are a good way to measure attitude; there are better methods for that. As for behaviour, people may claim they act in a certain way, but in reality, they may act differently. That's also more challenging." [P1]*

One participant provided a more balanced perspective, recognising the potential utility of KAB but warning of the complexity involved, mainly due to cultural or contextual modifiers that may affect how each dimension emerges: *"You might need to create a mapping of them. I think it's an interesting idea, but it's complex. For example, cultural influences affect all three dimensions in different ways. So it would be useful, but it would also be quite complicated to map out" [P7].*

These insights mirror the tensions in the literature between psychological precision and practical application. Although MMs and KAB may not map perfectly, particularly in terms of attitude or behavioural indicators, they remain interrelated components of how individuals perceive, interpret, and act within a security context. Given these reflections, the KAB structure is maintained as a guiding method for MM elicitation in this study, but with awareness of its limitations.

These reflections establish that, while KAB dimensions are conceptually useful in structuring MM assessments, particularly for ISA, they introduce challenges in practical measurement, especially attitude and behaviour. This naturally leads to a discussion of the tools most commonly employed to measure these dimensions. Among them, surveys are perhaps the most prevalent tools used to capture MM components.

Survey Suitability (10/8): Participants widely acknowledged surveys as a standard and accessible method to elicit MM components, especially knowledge and attitudes. Several interviewees recognised the utility of surveys to gather general insights: *"[...] if you want to measure mental models, I think we're doing a decent job with our annual survey. It contains questions about the organisation [...]. We also include knowledge-based questions that give insights into how people perceive themselves" [P3].*

However, participants also raised concerns about the limitations and potential biases of the surveys. One expert cautioned against overreliance on survey data, pointing out how many uncorrelated factors can distort responses: *"A hundred different factors could affect how I fill out a survey. So, while surveys can be helpful, I'm not their biggest fan. They do provide useful information, though "[P4].* This statement was further elaborated on with: *"[...] you'd need to conduct a survey beforehand [...]. But then you risk introducing bias. By presenting a scenario, you might be shaping the mental models you're trying to assess." [P4].*

The potential for desirability bias and limited behavioural insight was also raised. Participants noted that while surveys may be effective in assessing knowledge and, to some extent, attitudes, they fall short when it comes to behaviour: *"I think a survey can be used for knowledge; however, it has its own limitations. For attitude, I think a survey can be effective, especially if you want to conduct large-scale interviews. But for behaviour, there will be bias from people who say they act a certain way but actually don't, or they provide the answer they think they should give, while in reality, they act differently" [P1].* Another expert went further, critiquing the assumption that people are always consciously aware of the reasoning behind their own actions: *"Surveys are good for getting a general view, but they assume people fully understand why they do certain things. Many of these mental processes are unconscious, so it's tricky to get accurate results from a survey alone." [P7].*

To address these limitations, multiple participants advise combining surveys with additional validation techniques. Several participants mentioned direct observation as a necessary supporting method to gain a more reliable picture of users' MM, along with (follow-up) interviews: *"Since it's qualitative research, people might only choose from the available answers, which can limit the results. It's important to validate the survey findings through interviews." [P5].*

Interviews and Think-Aloud Approaches (6/5): Interviews, as an indirect elicitation method, were generally viewed as more suitable for uncovering the implicit or unconscious aspects of MMs. Several experts emphasised that interviews allow for nuanced follow-ups and deeper probing, something that is missing in surveys: *"I find that if you ask people through a survey, you rarely get deep insights. Some parts of mental models are things you're conscious of, but other parts are not. So it's hard to identify why people behave in certain ways. If you really want to understand someone's behaviour, you need to interview them, and have a skilled interviewer who can ask follow-up questions to truly understand" [P7].*

Others highlighted the need for psychological expertise, particularly when the goal is to capture underlying beliefs and behavioural drivers: *"To properly assess someone's mental model, you'd need a psychologist, likely a clinical psychologist, to speak to the individual. People often think they understand their mental model, but it's not always what they believe it to be." [P6].*

In line with the elicitation model proposed by Grenier and Dudzinska-P (2015), participants discussed the use of think-aloud interviews and graphical tasks as a means of encouraging the participants to express their internal reasoning in real time. This combination of narrative and visual elicitation was viewed as highly effective: *"We interviewed many people with open questions to let them describe their own mental model. In addition to that, we also asked them to make drawings during the interview and then explain their drawings out loud while they were drawing, in order to really grasp what their mental models were" [P1].*

These insights directly inform Activity *1.6.4 Conduct Interviews* and Activity *1.6.5 Conduct Pictorial tasks* in the proposed framework. Together, they reflect a layered approach to indirect and direct MM elicitation.

6.4.1 Additional Methods (20/8)

Beyond surveys and interviews, the participants highlighted several additional techniques to assess the KAB values that shape the MMs of employees. These approaches focus on leveraging objective data sources, contextual cues, and comparative analysis to obtain more profound and more reliable insights.

Although only discussed by one participant, the framework by Kruger, Drevin, et al. (2006) came forth with the mention of *System-Generated Data (1/1)* for behavioural assessment. This participant emphasised the potential by referring to the extent of behavioural data already embedded in enterprise systems: *"Microsoft Graph contains vast amounts of data regarding people's behaviour within systems." [P6].* These data can be analysed to assess patterns related to compliance, access control, or risky behaviour, providing a continuous, non-intrusive stream of indicators related to awareness.

An additional method that was discussed more frequently was *direct observation (4/4)*. Although closely related to system monitoring, this research considers direct observation, informal environmental audits, or controlled behaviour testing. Several experts noted that in-person observations and simulated attacks can reveal behavioural aspects that surveys or interviews often miss: *"For measuring knowledge and behaviour, I don't think it's sufficient. You need to observe people as well, either in the office or through system interactions" [P6]* and *"Or you could use phishing tests—they're the most well-known example of behavioural measurement." [P8]*

Another strategy suggested only once was to get insight by asking individuals not about themselves but about the behaviour of their peers. This *indirect Peer Review (1/1)* approach can mitigate self-enhancement bias, as individuals do not have to overestimate their own competence: *"To get honest answers about mental models, it's better to ask people about the behaviour of others in their office, not their own. There is a bias where people always think they're better at something than they are." [P3].*

Finally, one participant also introduced a comparative assessment method, in which individual MMs are evaluated against a reference model developed by subject matter experts. This comparative *Expert Comparison (1/1)* enables measurement of conceptual gaps: *"We created a mental model of how that system should be represented according to a few experts. Based on that, we explored the individual mental models and determined what was incorrect, what was missing, what was understood, and what the gaps were." [P1]*. Such gap analysis provides a structured way to assess not just the presence of knowledge, but its completeness and alignment with expert expectations. It also helps inform targeted training interventions to close specific cognitive or conceptual deficiencies.

These additional approaches are reflected in the framework as Activity *1.6.1 Direct Observe Employees* and *1.6.2 Retrieve Computer Logged Data*. Together, they broaden the scope of MM measurement beyond subjective self-reporting, integrating behavioural and comparative data sources for a more holistic assessment.

Group-Perspective Preference (7/6): In practical organisational settings, it is often neither feasible nor necessary to treat each employee's MM as wholly unique. Instead, a more scalable approach is to group individuals based on similarities in their MMs, enabling more efficient intervention strategies. This concept of segmentation was widely supported by the interview participants, both from a pragmatic and security-focused perspective.

One expert emphasised the trade-off between theoretical precision and organisational feasibility: *"It's difficult to develop an individualised programme based on each person's mental model. So, theoretically, it's important, but practically, it's hard to implement. You might gather individuals' mental models, but ultimately group them into something more general for the whole department, and that doesn't always align precisely with each person's mental model. It's tough to achieve full accuracy in individual interventions" [P5].*

From a security management standpoint, grouping MMs was seen as a logical step toward targeted control mechanisms. Participants noted that once these groups are formed, organisations can develop tailored training, policy communication, and awareness programmes suited to the behavioural and conceptual profiles of each group: *From a security standpoint, yes, it makes sense to do that (grouping of Mental Models). This way, you can create specific controls tailored to certain groups of people" [P6].*

Another participant elaborated on how this could be operationalised: *"[...] you could segment them. That's probably the most practical approach. Once you have segmentation, you could build a segmentation model. Based on that segmentation model, you could classify employees [...]" [P8].*

The literature supports this practice. The grouping of MMs can be based on the cooccurrence of themes in participant responses, which provides a metric for the similarity of MMs between individuals (Bessette et al., 2017). Once themes are identified, the frequency or proportion of participants associated with each concept can be calculated to build MMs at the group level (J. K. Doyle et al., 2008). These quantified MMs can then be used to inform targeted action at the group level rather than trying individual tailoring for every employee.

This grouping process is formalised in the framework through Activities *1.3 Define MM Categories* and *1.8 Interpret Results Using Mapping of MMs*. The former activity establishes a set of predefined MM categories, while the latter maps the collected MM data onto these categories to determine the classification of an individual or group.

6.5 Illustrative Example

Based on the results of the NLR and the expert interviews, a preliminary framework was constructed to assess MMs in cybersecurity contexts. To demonstrate its conceptual utility, this section provides an illustrative example focused on Shadow IT-related MMs. Although this example does not constitute a formal empirical case study, it is grounded in insights drawn from both the NLR and the SSIs. The aim is to clarify the reasoning processes of the framework and to showcase how theoretical constructs and empirical evidence combine in practice.

1.1 Eliciting Input from Stakeholders: In alignment with Stanton et al. (2005), Parsons, McCormac, et al. (2014), and Kruger and Kearney (2006), the process begins with interviewing key stakeholders and senior management. These interviews seek to identify organisational priorities, perceived risks, and high-interest security domains. The expert interviews reinforced the value of this step, as several experts stressed that management participation is essential to effectively align awareness strategies with organisational culture and priorities.

2.1 Declare Focus Area(s): Through the interviews discussed, senior management will express their wishes and declare their focus area. In this illustrative example, senior management has identified organisational information security as the principal domain of interest due to concerns about the increase in internal threats and compliance gaps.

2.2 Specify Key Topic(s): Once the broader area is defined, interviews with stakeholders help refine the key topic(s). Examples of such key topics can be found within the HAIS model by Parsons, McCormac, et al. (2014). However, for this illustrative example, Shadow IT was selected as the key topic within organisational information security.

1.2 Explore Key Topic(s): With the key topic established, experts engage in preliminary exploratory activities to deepen their contextual understanding. This includes conducting pilot interviews with organisational actors such as IT staff, security managers, and other employees. These interactions serve to both inform the experts' domain understanding and to map how employees conceptualise the topic.

1.3 Define Mental Model Categories: Based on insights from Jansen et al. (2023), two categories of MMs are defined: *Risk-Averse* and *Risk-Taking*. Each category consists of distinct MMs that reflect how users internally rationalise Shadow IT, see Table 2.2. Within each group, the difference between MMs is based on several factors, focused on one or more KAB dimensions or prevalent challenges. These categories and underlying MMs were based on explicit expressions towards the perception of the focus area during interviews. With a general understanding of Shadow IT, experts will define the different MMs based on Jansen's work.

1.4 Determine Measurable Metrics: The next step involves identifying relevant measurable metrics, drawing on the literature and interview-derived themes (Bitton et al., 2018; Jansen et al., 2023). A concise example of determining metrics based on defined factors is provided in Table 6.5. This figure shows a list of specific criteria for measuring knowledge and attitudes related to mobile applications. Examples of general Shadow IT factors to consider, taken from the results of Jansen et al. (2023), are shown in Table 6.8.

Table 6.8: General factors related to Shadow IT MMs, taken from Jansen et al. (2023)

General MM factors
Attitude with regard to the handling of Shadow IT
Aversion to unfavourable implications
Aware of the concept of Shadow IT
Basic understanding of cybersecurity concepts
Determination to achieve desired results
Deviation from established rules or guidelines
Explicit prioritisation for cost savings
Expertise within the field of technology
False or misguided idea of being protected
General awareness
Overall understanding of technology
Perceived impact and consequences of risks
Reliance on established knowledge and wisdom
Risk behaviour
Self-sufficiency

In addition, expert interviews played an essential role in indicating the influence of prevalent challenges and employee-related influence factors, as well as in constructing and prioritising specific metrics accordingly. These factors can serve as potential proxy variables, helping with the premature assessment of employee MMs. For example, several participants pointed to employees' self-efficacy and threat appraisal, as well as pressure and stress, as prevalent challenges, indicating their importance and role in helping with selecting and adapting appropriate metrics. Table 6.9 illustrates these factors and challenges of influence, such as personality traits, work location, or organisational culture, that the interviewees cited as relevant to the formation of MM. It is essential to address that these factors are not static and can change based on role, experience, or exposure to SETA programmes.

1.5 Map Metrics to MM elements: The defined metrics are then systematically assigned to the individual MM elements. This stage is shown in Table 6.10, which details how *Consequence-Avoidance Orientation* can be assessed using KAB metrics related to policy awareness, risk aversion, tool adherence, and incident avoidance. The value of this mapping lies in making internal cognitive constructs externally observable through measurable and transparent metrics. This step operationalises the interview feedback, particularly the emphasis on multimethod elicitation. Experts suggested that combining self-reported data, such as KAB questionnaires, with behavioural indicators, such as system logs, offers a more holistic and credible view of an individual's MM.

Table 6.9: Assessment Methods for Workplace Factors

Factor	Assessment Method
Demographics	Observation, Questionnaire
Education	Questionnaire, Interview
Experience	Questionnaire, Interview, System Logs
Past security training	Interview, System Logs
Personality Factors	Questions
Employee Motivation	Questions
Organisational Culture	Observation, Questionnaire
Predominant Work Location	Observation of logs, Questionnaire
Stress & Pressure	Calendar, Deadline Logs, Questionnaire
Threat Appraisal & Self-Efficacy	Exercises, Questionnaire

Table 6.10: Example of Metrics for MM assessment

Mental Model		General Metrics	Concrete Metrics	Method
Consequence-Avoidance Orientation	Knowledge Metrics	Familiarity with Organisational IT Policies	Ability to name examples of banned or discouraged applications	Questionnaire, Interview
			Understanding internal reporting or approval processes for new tools	Questionnaire, Interview
		Awareness of Shadow IT Risks	Awareness of financial or reputational damage caused by Shadow IT incidents	Questionnaire, Interview
			Ability to identify common security risks of using unauthorised applications	Questionnaire, Interview
	Attitudinal Metrics	Trust in IT Governance	Belief that internal IT systems are secure and sufficient	Questionnaire, Interview
			Express more comfort with officially endorsed tools	Questionnaire, Interview
			Willingness to consult IT before adopting new tech	Questionnaire, Interview
		IT Policy and Rules Adherence	Adherence driven by fear of consequences	Questionnaire, Interview
			View rules as necessary to prevent mistakes or failures	Questionnaire, Interview
			Belief that non-compliance leads to serious consequences	Questionnaire, Interview
		Risk Aversion	Avoidance of convenient but unauthorised apps due to fear	Scenario-based Questionnaire, Interview
			Expressed concern over the unknown consequences of Shadow IT	Questionnaire, Interview, Helpdesk tickets
			Perceived consequences of using unauthorised tools	Questionnaire, Interview
			Anxiety Levels about Cybersecurity	Questionnaire, HR, Helpdesk data, Monitoring
	Behaviour Metrics	Decision-making in simulation	Choosing secure and trusted options when faced with Shadow IT simulation	Simulation/test, Picotiral task
		Shadow IT usage frequency	The proportion between official and unofficial applications accessed or installed	System data, Monitoring
		IT Policy and Rules Adherence	Completion rates of IT security training programmes	System data
			Deletion of tools flagged as risky	System data, Monitoring
		Tool Adoption Behaviour	Regular checks for whether tools are approved before use	Helpdesk logs, Monitoring
			Seeking formal approval before installing or using new software	Helpdesk logs
		Incident Avoidance	Frequency of IT/security consults	Helpdesk logs, Monitoring
			Frequency of reporting security concerns about suspicious applications to IT	Helpdesk logs
		Information Security Practices	Avoid storing sensitive data in unapproved cloud services	System data, Monitoring

1.6 Gather Mental Model Metrics: With a robust set of metrics defined, the next phase involves gathering empirical data through various elicitation methods. As outlined in Figure 6.1, five primary data collection methods are proposed. These approaches are both theoretically grounded and empirically validated, aligning with interview findings that stressed the need for multimodal evidence to reliably assess MMs. In Table 6.10, the final column exemplifies which elicitation methods correspond to particular metrics, specifically for the *Consequence-Avoidance Orientation MM*.

1.6.1 Direct Observe Employees and *1.6.2 Retrieve Computer Logged Data*: Although self-reports and interviews are effective in determining knowledge and attitudes, they can be less reliable in measuring actual behaviour due to recall limitations or social desirability bias. Kruger, Drevin, et al. (2006) advise for behavioural metrics captured via system logs, such as software usage frequency, or violation of access policies. In addition to the literature, the interviewees also stressed this limitation, highlighting that behaviour is often best captured through unobtrusive monitoring or real-time observation. Several experts interviewed in this study supported this by recommending office walk-arounds and staged social engineering tests to surface real-world behaviours. The three categories of monitoring methods, with examples of actionable metrics, are summarised in Table 6.11

Table 6.11: Security-Related Monitoring Methods

Monitoring Methods	Monitored Activities/Metrics
Computer/System Logs	Illegal web page visits Installation of anti-virus software Switching off PCs after hours Downloading large files from the Internet
Testing	Password security checks Detection of suspicious email messages Phishing awareness testing
Direct Observation	Close laptops when away Notes with passwords Peer-interviews

1.6.3 Conduct KAB Questionnaire: Once measurable dimensions and related MMs are defined, experts develop KAB questionnaires to assess employee reasoning and perceptions. A source of such questions and statements for Shadow IT is found in the study by Banken (2025), offering four items per MM, covering both *knowledge* and *attitude*, see Appendix A.4. Furthermore, a simplified example of general KAB statements adapted from Kruger, Drevin, et al. (2006) is provided in Table 6.12. These statements, tailored to the Shadow IT context, help assess the metrics.

Table 6.12: Examples of General KAB Statements, focused on Shadow IT

Category	Statement
Knowledge	I am allowed to download any programme onto my work computer if it helps me do my job
Attitude	It can be risky to download programmes on my work computer
Behaviour	I download any programme onto my work computer that will help me get the job done

1.6.4 Conduct Interviews: Although structured questionnaires provide scalable insights, interviews offer depth and context. Jansen et al. (2023) emphasises that individuals can express different MMs in different contexts, making interviews a valuable method to explore the underlying mental representations and inconsistencies in reasoning. As supported by the interviewees in this study, interviews are adaptive as they allow follow-up questions, enabling the expert to clarify ambiguous responses.

A recommended interview structure includes a mix of open and guided questions, aligned with the KAB dimensions. Table 6.13 outlines example questions tailored to Shadow IT, derived from the pre-identified metrics. These questions help to surface the rationale behind the observed behaviours, thereby reinforcing the empirical foundation of the MM analysis.

Table 6.13: Interview Example Questions on Shadow IT Awareness and Usage

Question	Category
What is Shadow IT for you?	Knowledge
Have you ever used any shadow IT instances?	Behaviour
Why have you used these shadow IT instances?	Behaviour / Attitude
What do you think are the risks and implications of shadow IT?	Knowledge
What is the process of using non-standard solutions?	Knowledge
Are you aware of the [organization policy name]?	Knowledge / Attitude
Have you discussed the use of technology among co-workers?	Behaviour / Attitude
Do you feel well-informed about the use of technology?	Knowledge / Attitude

1.6.5 Conduct Pictorial Tasks: Pictorial tasks offer an alternative means of cognitive elicitation to complement verbal and observational methods. As utilised by Jansen et al. (2023), and recommended by E. E. Doyle et al. (2022) and Grenier and Dudzinska-P (2015), these exercises ask participants to draw representations of their understanding of IT processes or security decision-making.

Interview feedback underscored the utility of such non-verbal tools in capturing intuitive, pre-reflective elements of an individual's MM. These outputs are analysed by comparing them with a validated expert baseline, allowing qualitative scoring of complexity, accuracy, and alignment with organisational expectations.

1.7 Analyse Gathered MM Metrics: Following the multimethod data collection, the expert team undertakes an analysis of the gathered metrics. For KAB questionnaire data, scoring mechanisms such as mean averages or weighted scorecards are used to generate a three-dimensional KAB profile (Androulidakis and Kandus, 2011b; Kruger, Drevin, et al., 2006). Visual data from pictorial tasks is evaluated qualitatively by comparing employee outputs to expert-generated benchmarks. Discrepancies are analysed, allowing nuanced judgements of MM alignment or deviation. Importantly, observed or system-logged behaviour is cross-checked against self-reported values to detect consistency gaps. If, for example, an employee reports cautious behaviour but the log data shows frequent installations of unauthorised applications, this discrepancy may signal a misguided sense of *Threat Appraisal & Self Efficacy*.

1.8 Interpret Results Using Mapping of MMs: The quantification of KAB dimensions provides a structured overview of an employee's security mindset. However, the interpretation phase goes beyond the numerical scoring. Experts must holistically assess how the combination of responses, behavioural indicators, and contextual variables align with predefined MM categories. This mapping process involves cross-referencing elicited data with the distinctive characteristics of each MM, as outlined by Jansen et al. (2023). For example, high levels of risk awareness coupled with procedural compliance may correspond with *Consequence-Avoidance*

Orientation MM. In contrast, indicators of adoption of information tools, despite risk awareness, may align with the *Performance-Driven Rule Bending* MM.

During expert interviews, multiple participants stressed that interpretative judgment is essential in this step. They argued that rigid metrics alone cannot fully capture the nuance of employee reasoning or intention. Instead, they recommended incorporating system logs, verbal justifications from interviews, and observed behaviour into a unified cognitive profile. This also helps to surface ambiguous and hybrid profiles, where an employee exhibits elements of multiple MMs. For example, an individual might express strong procedural awareness, yet bypass official IT channels for reasons of efficiency. The goal is to assign each employee to one or more MM categories based on pattern recognition, not just isolated data points.

1.9 Assessing Final MMs: After combining the insights from all elicitation activities, experts make a final determination of an employee's MM classification. As stated above, this is not a binary outcome, but rather a nuanced assignment across one or more MMs, acknowledging that MMs are dynamic and context-sensitive rather than fixed or mutually exclusive.

Ultimately, the outcome of this process is a well-documented MM profile for each employee, grounded in both subjective and objective metrics. These profiles can inform: *Targeted security training, Policy communication strategies* and *Risk monitoring approaches*.

6.6 Summary

This chapter proposed a structured framework for eliciting and assessing employee MMs related to ISA. The foundation of this framework lies in identifying appropriate elicitation techniques and aligning them with key focus areas within organisational security contexts.

The literature review distinguished between two broad categories of MM elicitation methods, direct and indirect. Direct techniques, such as drawing tasks, card sorting, and concept mapping, offer visual or interactive representations of MMs. Indirect approaches, which were more commonly cited and preferred by interview participants, include surveys, interviews, and the use of system-generated behaviour data. Although direct techniques were less frequently mentioned in interviews, they remain valuable for specific use cases and deeper cognitive insight.

After determining suitable elicitation techniques, the framework focuses on identifying relevant security topics through interviews with senior management. These topics help define the scope and direction of the MM assessment. Within these domains, experts develop categories of MMs, informed by preliminary interviews and subject-specific concerns. The dimensions of ISA, Knowledge, Attitude, and Behaviour, serve as the primary base for measurement, consistent with established ISA taxonomies.

Quantitative data on KAB dimensions are gathered primarily through structured questionnaires and interview prompts developed by experts. These are complemented by observational methods such as system log analysis, social engineering simulations, and direct behavioural observations. The inclusion of these diverse data sources ensures that both subjective and objective perspectives are captured. Pictorial tasks, though less common, offer additional nuance in assessing conceptual understanding in targeted areas.

The integration of data from these multiple sources allows experts to construct a holistic ISA profile. For this research, all measured criteria are weighted equally. Although assigning differential weights might enhance precision and accuracy, this is beyond the scope of this research. The selected scoring approach, whether through scorecards or mean-based methods, produces interpretable outputs for each of the KAB dimensions, enabling the inference of an individual's MM profile.

Finally, the KAB scores are assigned to representative MMs. This mapping process considers contextual influence factors such as employee demographics, work context, and known organisational challenges. These factors represent valuable proxies for early-stage assessment and hypothesis generation. Their actual predictive impact remains a subject for future research.

In general, this chapter represents a replicable, flexible, and empirically grounded method for assessing MMs related to ISA. The proposed framework offers a foundational step towards more targeted, adaptive, and evidence-based cybersecurity awareness strategies.

7 Results

This chapter presents the findings of the study based on the SRQs introduced in Chapter 3. The results are derived from both the narrative literature review and expert interviews. Each subsection in the following addresses one of the SRQs and summarises the empirical and conceptual outcomes that inform the proposed assessment framework.

7.1 Findings

We present the key findings of this study by addressing each of the SRQs introduced in Chapter 3. These findings form the foundation upon which the developed framework for assessing employee mindsets toward Shadow IT was constructed.

SRQ1: *What are the predominant challenges related to organisational information security awareness?*

The literature review and interviews revealed several recurring challenges that affect ISA in organisations. A key issue is the limited effectiveness of current SETA programmes, which are often perceived as generic, unengaging, and poorly aligned with the actual decision-making processes of employees. Employees tend to lack the motivation to participate in these programmes and frequently do not see the relevance of the material to their day-to-day work. Furthermore, the absence of clear methodologies to tailor training to employee needs hinders long-term behavioural change. Factors such as unclear communication of security policies, low perceived personal responsibility, and insufficient reinforcement mechanisms were also identified as contributors to weak ISA. These findings highlight the need for more personalised and context-sensitive approaches to improving ISA.

SRQ2: *What employee-related aspects need to be considered in the assessments of Mental Models?*

Employee-related factors that influence the formation and structure of MMs include demographics, such as gender and age, personality traits, such as conscientiousness and extraversion, and work location, decentralised or not. For example, demographic findings found that experience plays an important role. This supports the distinction between expert and "folk" users, where the latter group constructs MMs based on assumptions, intuition, or incomplete information due to a lack of formal training and knowledge. Such users often operate under inaccurate beliefs, such as overconfidence in personal judgment or reliance on informal rules, which can lead to risky behaviour. Importantly, we found that even inaccurate MMs are valuable and can still guide positive behaviour. Thus, assessing these internal representations can offer predictive insights into employee actions. This is especially true when evaluating compliance with security policies or the use of unauthorised tools such as Shadow IT.

SRQ3: *How can Mental Models of cybersecurity be elicited and classified in the context of organisations?*

The study found that, while MMs are inherently abstract and subjective, they can be approximated and assessed through structured instruments. Semi-structured interviews, guided questionnaires, and thematic coding of employee responses emerged as practical tools to identify MM patterns. A key methodological insight was that assessing not only what employees know, but also how they reason about risks and behave in risk-filled information security areas, provides a richer understanding of their general awareness. Based on the literature, we developed a measurable hierarchy of security awareness indicators. These indicators are mapped across the knowledge, attitude, and behaviour dimensions (KAB). These dimensions serve as the backbone of the proposed assessment framework.

Existing security behaviour taxonomies, such as those proposed by Kruger, Drevin, et al. (2006) and Bitton et al. (2018), provide a helpful starting point but require extension to capture the nuances of MMs. The key insight was that behavioural indicators alone do not fully account for the reasoning process underlying security-related decisions. Integrating MM aspects adds a cognitive dimension to the behavioural taxonomies, allowing for more precise categorisation of users. Furthermore, we found that MM types, such as risk-averse versus risk-taking, can be layered on traditional behavioural metrics to improve precision and intervention design. This enhanced taxonomy forms a critical component of the developed framework, enabling organisations to align SETA strategies better with both observed behaviour and the underlying MMs.

7.2 Interpretation of Results

The findings above reveal several important insights into the cognitive and organisational dynamics that influence employee security behaviour. At the core of these results is the recognition that MMs offer a powerful lens through which to interpret the persistent challenges organisations face in strengthening ISA. Rather than focusing solely on behavioural outcomes or knowledge levels, this study shifts attention to how employees reason about cybersecurity, so what risks they perceive, how they interpret guidance, and why they make particular decisions in their daily workflows.

The results suggest that weak ISA is not merely a function of low knowledge but often stems from cognitive mismatches between employees' internal models and the formal expectations of the organisation. Security practices are frequently filtered through assumptions about perceived personal responsibility, risk likelihood, or efficiency trade-offs. These are examples of shortcuts that may conflict with policy, but appear rational from the employee's perspective. This highlights a critical limitation of traditional SETA programmes, as even well-designed training materials can fail if they do not align with how individuals internally frame and prioritise information.

Notably, the study confirms that these MMs are shaped by individual and contextual factors such as role, experience, and incident exposure. This variation explains why employees respond differently to the same security protocols or programmes. A technically skilled employee may ignore security controls due to overconfidence, while a non-technical colleague may unknowingly bypass procedures due to a simplified or incomplete understanding of risk. These distinctions are not easily captured by standard behavioural metrics, but become visible through the proposed MM framework.

Furthermore, the research illustrates that it is feasible to identify and classify these internal models using structured elicitation methods, primarily when supported by the KAB framework. Although MMs are subjective and often implicit, their influence is observable and can be systematically interpreted through a combination of methods such as interviews, surveys, and behavioural data. This creates new opportunities for organisations to proactively map cognitive risk profiles and tailor awareness interventions accordingly.

Taken together, the results support the broader argument that eliciting MMs and improving ISA requires a more nuanced, human-centric approach, one that recognises employees not just as compliance targets, but as reasoning agents. By focusing on how users make sense of security, rather than simply what they do or know, organisations can begin to design SETA programmes that truly engage with the cognitive drivers of behaviour. This cognitive focus does not replace technical or procedural safeguards, but rather complements them by addressing the psychological layer of IS defence, arguably the most variable and underdeveloped aspect of organisational security.

Although these findings offer a foundation for more targeted ISA interventions, they also raise important questions about the practical application and generalisability of the proposed approach. The following chapter discusses these contributions in more detail, reflects on the study's limitations, and outlines implications for future research and practice.

8 Discussion

This chapter positions the key findings of the research within the broader academic and practical landscape of cybersecurity awareness. The chapter also explores the practical implications of the proposed framework, and examines its limitations, in addition to ethical remarks.

8.1 Contributions to the Field

This research contributes to both the scientific understanding of human-centred cybersecurity and its practical application in organisational contexts. In particular, it addresses the underexplored intersection of MMs, security behaviour, and Shadow IT by offering a structured framework for assessing employee perceptions and behaviours.

Scientific Contributions: From a scientific perspective, this thesis extends existing research on MMs by examining how internal cognitive representations shape employee reasoning and behaviour in cybersecurity contexts, specifically when looking at Shadow IT. Although earlier studies have highlighted the role of MMs in explaining user behaviour, this study advances the field by proposing a structured, analytical framework to assess them in practice.

The framework integrates subjective KAB data with behavioural proxies, such as system-generated data, providing a multidimensional and methodologically grounded approach to the assessment of MM. This interpretive approach avoids oversimplification, offering a realistic lens through which the often invisible reasoning behind security-related actions can be explored.

In addition, the research refines existing behavioural taxonomies by embedding cognitive and motivational dimensions. This enables more holistic categorisation of user behaviour, as it captures not only what users do, but why they do it. As a result, the framework contributes to the field of human-centric risk modelling by supporting a deeper understanding of how MMs evolve and how awareness initiatives may shape them over time.

Practical Contributions: Practically, this research offers a novel and applicable framework to help organisations tailor SETA programmes more effectively. By enabling the identification and categorisation of MMs, the framework allows organisations to align awareness strategies with the cognitive needs and risk profiles of different employee groups. This approach improves the relevance and engagement of ISA initiatives, allowing resources to be allocated more efficiently towards areas of highest risk. It also provides organisations with a mechanism to interpret the reasoning behind policy non-compliance, helping to surface cognitive barriers such as misconceptions, risk misperception, or lack of trust in official systems.

Furthermore, the framework supports the ongoing evaluation of ISA efforts. By tracking changes in employee understanding and attitudes, organisations can better assess the impact of awareness programmes and refine them based on cognitive and behavioural outcomes.

Ultimately, this research contributes to a more adaptive and psychologically grounded approach to cybersecurity training and culture building, helping organisations beyond surface-level compliance and toward genuine behavioural change.

8.2 Limitations

Although the study offers valuable contributions, it is not without limitations. First, the empirical data were drawn from a limited sample of interview participants. Although efforts were made to ensure diversity across roles and sectors, the sample may not capture the full range of employee mindsets or organisational cultures, particularly in non-Western contexts or sectors with distinct regulatory environments.

There is also the limitation of not saturating the interviews. Although the intended sample size was twelve participants, only eight interviews were ultimately conducted due to scheduling constraints. Although this number is within the acceptable range for qualitative research and core themes emerged consistently, a smaller sample increases the risk that some perspectives, particularly from the less-represented subgroup, may have been underexplored. This limitation may affect the transferability of the findings to other organisational or disciplinary contexts.

Secondly, the inherently abstract nature of MMs introduces challenges in validation and elicitation. Although the framework provides an initial structure for quantification, further empirical testing is required to ensure reliability and scalability.

Third, the study's scope was confined to Shadow IT, a critical but specific domain within broader cybersecurity concerns. The findings may not fully generalise to other contexts such as phishing, password hygiene, or social engineering.

Lastly, while the literature review and interviews were comprehensive, the integration of findings may reflect certain interpretive biases. Translating qualitative insights into a structured framework involved subjective judgment, which may affect applicability.

8.3 Ethical Discussion

As this study proposes a framework for assessing employee MM, using both self-reported data and system-generated metrics, it is important to critically reflect on the ethical implications of collecting and using such data in organisational settings. Although such assessments may offer substantial value to improve cybersecurity awareness and mitigate internal threats, they also raise significant concerns about employee privacy, informed consent, data governance, and potential misuse of surveillance-like practices.

A core ethical issue is related to the risk of employee surveillance. Frameworks that monitor behaviour through system logs, or application usage, may blur the boundary between security oversight and intrusive surveillance (Ball, 2010). Although the intention may be to mitigate risk, employees may perceive such data collection as violating their personal space, especially if they are unaware of what is being monitored. Excessive or unclear monitoring practices have resulted in decreased trust, reduced morale, and resistance to compliance initiatives.

Ethical data collection must be grounded in transparency and informed consent. Employees should be clearly informed about the nature, scope, purpose, and retention of the collected data. This includes both subjective input, such as survey responses, and objective digital observations, such as software usage and access logs. Following ethical frameworks such as the General Data Protection Regulation (GDPR), organisations must ensure that data subjects, here employees, are fully aware of what data is being processed, for what purpose, and how it will be protected (Voigt and Von dem Bussche, 2017). Simply including consent forms may not be sufficient if consent is not meaningfully voluntary, particularly in hierarchical employment contexts where refusal may be implicitly discouraged.

There is an inherent power asymmetry in employer-employee relationships. Although organisations may seek to use MM assessments to tailor awareness programmes, such frameworks also create opportunities for profiling, categorisation, and even disciplinary action based on perceived profiles. This can lead to ethical misuse if insights into employees' cognitive models or behavioural patterns are employed for performance evaluation or disciplinary decision-making, rather than constructive awareness-building. It is therefore ethically imperative that the use of such frameworks be clearly limited to educational and developmental purposes. Clear boundaries must be established to prevent the misuse of psychological profiling or behaviour prediction in ways that would undermine the dignity or autonomy of the employee.

Another ethical principle is data minimisation, collecting only what is necessary and proportionate (Antignac et al., 2017). This is particularly important when assessing MMs, which are inherently subjective, fluid, and context-dependent. Overreliance on system-generated proxy metrics for behaviour may risk overgeneralisation or misinterpretation. Therefore, any MM assessment framework must ensure that the data collected is both relevant and proportionate, and that interpretations are treated probabilistically rather than deterministically.

Lastly, it is worth noting that ethical data collection itself might be a driver of improved ISA. Transparent and participatory assessment practices can improve trust and engagement. Involving employees in the design and implementation of MM assessments and communicating how data insights will be used to benefit their security posture can transform ethical risks into opportunities for collaborative awareness building.

In conclusion, while this thesis contributes to a novel method for assessing the human dimension of cybersecurity through MMs, it must be understood that ethical considerations are not peripheral but central to the responsible deployment of such frameworks. Balancing the need for organisational security with the rights and dignity of employees is not only a legal requirement, but a prerequisite for a sustainable and trustworthy cybersecurity culture.

9 Conclusion

This research was set out to answer the main research question:

MRQ: *"How can Employee Mental Models toward Shadow IT Usage be Identified and Assessed in an Organisational Context?"*

This study has proposed a framework to assess employee MMs about Shadow IT by translating the abstract MMs into structured, measurable proxy metrics on ISA. These metrics are grounded in the Knowledge, Attitude, and Behaviour (KAB) dimensions, obtained through direct, indirect, and system-generated data elicitation. This forms the basis for a robust hybrid approach that captures what employees do and how and why they make certain security-related decisions.

Through literature and expert interviews, this research identified that MMs, internal individual representations of reality, shape how employees perceive cybersecurity risks and justify their behaviour. Shadow IT usage often results from reasoning that seems rational to the employee, but conflicts with formal IT policies. These internal justifications are not easily visible, yet they are critical in security-related decision-making.

This thesis developed a novel framework that connects qualitative information with measurable indicators to make such MMs assessable. The framework addresses the importance of determining the focus area and the key topic selection that results from it. For metrics, the frameworks include those that focus on KAB. By combining elicitation methods such as interviews, questionnaires, system-generated data tracking, and general monitoring, we can obtain these metrics and evaluate and map them to specific MMs.

This approach also allows organisations to identify patterns between employee groups, such as whether certain MMs are associated with specific roles, experience levels, or types of risky behaviour. It also helps to expose discrepancies between self-perception and actual behaviour, a critical insight for designing more effective SETA programmes.

Ultimately, this research concludes that prevalent employee MMs, in this research focused on Shadow IT, can be systematically assessed using a metrics-based framework that combines cognitive, behavioural, and system-level indicators. The operationalisation of MMs into quantifiable KAB dimensions offers a practical and scalable way to enhance understanding of human security behaviour. This assessment enables organisations to better personalise cybersecurity awareness efforts, improve resource allocations, and proactively mitigate internal security risks from misunderstood or misapplied reasoning.

This created framework is not merely a descriptive insight into employee behaviour, but an action model for risk detection and intervention. This shifts cybersecurity management from reactive to proactive, and from technical to more human-centric.

9.1 Future Research and Practice

Although this study lays a solid foundation, several recommendations for further exploration and development can be proposed to improve practical relevance and theoretical robustness.

- Empirical Validation and Integration with SETA programmes: Practical implementations of the framework in organisational environments could provide feedback on its usability, integration challenges, and the effectiveness of SETA programmes informed by MMs.
- Cross-Contextual Studies: Research comparing different organisational sizes, sectors, or cultural environments could examine the generalisability of MM typologies and metric structures.
- Broader Behaviour Scope: This study focused on Shadow IT, meaning that future studies could apply the framework to other risk behaviours such as phishing susceptibility, password management, or data handling practices.
- Determine Specific Factor and Metric Influence: As discussed, it was considered outside the scope of this study to look at the specific weights that all metrics have on the final determination of MMs. Furthermore, while we made assumptions about the influence of employee-related factors, such as age, personality, and experience, on MMs, we did not validate this. Future research could examine whether these factors can indeed be used as proxy variables and the specifics of how they influence MMs.

Bibliography

Abawajy, Jemal (2014). "User preference of cyber security awareness delivery methods". In: *Behaviour & information technology* 33.3, pp. 237–248.

Adams, William C (2015). "Conducting semi-structured interviews". In: *Handbook of practical program evaluation*, pp. 492–505.

Adeoye-Olatunde, Omolola A and Nicole L Olenik (2021). "Research and scholarly methods: Semi-structured interviews". In: *Journal of the american college of clinical pharmacy* 4.10, pp. 1358–1367.

Ahmad, SHOEB (2014). "Technology in organizations". In: *International Journal of Research in Business Management* 2.7, pp. 73–80.

Albladi, Samar Muslah and George RS Weir (2017). "Personality traits and cyber-attack victimisation: Multiple mediation analysis". In: *2017 Internet of Things Business Models, Users, and Networks*. IEEE, pp. 1–6.

Aldawood, Hussain and Geoffrey Skinner (2019). "Reviewing cyber security social engineering training and awareness programs—Pitfalls and ongoing issues". In: *Future internet* 11.3, p. 73.

Alladi, Tejasvi, Vinay Chamola, and Sherali Zeadally (2020). "Industrial control systems: Cyberattack trends and countermeasures". In: *Computer Communications* 155, pp. 1–8.

Alotaibi, Mutlaq, Steven Furnell, and Nathan Clarke (2016). "Information security policies: A review of challenges and influencing factors". In: *2016 11th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, pp. 352–358.

Alseadoon, Ibrahim, MFI Othman, and Taizan Chan (2015). "What is the influence of users' characteristics on their ability to detect phishing emails?" In: *Advanced Computer and Communication Engineering Technology: Proceedings of the 1st International Conference on Communication and Computer Engineering*. Springer, pp. 949–962.

Alsharif, Maher, Shailendra Mishra, and Mohammed AlShehri (2022). "Impact of Human Vulnerabilities on Cybersecurity." In: *Computer Systems Science & Engineering* 40.3.

Alshenqeeti, Hamza (2014). "Interviewing as a data collection method: A critical review". In: *English linguistics research* 3.1, pp. 39–45.

Ampatzoglou, Apostolos, Stamatia Bibi, Paris Avgeriou, Marijn Verbeek, and Alexander Chatzigeorgiou (2019). "Identifying, categorizing and mitigating threats to validity in software engineering secondary studies". In: *Information and Software Technology* 106, pp. 201–230.

Anderson, Catherine L and Ritu Agarwal (2010). "Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions". In: *MIS quarterly*, pp. 613–643.

Androulidakis, Iosif and Gorazd Kandus (2011a). "A survey on saving personal data in the mobile phone". In: *2011 Sixth International Conference on Availability, Reliability and Security*. IEEE, pp. 633–638.

– (2011b). "Bluetooth® usage among students as an indicator of security awareness and feeling". In: *Proceedings ELMAR-2011*. IEEE, pp. 157–160.

Antignac, Thibaud, David Sands, and Gerardo Schneider (2017). "Data minimisation: a language-based approach". In: *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, pp. 442–456.

Bada, Maria and Jason RC Nurse (2019). "Developing cybersecurity education and awareness programmes for small and medium-sized enterprises (SMEs)". In: *Information & Computer Security* 27.3, pp. 393–410.

Baker, Sarah Elsie and Rosalind Edwards (2012). "How many qualitative interviews is enough". In.

Ball, Kirstie (2010). "Workplace surveillance: An overview". In: *Labor History* 51.1, pp. 87–106.

Banken, A (2025). "Illuminating ShadowIT: A Study to Understand and Assess Unauthorised IT Behaviour [Master's thesis, Utrecht University]". In: Utrecht University Student Theses Repository.

Baumeister, Roy F and Mark R Leary (1997). "Writing narrative literature reviews". In: *Review of general psychology* 1.3, pp. 311–320.

Baumer N, Frueh J (Nov. 2021). *What is neurodiversity?* URL: https://www.health.harvard.edu/blog/what-is-neurodiversity202111232645#:~:text=Neurodiversity%20describes%20the%20idea%20that,are%20not%20viewed%20as%20deficits. (visited on 01/09/2025).

Bendovschi, Andreea (2015). "Cyber-attacks–trends, patterns and security countermeasures". In: *Procedia Economics and Finance* 28, pp. 24–31.

Bessette, Douglas L, Lauren A Mayer, Bryan Cwik, Martin Vezér, Klaus Keller, Robert J Lempert, and Nancy Tuana (2017). "Building a values-informed mental model for New Orleans climate risk management". In: *Risk Analysis* 37.10, pp. 1993–2004.

Bitton, Ron, Andrey Finkelshtein, Lior Sidi, Rami Puzis, Lior Rokach, and Asaf Shabtai (2018). "Taxonomy of mobile users' security awareness". In: *Computers & Security* 73, pp. 266–293.

Brereton, Pearl, Barbara A Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil (2007). "Lessons from applying the systematic literature review process within the software engineering domain". In: *Journal of systems and software* 80.4, pp. 571–583.

Bruijn, Hans de and Marijn Janssen (2017). "Building cybersecurity awareness: The need for evidence-based framing strategies". In: *Government Information Quarterly* 34.1, pp. 1–7.

Bulgurcu, Burcu, Hasan Cavusoglu, and Izak Benbasat (2010). "Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness". In: *MIS quarterly*, pp. 523–548.

Chang, Lennon YC and Nicholas Coppel (2020). "Building cyber security awareness in a developing country: lessons from Myanmar". In: *Computers & Security* 97, p. 101959.

Cho, Jin-Hee, Hasan Cam, and Alessandro Oltramari (2016). "Effect of personality traits on trust and risk to phishing vulnerability: Modeling and analysis". In: *2016 IEEE International*

Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). IEEE, pp. 7–13.

Chua, Hui Na, Siew Fan Wong, Yeh Ching Low, and Younghoon Chang (2018). "Impact of employees' demographic characteristics on the awareness and compliance of information security policy in organizations". In: *Telematics and Informatics* 35.6, pp. 1770–1780.

Conteh, Nabie Y and Paul J Schmick (2021). "Cybersecurity risks, vulnerabilities, and countermeasures to prevent social engineering attacks". In: *Ethical hacking techniques and countermeasures for cybercrime prevention*, pp. 19–31.

Corbin, Kenneth (Aug. 2015). *CIOs vastly underestimate extent of Shadow IT*. URL: https://www.cio.com/article/244776/cios-vastly-underestimate-extent-of-shadow-it.html%20%5Chttps://www.cio.com/article/244776/cios-vastly-underestimate-extent-of-shadow-it.html (visited on 01/10/2024).

Craigen, Dan, N Diakun-Thibault, and R Purse (2014). "Defining Cybersecurity. Technology Innovation Management Review". In: *Technology Innovation Management Review*.

Crossler, Robert E and France Bélanger (2009). "The Effects of Security Education Training and Awareness Programs and Individual Characteristics on End User Security Tool Usage." In: *Journal of Information System Security* 5.3.

Darwish, Ali, Ahmed El Zarka, and Fadi Aloul (2012). "Towards understanding phishing victims' profile". In: *2012 international conference on computer systems and industrial informatics*. IEEE, pp. 1–5.

DeChurch, Leslie A and Jessica R Mesmer-M (2010). "Measuring shared team mental models: A meta-analysis." In: *Group dynamics: Theory, research, and practice* 14.1, p. 1.

Department for Digital Culture, Media Sport (Mar. 2020). *Cyber Security Breaches Survey 2020*. URL: https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2020/cyber-security-breaches-survey-2020 (visited on 01/14/2024).

Derbyshire, Richard, Benjamin Green, Daniel Prince, Andreas Mauthe, and David Hutchison (2018). "An analysis of cyber security attack taxonomies". In: *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, pp. 153–161.

Doyle, Emma EH, Sara E Harrison, Stephen R Hill, Matt Williams, Douglas Paton, and Ann Bostrom (2022). "Eliciting mental models of science and risk for disaster communication: A scoping review of methodologies". In: *International Journal of Disaster Risk Reduction* 77, p. 103084.

Doyle, James K, Michael J Radzicki, and W Scott Trees (2008). "Measuring change in mental models of complex dynamic systems". In: *Complex decision making: Theory and practice*, pp. 269–294.

Drevin, Lynette, Hennie A Kruger, and Tjaart Steyn (2007). "Value-focused assessment of ICT security awareness in an academic environment". In: *Computers & Security* 26.1, pp. 36–43.

Farahbod, Kamy, Conrad Shayo, and Jay Varzandeh (2020). "Cybersecurity indices and cybercrime annual loss and economic impacts". In: *Journal of Business and Behavioral Sciences* 32.1, pp. 63–71.

Farooq, Ali, Johanna Isoaho, Seppo Virtanen, and Jouni Isoaho (2015). "Information security awareness in educational institution: An analysis of students' individual factors". In: *2015 IEEE Trustcom/BigDataSE/ISPA*. IEEE, pp. 352–359.

Feldt, Robert and Ana Magazinius (2010). "Validity threats in empirical software engineering research-an initial survey." In: *Seke*, pp. 374–379.

Ferrari, Rossella (2015). "Writing narrative style literature reviews". In: *Medical writing* 24.4, pp. 230–235.

Ferreira, Daniel Jorge, Nuno Mateus-Coelho, and Henrique S Mamede (2023). "Methodology for predictive cyber security risk assessment (PCSRA)". In: *Procedia Computer Science* 219, pp. 1555–1563.

Furnell, Steven and Nathan Clarke (2012). "Power to the people? The evolving recognition of human aspects of security". In: *computers & security* 31.8, pp. 983–988.

Furnell, Steven, Harry Heyburn, Andrew Whitehead, and Jayesh Navin Shah (2020). "Understanding the full cost of cyber security breaches". In: *Computer fraud & security* 2020.12, pp. 6–12.

Fürstenau, Daniel and Hannes Rothe (2014). "Shadow IT Systems: Discerning the Good and the Evil". In: *ECIS*.

González-Pereira, Borja, Vicente P Guerrero-Bote, and Félix Moya-Anegón (2010). "A new approach to the metric of journals' scientific prestige: The SJR indicator". In: *Journal of informetrics* 4.3, pp. 379–391.

Greitzer, Frank L, Jeremy R Strozer, Sholom Cohen, Andrew P Moore, David Mundie, and Jennifer Cowley (2014). "Analysis of unintentional insider threats deriving from social engineering exploits". In: *2014 IEEE Security and Privacy Workshops*. IEEE, pp. 236–250.

Grenier, Robin S and Dana Dudzinska-P (2015). "A conceptual model for eliciting mental models using a composite methodology". In: *Human Resource Development Review* 14.2, pp. 163–184.

Gusenbauer, Michael and Neal R Haddaway (2020). "Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources". In: *Research synthesis methods* 11.2, pp. 181–217.

Györy, Andreas, Anne Cleven, Falk Uebernickel, and Walter Brenner (2012). "Exploring the shadows: IT governance approaches to user-driven innovation". In.

Haag, Steffi and Andreas Eckhardt (2017). "Shadow it". In: *Business & Information Systems Engineering* 59, pp. 469–473.

Hadlington, Lee, Maša Popovac, Helge Janicke, Iryna Yevseyeva, and Kevin Jones (2019). "Exploring the role of work identity and work locus of control in information security awareness". In: *Computers & Security* 81, pp. 41–48.

Halevi, Tzipora, Nasir Memon, James Lewis, Ponnurangam Kumaraguru, Sumit Arora, Nikita Dagar, Fadi Aloul, and Jay Chen (2016). "Cultural and psychological factors in cyber-security". In: *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services*, pp. 318–324.

He, Wu and Zuopeng Zhang (2019). "Enterprise cybersecurity training and awareness programs: Recommendations for success". In: *Journal of Organizational Computing and Electronic Commerce* 29.4, pp. 249–257.

Henshel, Diane, Char Sample, Mariana Cains, and Blaine Hoffman (2016). "Integrating cultural factors into human factors framework and ontology for cyber attackers". In: *Advances in Human Factors in Cybersecurity: Proceedings of the AHFE 2016 International Conference on Human Factors in Cybersecurity, July 27-31, 2016, Walt Disney World®, Florida, USA*. Springer, pp. 123–137.

Hijji, Mohammad and Gulzar Alam (2021). "A multivocal literature review on growing social engineering based cyber-attacks/threats during the COVID-19 pandemic: challenges and prospective solutions". In: *Ieee Access* 9, pp. 7152–7169.

Hovav, Anat and John D'Arcy (2012). "Applying an extended model of deterrence across cultures: An investigation of information systems misuse in the US and South Korea". In: *Information & Management* 49.2, pp. 99–110.

IBM (2015). *IBM 2015 Cyber Security Intelligence Index*. URL: https://www.ospi.es/export/sites/ospi/documents/IBM_2015_CyberSecurity_Intelligence_Index.pdf (visited on 01/15/2024).

Ifinedo, Princely (2014). "Information systems security policy compliance: An empirical study of the effects of socialisation, influence, and cognition". In: *Information & Management* 51.1, pp. 69–79.

Jagatic, Tom N, Nathaniel A Johnson, Markus Jakobsson, and Filippo Menczer (2007). "Social phishing". In: *Communications of the ACM* 50.10, pp. 94–100.

Jahan, Nusrat, Sadiq Naveed, Muhammad Zeshan, and Muhammad A Tahir (2016). "How to conduct a systematic review: a narrative literature review". In: *Cureus* 8.11.

Jansen, Floris, Katsiaryna Labunets, Slinger Jansen, and Niels Wagenaar (2023). "Mental models of Shadow IT". In: Utrecht University.

Jeong, Jongkil, Joanne Mihelcic, Gillian Oliver, and Carsten Rudolph (2019). "Towards an improved understanding of human factors in cybersecurity". In: *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*. IEEE, pp. 338–345.

Johnson-Laird, Philip N, Vittorio Girotto, and Paolo Legrenzi (1998). "Mental models: a gentle guide for outsiders". In: *Sistemi Intelligenti* 9.68, p. 33.

Jones, Natalie A, Helen Ross, Timothy Lynam, Pascal Perez, and Anne Leitch (2011). "Mental models: an interdisciplinary synthesis of theory and methods". In: *Ecology and society* 16.1.

Junglas, Iris and Christiane Spitzmuller (2006). "Personality traits and privacy perceptions: an empirical study in the context of location-based services". In: *2006 International Conference on Mobile Business*. IEEE, pp. 36–36.

Kajzer, Mitchell, John D'Arcy, Charles R Crowell, Aaron Striegel, and Dirk Van Bruggen (2014). "An exploratory investigation of message-person congruence in information security awareness campaigns". In: *Computers & security* 43, pp. 64–76.

Kapko, Matt (Oct. 2, 2023). *Cyber investments on pace to reach* $215 Bin 2024 : Gartner$. URL: https://www.cybersecuritydive.com/news/cyber-spending-gartner-moodys/695343/ (visited on 01/17/2024).

Khando, Khando, Shang Gao, Sirajul M Islam, and Ali Salman (2021). "Enhancing employees information security awareness in private and public organisations: A systematic literature review". In: *Computers & security* 106, p. 102267.

Kirlappos, Iacovos, Simon Parkin, and M Angela Sasse (2015). "" Shadow security" as a tool for the learning organization". In: *Acm Sigcas Computers and Society* 45.1, pp. 29–37.

Kirsch, Laurie and Scott Boss (2007). "The last line of defense: motivating employees to follow corporate security guidelines". In: *ICIS 2007 proceedings*, p. 103.

Klotz, Stefan, Andreas Kopper, Markus Westner, and Susanne Strahringer (2019). "Causing factors, outcomes, and governance of Shadow IT and business-managed IT: a systematic literature review". In: *International Journal of Information Systems and Project Management* 7.1, pp. 15–43.

Kopper, Andreas and Markus Westner (2016). "Towards a taxonomy for shadow IT". In.

Kraus, Sascha, Carolin Palmer, Norbert Kailer, Friedrich Lukas Kallinger, and Jonathan Spitzer (2019). "Digital entrepreneurship: A research agenda on new business models for the twenty-first century". In: *International Journal of Entrepreneurial Behavior & Research* 25.2, pp. 353–375.

Kretzer, Martin and Alexander Maedche (2014). *Generativity of business intelligence platforms: a research agenda guided by lessons from shadow IT*. Univ.

Krombholz, Katharina, Karoline Busse, Katharina Pfeffer, Matthew Smith, and Emanuel Von Zezschwitz (2019). "" If HTTPS Were Secure, I Wouldn't Need 2FA"-End User and Administrator Mental Models of HTTPS". In: *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, pp. 246–263.

Kruger, Hennie A, Lynette Drevin, and Tjaart Steyn (2006). "A Framework for Evaluating ICT Security Awareness." In: *ISSA*, pp. 1–11.

Kruger, Hennie A and Wayne D Kearney (2006). "A prototype for assessing information security awareness". In: *Computers & security* 25.4, pp. 289–296.

Kumaraguru, Ponnurangam, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong (2010). "Teaching Johnny not to fall for phish". In: *ACM Transactions on Internet Technology (TOIT)* 10.2, pp. 1–31.

Kure, Halima Ibrahim, Shareeful Islam, Mustansar Ghazanfar, Asad Raza, and Maruf Pasha (2022). "Asset criticality and risk prediction for an effective cybersecurity risk management of cyber-physical system". In: *Neural Computing and Applications* 34.1, pp. 493–514.

Lacey, David (2010). "Understanding and transforming organizational security culture". In: *Information Management & Computer Security* 18.1, pp. 4–13.

Lallie, Harjinder Singh, Lynsay A Shepherd, Jason RC Nurse, Arnau Erola, Gregory Epiphaniou, Carsten Maple, and Xavier Bellekens (2021). "Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic". In: *Computers & security* 105, p. 102248.

Lawson, Patrick, Olga Zielinska, Carl Pearson, and Christopher B Mayhorn (2017). "Interaction of personality and persuasion tactics in email phishing attacks". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1331–1333.

Leach, John (2003). “Improving user security behaviour”. In: *Computers & Security* 22.8, pp. 685–692.

Li, Ling, Wu He, Li Xu, Ivan Ash, Mohd Anwar, and Xiaohong Yuan (2019). “Investigating the impact of cybersecurity policy awareness on employees’ cybersecurity behavior”. In: *International Journal of Information Management* 45, pp. 13–24.

Mallmann, Gabriela Labres, Aline de Vargas Pinto, and Antônio Carlos Gastaud Maçada (2019). “Shedding Light on Shadow IT: Definition, Related Concepts, and Consequences”. In: *Information Systems for Industry 4.0: Proceedings of the 18th Conference of the Portuguese Association for Information Systems*. Springer, pp. 63–79.

Martocchio, Joseph J and Timothy A Judge (1997). “Relationship between conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy.” In: *Journal of applied psychology* 82.5, p. 764.

McCormac, Agata, Tara Zwaans, Kathryn Parsons, Dragana Calic, Marcus Butavicius, and Malcolm Pattinson (2017). “Individual differences and information security awareness”. In: *Computers in Human Behavior* 69, pp. 151–156.

McCrae, Robert R and Oliver P John (1992). “An introduction to the five-factor model and its applications”. In: *Journal of personality* 60.2, pp. 175–215.

McGrath, Cormac, Per J Palmgren, and Matilda Liljedahl (2019). “Twelve tips for conducting qualitative research interviews”. In: *Medical teacher* 41.9, pp. 1002–1006.

Moody, Gregory D, Dennis F Galletta, and Brian Kimball Dunn (2017). “Which phish get caught? An exploratory study of individuals susceptibility to phishing”. In: *European Journal of Information Systems* 26, pp. 564–584.

Morgan, Steve (Dec. 2022). *Cybersecurity research: All in one place*. URL: `https://cybersecurityventures.com/research/` (visited on 01/15/2024).

Mourby, Miranda, Elaine Mackey, Mark Elliot, Heather Gowans, Susan E Wallace, Jessica Bell, Hannah Smith, Stergios Aidinlis, and Jane Kaye (2018). “Are ‘pseudonymised’data always personal data? Implications of the GDPR for administrative data research in the UK”. In: *Computer Law & Security Review* 34.2, pp. 222–233.

Nayak, MSDP and KA Narayan (2019). “Strengths and weaknesses of online surveys”. In: *technology* 6.7, pp. 0837–2405053138.

Öğütçü, Gizem, Özlem Müge Testik, and Oumout Chouseinoglou (2016). “Analysis of personal information security behavior and awareness”. In: *Computers & Security* 56, pp. 83–93.

Parker, Heather J and Stephen V Flowerday (2020). “Contributing factors to increased susceptibility to social media phishing attacks”. In: *South African Journal of Information Management* 22.1, pp. 1–10.

Parrish Jr, James L, Janet L Bailey, and James F Courtney (2009). “A personality based model for determining susceptibility to phishing attacks”. In: *Little Rock: University of Arkansas*, pp. 285–296.

Parsons, Kathryn, Dragana Calic, Malcolm Pattinson, Marcus Butavicius, Agata McCormac, and Tara Zwaans (2017). “The human aspects of information security questionnaire (HAIS-Q): two further validation studies”. In: *Computers & Security* 66, pp. 40–51.

Parsons, Kathryn, Agata McCormac, Marcus Butavicius, Malcolm Pattinson, and Cate Jerram (2014). "Determining employee awareness using the human aspects of information security questionnaire (HAIS-Q)". In: *Computers & security* 42, pp. 165–176.

Patten, Mildred (2016). *Questionnaire research: A practical guide*. routledge.

Pattinson, Malcolm, Marcus Butavicius, Kathryn Parsons, Agata McCormac, and Dragana Calic (2015). "Factors that influence information security behavior: An Australian web-based study". In: *Human Aspects of Information Security, Privacy, and Trust: Third International Conference, HAS 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015. Proceedings 3*. Springer, pp. 231–241.

Pattinson, Malcolm, Cate Jerram, Kathryn Parsons, Agata McCormac, and Marcus Butavicius (2012). "Why do some people manage phishing e-mails better than others?" In: *Information Management & Computer Security* 20.1, pp. 18–28.

Rafaiani, Giulia, Massimo Battaglioni, Marco Baldi, Franco Chiaraluce, Giovanni Libertini, Luca Spalazzi, and Giovanni Cancellieri (2021). "A functional approach to cyber risk assessment". In: *2021 AEIT International Annual Conference (AEIT)*. IEEE, pp. 1–6.

Rea-Guaman, AM, T San Feliu, JA Calvo-Manzano, and Isaac Daniel Sánchez-García (2018). "Systematic review: Cybersecurity risk taxonomy". In: *Trends and Applications in Software Engineering: Proceedings of the 6th International Conference on Software Process Improvement (CIMPS 2017) 6*. Springer, pp. 137–146.

Reeves, Andrew, Dragana Calic, and P Delfabbro (2021). ""Get a red-hot poker and open up my eyes, it's so boring" 1: Employee perceptions of cybersecurity training". In: *Computers & security* 106, p. 102281.

Reicher, Hannelore (2010). "Building inclusive education on social and emotional learning: Challenges and perspectives–a review". In: *International Journal of Inclusive Education* 14.3, pp. 213–246.

Ritchie, Jane and Liz Spencer (2002). "Qualitative data analysis for applied policy research". In: *Analyzing qualitative data*, pp. 173–194.

Safa, Nader Sohrabi, Mehdi Sookhak, Rossouw Von Solms, Steven Furnell, Norjihan Abdul Ghani, and Tutut Herawan (2015). "Information security conscious care behaviour formation in organizations". In: *Computers & Security* 53, pp. 65–78.

Salgado, Jesus F (2002). "The Big Five personality dimensions and counterproductive behaviors". In: *International journal of selection and assessment* 10.1-2, pp. 117–125.

Santos-Olmo, Antonio, Luis Enrique Sánchez, Ismael Caballero, Sara Camacho, and Eduardo Fernandez-Medina (2016). "The importance of the security culture in SMEs as regards the correct management of the security of their assets". In: *Future Internet* 8.3, p. 30.

Schmidt, Nikolaus, Bastian Zöller, and Christoph Rosenkranz (2016). "The clash of cultures in information technology outsourcing relationships: An institutional logics perspective". In: *Shared Services and Outsourcing: A Contemporary Outlook: 10th Global Sourcing Workshop 2016, Val d'Isère, France, February 16-19, 2016, Revised Selected Papers 10*. Springer, pp. 97–117.

Schoormann, Thorsten, Frederik Möller, Leona Chandra Kruse, and Boris Otto (2024). "BAUSTEIN—A design tool for configuring and representing design research". In: *Information Systems Journal*.

Sharif, Md Haris Uddin and Mehmood Ali Mohammed (2022). "A literature review of financial losses statistics for cyber security and future trend". In: *World Journal of Advanced Research and Reviews* 15.1, pp. 138–156.

Shaw, Ruey Shiang, Charlie C Chen, Albert L Harris, and Hui-Jou Huang (2009). "The impact of information richness on information security awareness training effectiveness". In: *Computers & Education* 52.1, pp. 92–100.

Sheehan, Barry, Finbarr Murphy, Martin Mullins, and Cian Ryan (2019). "Connected and autonomous vehicles: A cyber-risk classification framework". In: *Transportation research part A: policy and practice* 124, pp. 523–536.

Sheng, Steve, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs (2010). "Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions". In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 373–382.

Shropshire, Jordan, Merrill Warkentin, Allen Johnston, and Mark Schmidt (2006). "Personality and IT security: An application of the five-factor model". In: *AMCIS 2006 Proceedings*, p. 415.

Silic, Mario and Andrea Back (2014). "Shadow IT–A view from behind the curtain". In: *Computers & Security* 45, pp. 274–283.

Silic, Mario and Paul Benjamin Lowry (2020). "Using design-science based gamification to improve organizational security training and compliance". In: *Journal of management information systems* 37.1, pp. 129–161.

Siponen, Mikko T (2001). "Five dimensions of information security awareness". In: *SIGCAS Comput. Soc.* 31.2, pp. 24–29.

Smahel, David, B Bradford Brown, and Lukas Blinka (2012). "Associations between online friendship and Internet addiction among adolescents and emerging adults." In: *Developmental psychology* 48.2, p. 381.

Snyder, Hannah (2019). "Literature review as a research methodology: An overview and guidelines". In: *Journal of business research* 104, pp. 333–339.

Son, Jai-Yeol (2011). "Out of fear or desire? Toward a better understanding of employees' motivation to follow IS security policies". In: *Information & Management* 48.7, pp. 296–302.

Spears, Janine L and Henri Barki (2010). "User participation in information systems security risk management". In: *MIS quarterly*, pp. 503–522.

Spector, Paul E (1988). "Development of the work locus of control scale". In: *Journal of occupational psychology* 61.4, pp. 335–340.

Staggers, Nancy and Anthony F. Norcio (1993). "Mental models: concepts for human-computer interaction research". In: *International Journal of Man-machine studies* 38.4, pp. 587–605.

Stanton, Jeffrey M, Kathryn R Stam, Paul Mastrangelo, and Jeffrey Jolton (2005). "Analysis of end user security behaviors". In: *Computers & security* 24.2, pp. 124–133.

Stewart, Geordie and David Lacey (2012). "Death by a thousand facts: Criticising the technocratic approach to information security awareness". In: *Information Management & Computer Security* 20.1, pp. 29–38.

Talib, Shuhaili, Nathan L Clarke, and Steven M Furnell (2010). "An analysis of information security awareness within home and work environments". In: *2010 International Conference on Availability, Reliability and Security*. IEEE, pp. 196–203.

Tsohou, Aggeliki, Maria Karyda, and Spyros Kokolakis (2015). "Analyzing the role of cognitive and cultural biases in the internalization of information security policies: Recommendations for information security awareness programs". In: *Computers & security* 52, pp. 128–141.

Turner III, Daniel W and Nicole Hagstrom-Schmidt (2022). "Qualitative interview design". In: *Howdy or Hello? Technical and professional communication*.

Uebelacker, Sven and Susanne Quiel (2014). "The social engineering personality framework". In: *2014 Workshop on Socio-Technical Aspects in Security and Trust*. IEEE, pp. 24–30.

Utrecht University (2024). *How to write an informed consent form*. URL: `https://www.uu.nl/en/research/research-data-management/guides/legal-considerations/how-to-write-an-informed-consent-form` (visited on 02/29/2024).

Valentine, J Andrew (2006). "Enhancing the employee security awareness model". In: *Computer fraud & security* 2006.6, pp. 17–19.

Van Steen, Tommy, Emma Norris, Kirsty Atha, and Adam Joinson (2020). "What (if any) behaviour change techniques do government-led cybersecurity awareness campaigns use?" In: *Journal of Cybersecurity* 6.1, tyaa019.

Voigt, Paul and Axel Von dem Bussche (2017). "The eu general data protection regulation (gdpr)". In: *A practical guide, 1st ed., Cham: Springer International Publishing* 10.3152676, pp. 10–5555.

Von Solms, Rossouw and Johan Van Niekerk (2013). "From information security to cyber security". In: *computers & security* 38, pp. 97–102.

Wash, Rick (2010). "Folk models of home computer security". In: *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pp. 1–16.

Wash, Rick and Emilee Rader (2011). "Influencing mental models of security: a research agenda". In: *Proceedings of the 2011 New Security Paradigms Workshop*, pp. 57–66.

WEF (Jan. 2024). *The Global Risks Report 2024*. World Economic Forum. URL: `https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf` (visited on 03/01/2024).

Whitman, Michael E, Anthony M Townsend, and Robert J Aalberts (2001). "Information systems security and the need for policy". In: *Information security management: Global challenges in the new millennium*, pp. 9–18.

Wiley, Ashleigh, Agata McCormac, and Dragana Calic (2020). "More than the individual: Examining the relationship between culture and Information Security Awareness". In: *Computers & security* 88, p. 101640.

Willison, Robert and Merrill Warkentin (2013). "Beyond deterrence: An expanded view of employee computer abuse". In: *MIS quarterly*, pp. 1–20.

Wohlin, Claes, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén (2012). *Experimentation in software engineering*. Springer Science & Business Media.

Wong, Geoff, Trish Greenhalgh, Gill Westhorp, Jeanette Buckingham, and Ray Pawson (2013). "RAMESES publication standards: Meta-narrative reviews". In: *Journal of Advanced Nursing* 69.5, pp. 987–1004.

Wong, Lai-Wan, Voon-Hsien Lee, Garry Wei-Han Tan, Keng-Boon Ooi, and Amrik Sohal (2022). "The role of cybersecurity and policy awareness in shifting employee compliance attitudes: Building supply chain capabilities". In: *International Journal of Information Management* 66, p. 102520.

Yan, Zheng, Thomas Robertson, River Yan, Sung Yong Park, Samantha Bordoff, Quan Chen, and Ethan Sprissler (2018). "Finding the weakest links in the weakest link: How well do undergraduate students make cybersecurity judgment?" In: *Computers in Human Behavior* 84, pp. 375–382.

Zaytsev, Anton, Anatoly Malyuk, and Natalia Miloslavskaya (2017). "Critical analysis in the research area of insider threats". In: *2017 IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)*. IEEE, pp. 288–296.

Zhang, Li-fang (2002). "Measuring thinking styles in addition to measuring personality traits?" In: *Personality and Individual Differences* 33.3, pp. 445–458.

Zhou, Xin, Yuqin Jin, He Zhang, Shanshan Li, and Xin Huang (2016). "A map of threats to validity of systematic literature reviews in software engineering". In: *2016 23rd Asia-Pacific Software Engineering Conference (APSEC)*. IEEE, pp. 153–160.

A Appendix

A.1 Interview Guideline

Version 2, 18-07-2024

Introduction
This interview will not take more than 45 minutes. As already communicated, this interview will be recorded unless you specify otherwise. All comments, responses, and recordings will be kept strictly confidential, which means that your responses will only be shared with the direct research team. Data from your interviews used within the final research will be pseudonymised to ensure that you cannot be directly identified. Before we start, do you have any questions already?

Introductory questions | organisation

- Can you give demographics about your organisation (so how many people are employed, what are the main activities or services of the organisation, what is the business domain of your organisation)?
- More specifically, can you give some demographics about your department (again, how many people are employed, what are the main activities or services of the department, and where is your department positioned within the organisation)?
- More focused on you, what is your position within the company, and what are your main activities?
- **If applicable/needed**: Can you provide additional information about your company?

Introductory questions | researcher

- Please state your highest completed level of education
  - Was this education related to cybersecurity?
- What cybersecurity research have you done in the past?
- Are you currently researching cybersecurity, and if so, are you willing to elaborate on this?
- In which research institution are you currently conducting research?
- How many years of experience do you have on this topic (cybersecurity in general and more ISA/mental model focused)?

Questions on cybersecurity

- **Only for organisational experts**: What kind of cybersecurity experience do you have?
- How would you describe cybersecurity?

- Have you ever experienced or observed a cybersecurity breach? (If so, elaborate.)
- Which of the following different cybersecurity types do you think are the main threat to your/an organisation? Also, which are the least threatening?
  - **Intentional destruction** (Technical expertise together with a strong intention to harm the organisation)
  - **Detrimental misuse** (minimal technical expertise but intention to harm through annoyance, harassment, rule-breaking, etc.)
  - **Dangerous tinkering** (Technical expertise but no clear intention to harm the organisation)
  - **Naïve mistakes** (minimal technical expertise and no clear intention to harm)
  - **Outsider threat**s (threats focused on systems instead of individuals)

Questions focusing on SRQ 1

- How familiar are you with information security awareness (ISA for short)?
  - **If familiar**, let them explain their meaning and, if different from ours, add to it.
  - **If not**, briefly address / explain what ISA is.
    - *ISA is the cognizance of an employee about the organisation's Information Security Policy (ISP), and their general knowledge of information security. ISA is essential to mitigate cybersecurity risk, as cyberattacks are shifting away from the exploitation of systems or information security vulnerabilities and moving towards human factors.*
- What are the problems related to ISA and the secure behaviour among employees in general? Examples of problems are:
  - *Lack of Clear Methodology*
  - *Motivation*
  - *Organisational Culture*
  - *Constant Change*
  - *Pressure and Stress*
  - *Threat Appraisal & Self-Efficacy*
  - *Keep Informed*
  - *Medium Richness*
  - *Program Personalisation*
  - *Limited Resources*
  - *Technical Standpoint*
- **(Only to an expert, not a researcher)**:
  - Which of those are specific to your organisation?
  - How does your organisation raise ISA and facilitate secure behaviour among employees?
- **(Only to a researcher, not an expert)**:
  - Which of those would you say are most apparent within organisations?
  - How do organisations raise ISA and facilitate secure behaviour among employees?

Questions focusing on SRQ 2

- Understanding employees is important to improve ISA within organisations. However, for this, it is necessary to assess the current attitude of employees towards cybersecurity. Related to this assessment are mental models. Are you aware of the term "mental models"?
  - **If so**, let them explain their meaning and, if different from ours, add to it

  - **If not**, shortly address/explain what mental models are.
    * *Mental models are a concept widely accepted and applied in human-computer interaction research. These models are context-dependent psychological representations of external reality, often regarded as incomplete and inconsistent. The structure of a mental model, the model inside a user's mind of how things work, corresponds to the structure of what they represent. These simplified psychological representations of reality allow for interaction with the real world, which can be used to anticipate events, reason, or underlie explanations.*
- Which of the following factors do you think influence the difference in mental models on the same topic between individuals?
  - Personality Traits (The big 5)
    * Conscientiousness
    * Extraversion
    * Agreeableness
    * Openness to Experience
    * Neuroticism
  - Demographics
    * Age
    * Gender
    * Working Industry
    * Education level
    * Culture
    * Ethnic background/group
  - Work Location (home or in office)
- Do you think these, or potentially others, such as job rank, employee-related aspects, need to be considered within cybersecurity?
  - **If so**, can you elaborate on which ones you think are more important?

Questions focusing on SRQ 3 and 4

- Do you think understanding an individual's mental model is important for measuring ISA and secure behaviour?
- Do you have prior experience with measuring mental models?
  - **If so**, could you elaborate on how you measured mental models? How did you quantify the metrics of the Mental Model, and on what did you base your decisions? What exactly did you measure, and how did you measure it?
  - **If not**, could you think of a method or way to measure an individual's mental model on ISA and behaviour? How would you select and quantify metrics on the relevant topic, and how would you measure them?
- In past studies, several methods for analysing and measuring end-user ISA and related behaviour are discussed. One method, covered in most papers, is the use of the KAB dimensions, so Knowledge, Attitude, and Behaviour dimensions. Together, these threefold dimensions are believed to shed light on ISA. However, do you think these dimensions can also be used to measure cybersecurity mental models more directly?
- Most past studies use self-reported Likert-scale questionnaires to obtain employees' KAB levels. Do you agree with this method, or do you see potential flaws with using a survey?

- **Independent of the answer above**, what would you say are other effective methods to collect KAB dimensions data on individuals?
  - Monitoring
  - Activity logging

Interview closing

- Do you still have any questions, or would you like to add anything?
- If not, I will analyze the recording along with the other interviews in the following weeks. If you want, I will send you a copy of the transcribed interview for you to review. Thank you for your time!

General probes used during the interview

- Would you please explain why?
- Could you give an example?
- I do not fully understand; could you repeat that?

A.2 Consent Form

Version 1, 19-03-2024

Interview on Cybersecurity assessment
You have been asked to participate in an interview as part of scientific research to assess prevalent employee mindsets towards Shadow IT usage in an organisational setting. Shadow IT is *" hardware, software, or services built, introduced, and/or used for the job without explicit approval or even knowledge of the organisation"* (Haag and Eckhardt, 2017). Using Shadow IT brings cybersecurity risks, which are sometimes not well understood by those creating them. Understanding the reasoning behind cybersecurity decisions is essential to assess prevalent mindsets. For this, mental models, or context-dependent representations of external reality, are important. This research tries to determine whether mental models of cybersecurity can be quantified into general taxonomies and guidelines. This research is conducted as a master's thesis project in collaboration with Utrecht University. The leading researcher of this project is Tjerk de Haan (t.dehaan@uu.nl), who is under the supervision of Kate Labunets (k.labunets@uu.nl).

What are the details of the interview?
The interview will have a time limit of half an hour. During this timeframe, the researcher will guide the conversation, asking you to provide insights into how you would approach the selection and what factors influence this decision. The interview will take place in person at the location of your preference or can be conducted online. In addition, the audio of the conversation will be recorded. The sole purpose of the recording is to post-process the interview into reportable results. You will not receive any compensation for participating in the study.

What is expected from you?
You have been selected due to your experience and familiarity with either cybersecurity or mental models, or both. The questions we will ask do not have a right or a wrong answer. To draw sound conclusions, it would be helpful if you provided short reasoning. If you have previously looked more into the assessment of mental models within the field of cybersecurity, it is appreciated to talk about this experience, but not mandatory.

What happens to your data?
The interview aims to obtain information on cybersecurity awareness generalisable mental model assessments. However, there is an overhead of potentially personally identifiable information. This is dealt with accordingly. The audio recording will serve to obtain a transcript of the interview and will be disposed of within six months after the interview date. The non-anonymised transcript will be solely handled by the research team at Utrecht University and used for the analysis. The report will contain general information without company-specific details or privacy-sensitive information.

After analysis, the transcripts are pseudonymised. It means that any information pointing to you or your organisation will be de-identified and coded so that it can no longer lead to identification without additional information. The principal investigator will keep a link that identifies you to your coded information. However, this link will remain secure and will be available only to the principal investigator or selected members of the research team. Any information that can identify you will remain confidential. The pseudonym identifiers will be deleted after the end of the project, ensuring that the data is no longer considered personal.

Utrecht University will store the processed data and de-identified transcripts for at least 10 years for research integrity. The data is stored and processed solely within the EU. When third-party applications are used, they are required to have an appropriate data processing agreement with Utrecht University. The results will be published in a Master's thesis.

What are your rights?
Participation is voluntary, but very much appreciated. We are only allowed to collect data for our study if signed consent is given. Signing the form is voluntary, and choosing not to engage will not require any further steps. You can opt out anytime before or during the interview. Data collected from your interview will be immediately disregarded without any consequence. In addition, you can withdraw your consent after participating. Then your transcript will be removed from the dataset. However, data from your interview that is already processed into a global result, such as a percentage, will not be retracted.

Potential risk
During the interview, you can share organisation-specific details that might affect your relationship with your employer. To mitigate this risk and protect you, the organisation-specific information is not explicitly reported, and the company in question will be de-identified. However, external factors might still cause this risk to exist. Therefore, you have the right and are encouraged to review the transcript and retract or eliminate any statements that might cause a risk or that you do not want to share.

Approval and complaints
A Quick Ethics scan provided by the Science-Geo Ethics Review Board (SG ERB) at Utrecht University did not reveal any ethical problems for this research. In case you have a complaint about the execution or ethics of the study, send an email to the secretary of SG ERB: etc-beta-geo@uu.nl. If you have questions or complaints about the use of personal data or the data processing agreement, please email the Data Protection Officer of Utrecht University: privacy@uu.nl. The Data Protection Officer will also be able to inform you about the GDPR, which ensures the careful processing of personal data. If you would like to file a complaint with a third party, you can do so with the Dutch Data Protection Authority [1].

Want to know more?
If you want to know more about the study, have any comments, or wish to receive updates on the publication, please get in touch with Tjerk de Haan (Researcher/Interviewer) at t.dehaan@students.uu.nl or Kate Labunets (Project supervisor) at k.labunets@uu.nl.

Privacy policy I have read and understood the study information dated (X). I have been able to ask questions about the study, and my questions have been answered to my satisfaction.

- Yes

[1]https://www.autoriteitpersoonsgegevens.nl/en

- No

I consent voluntarily to participate in this study and understand that I can refuse to answer questions and withdraw from the study at any time.

- Yes
- No

I understand that the information I provide will be used for the report and publications in academic venues (like conferences or journals).

- Yes
- No

I understand that personal information collected about me that can identify me, such as my name or email address, will not be shared beyond the research team.

- Yes
- No

I agree that my information can be quoted in research outputs.

- Yes
- No

Your name

- Name:

Permission to record and store data

I give permission to record the interview for the sole purpose of producing a transcript.

- Yes
- No

I give additional permission that the research data, without any personal information that could identify me (not linked to me), may be shared with others and archived in UU's Your-Data so it can be used for future research and learning.

- Yes
- No

Your name

- Name:

A.3 Codebook

Table A.1: Interview Codebook with Coverage and Grounding

Codebook Levels				
Lvl 1	Lvl 2	Level 3	G	C
ISA and Behaviour			62	8
	Prevalent Challenges		46	8
		Organisational Culture	12	7
		Programme Personalisation	8	5
		Resource Availablity	7	5
		Technical Standpoint	6	5
		Employee's Motivation	5	5
		Clear Methodology	4	4
		Time and Stress	4	4
		Constant Change	3	2
		Risk Communication	3	2
		Medium Richness	2	2
		Self-Efficacy and Threat Appraisal	1	1
		Equal Prevalence	2	2
	Less Prevalent Challenges		6	5
		Constant Change	3	3
		Employee's Motivation	3	3
		Pressure and Stress	1	1
		Self-Efficacy and Threat Appraisal	1	1
	Additional Challenges		10	6
		Focus on the Human Aspect of Security	4	3
		Difference in Expected and Actual Behaviour	2	2
		Focus on Higher-Risk Individuals	2	2
		Measuring ISA Performance Results	2	2
	Influence Factors		13	7
		Lead by Example	7	5
		Education	6	5
MM			73	8
	Influential Factors		23	8
		Demographics	11	8
		Personality Traits	7	7
		Work Location	6	6
		Shared Importance	2	2
	Additional Factors		3	2
		Neurodiversity	3	2
	Assessment Challenges		10	6
		Individual-Level Assessment	6	6
		Individual-Level Assessment Costs	1	1
		Relation MM and Knowledge	1	1
		More than Behaviour Tracking	1	1
		Privacy Issues Regarding Data	1	1
		Context Influence	1	1
	KAB Suitability		7	6
		Knowledge Suitability	1	1
		Attitude and Behaviour Suitability	2	2
		General Suitability	4	3
	Survey Suitability		10	8
		Response Bias	7	6
		General Suitability	2	2
		Limited Results	2	2
		External Factor Influence	1	1
	Additional Methods		20	8
		Group-Perspective Preference	7	6
		Interviews	6	5
		Direct Observation	4	4
		Anonymous Open-Text	1	1
		Drawing	1	1
		Expert MM Comparison	1	1
		Peer Reviews Among Colleagues	1	1
		Seven Dimensions Model	1	1
		System-Generated Data	1	1

Note: G abbreviates grounding, the number of times a code occurred in total; C abbreviates coverage, the number of different experts mentioning the concept.

A.4 MM Questionnaire questions

Table A.2: MM Questionnaire questions

Mental Model	MM Summary	Question/Statement
Risk-Averse MM's		
Consequence-Avoidance Orientation	Avoiding potential negative consequences/ impact of shadow IT (e.g., confidentiality break, data leakage)	It is not necessary to consider potential risks before using an external IT solution for work.
		It is acceptable to use some IT solutions for work, even if I did not check whether my organisation has any security requirements relating to those solutions.
		It can be risky to use new IT solutions in the workplace without considering potential downsides.
		It is safer to avoid using any external IT solutions for work, even if they seem helpful.
Knowledge-Based Conservatism	Avoiding shadow IT usage based on the background expertise knowledge in cybersecurity-related topics	Personal technology expertise should not influence the choice of work-related IT solutions.
		Regardless of my IT knowledge, I am open to using new IT solutions at work.
		My knowledge of organizational IT security policies helps me decide whether to use new IT solutions.
		I can rely on my IT knowledge when deciding whether to use new IT solutions.
Risk Transfer Mindset	By using client-provided devices, shadow IT risks are avoided since the responsibility is transferred to the client	When managing a project using an IT solution not controlled by your organization, do you agree with this statement: It can be problematic to assume that an external organization should bear all related risks, which occurred while working on their behalf.
		When managing a project using an IT solution not controlled by your organization, do you agree with this statement: Imposing IT risks to another organization is not a safe or responsible practice.
		When managing a project using an IT solution not controlled by your organization, do you agree with this statement: I believe transferring risks to an external organization is an acceptable way to manage potential risks that can occur.
		When managing a project using an IT solution not controlled by your organization, do you agree with this statement: It is appropriate for an external organization to bear risks ①associated with IT solutions they request.
Cautious Seasoned Judgement	Avoiding shadow IT usage based on the previous personal or group experience	My experience with cybersecurity incidents is irrelevant to my current IT usage.
		Considering past IT-related mistakes and incidents from experienced coworkers limits work performance.
		Applying personal knowledge and experience is crucial when making IT-related decisions for my team.
		When selecting IT solutions, we should value cybersecurity insights from experienced coworkers.

Risk-Taking MM's

Common Sense Fallacy	Using shadow IT based on common sense about IT usage	Relying on common sense alone provides incomplete protection against unintended harms.
		My organization needs more knowledge of security best practices in the workplace.
		IT security practices are common knowledge for everyone in my organization.
		If an IT solution seems straightforward to use, it is appropriate for work tasks.
Illusion of Self-Sufficiency	Using shadow IT while believing that all IT (incl. shadow IT) is managed by the IT department	I suspect that some of the IT solutions provided by my organization are not secure.
		Not every IT solution I install is managed by my organization.
		My employer provides all the IT solutions I need, so I never require other software or hardware other than what they provide.
		It is safe to use any IT solution on my work computer because my organization secures all of them.
Misguided Sense of Protection	Using shadow IT while believing that the existing security countermeasures protect against potential problems	Detection systems in organisations are limited in their ability to spot all cybersecurity threats.
		Organizational cybersecurity systems can miss some threats.
		My organization has good security practices for risk prevention, which protects me from any harm that could be caused by IT solutions.
		Relying solely on the organization's cybersecurity measures for IT is sufficient to safeguard me from all IT risks.
Performance-Driven Rule Bending	Using shadow IT consciously to meet goals	Circumventing IT policies for work objectives can increase the chance of security risks.
	Goals (deadlines, better quality)	Compliance with IT policies is more important than completing work tasks.
		Using unapproved IT solutions to get my work goals done more quickly is acceptable.
		Bending IT policy rules to meet client expectations can be necessary.
Longevity-Based Invincibility	Adopting shadow IT which is believed to be safe based on team experience/culture	An incident-free history does not indicate an IT solution's security.
		Long-term IT solution usage offers no security guarantee.
		Established IT practices in my organization are secure and do not need to be questioned.
		A history of problem-free usage eliminates the need to review IT solutions.
Cost-Driven Compromise	Using shadow IT based on saving resources	It is important to get IT solutions approved by the IT department, even if a free alternative is available.
		High costs and lengthy procurement processes should not compromise IT solution choices.
		Using free IT solutions is a good choice to save time and money, because I do not need to wait for approval by the IT department.
		Choosing cheap or free IT solutions is acceptable, even if they might have lower security standards.