Cyber security awareness campaigns in the age of Artificial Intelligence (AI): Do they still fail to change behaviour?

Maria Bada[1] and Jason R.C. Nurse[2]

[1] Queen Mary University of London, UK
m.bada@qmul.ac.uk
[2] University of Kent, UK
j.r.c.nurse@kent.ac.uk

Abstract. Cyber security awareness and training programmes are the primary mechanism used to manage risks associated with people within organisations and across countries. Ten years ago, we examined such campaigns and identified several factors that explained why they often fail to bring about the intended behaviour change. Since then, the security landscape has evolved substantially, with developments such as generative artificial intelligence (AI), remote working, and increasing security fatigue introducing new and significant risks. In this paper, we conduct an updated analysis to examine whether more recent security awareness campaigns have improved and are consequently more likely to result in behaviour change. Specifically, we present a ten year retrospective of national cyber security awareness campaigns (2015–2025) across five global regions. Through thematic analysis of campaign materials, we identify continued reliance on fear based, information-heavy messaging, inconsistent calls to action, and persistent structural barriers. We conclude that without realignment to behavioural drivers and real-world constraints, particularly in an AI-augmented threat landscape, awareness campaigns are likely to repeat familiar failures.

Keywords: Security awareness · Behaviour change · Human aspects · Awareness and training · Human risk management · Artificial intelligence

1 Introduction and background

While security is one of the most extensively studied subjects in history, it remains the focus of fierce debate, particularly in relation to today's increasingly digital world. One such debate concerns the role of the individual, human psychology, and more broadly, human-centred approaches in cyber security. Within academia, this discussion has matured since seminal works such as "Users Are

This work is licensed under a Creative Commons "Attribution 4.0 International" license. To view a copy of this license visit https://creativecommons.org/licenses/by/4.0/deed.en.
©2026 Copyright held by the owner/author(s).

Not the Enemy" [1] and "Why Johnny Can't Encrypt" [40]. It could be argued that we have reached a point where the importance of understanding users, designing for their needs, and accounting for their limitations is well established.

The position adopted by industry and the media, however, is less clear. In 2025 alone, for example, we identified at least 10 articles in mainstream industry and media outlets that referred to users as "the weakest link" [36]. Motivated by this observation, and as a thought exercise given the growing prominence of generative AI platforms, we posed the question "Are humans the weakest link?" to Google's Gemini. While the initial response appeared balanced, further prompting for a more definitive answer resulted in, "*Yes, because attackers find it far easier to exploit human psychology through deception and urgency than to bypass sophisticated technical security measures.*" This serves as a simple illustration of security perceptions that persist and are frequently reinforced.

Within the people and cyber security domain, a question that has received substantial attention over the past two decades is how to encourage/sustain secure behaviours among individuals. Common examples include choosing strong passwords, adopting multi-factor authentication, and identifying scams and social engineering attempts [13,15,26]. Ten years ago, we approached this problem through the lens of awareness campaigns, examining the question of why cyber security awareness campaigns fail to change behaviour [3]. This focus was motivated by the observation that, while security awareness and training (SAT) campaigns are the primary method used by businesses and governments to influence behaviour, they rarely achieve the desired outcomes. This is largely because core components, such as personal, social, and environmental factors, are often neglected in the design and implementation stages [5,9,38]. More specifically, such initiatives tend to articulate risks and prescribe 'good' behaviours, but overlook other critical factors required to bring about sustained behaviour change.

Since our initial research, the cyber security landscape has changed dramatically. The Covid-19 pandemic ushered in an era of remote working, which contributed to a rise in cyber-attacks worldwide [27]. Threat actors have become increasingly innovative and brazen in their use of social engineering techniques, including attacks on organisational help desks, as well as ransomware campaigns involving multiple forms of extortion [4,30]. Artificial intelligence (AI) also poses a distinct set of challenges, both in terms of corporate and personal data exposure through shadow IT, and through its use by attackers as an enabling tool, for example to craft phishing messages or develop malware [35]. The harms experienced by victims of cyber-attacks are more significant than ever due to our increasingly connected organisations and personal lives [22,29]. These challenges are further compounded by the broader context in which individuals operate, characterised by increasing work and personal demands, heightened levels of information overload, and persistent experiences of digital and security fatigue.

In this article, we therefore undertake an updated analysis to examine whether current security awareness campaigns are more likely to result in behaviour change. This analysis considers the extent to which the security landscape has evolved, alongside more recent research in the people and security domain, in-

cluding personal, social, and environmental factors that influence behaviour. At the centre of our analysis is a series of national campaigns from 2015-2025. We further scope this study to AI-enabled threats, including deepfakes, phishing, or voice synthesis. This was done given the prominence of these threats today [25], as well as the space limitations of this article.

The remainder of the article is structured as follows. Section 2 outlines the methodology adopted for this study. Section 3 presents the results, while Section 4 reflects on them and wider research. Finally, in Section 5 we present implications for future campaigns, with a focus on behavioural aspects.

2 Methodology

This study adopts a qualitative, interpretive approach to examine how national cyber security awareness campaigns have incorporated AI-related threats into their messaging over the past decade. Consistent with prior work on awareness campaigns [3], the goal is not to measure behavioural outcomes but to assess the assumptions, framing, and behavioural logic embedded in campaign materials.

2.1 Information sources

Data were drawn from two primary sources. Firstly, national cyber security awareness campaign materials produced between 2015 and 2025 by government cyber security agencies and secondly national Computer Emergency Response Teams (CERTs). Materials included publicly available web pages, downloadable toolkits and posters, intended for citizens, small and medium-sized enterprises (SMEs), or the general workforce. Campaign materials were collected directly from official agency websites and archived resources to ensure authenticity and consistency.

Campaigns were included in the analysis if they met all of the following criteria: (1) Public-facing scope: The campaign targeted citizens, SMEs, or a broad workforce audience, rather than specialised technical or sector specific groups. (2) National authority: Materials were produced or formally endorsed by a government body, national cyber security agency, or national CERT. (3) Accessibility: Campaign content was available in English or provided reliable official translations. (4) Temporal relevance: Materials were published or actively maintained between 2015 and 2025. (5) Explicit reference to AI-enabled threats: Campaigns explicitly mentioned artificial intelligence, deepfakes, voice synthesis, or AI-driven phishing or impersonation in their public messaging.

Campaign identification followed a systematic, multi-stage process inspired by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [28]. First, we conducted a comprehensive search of official national cyber security agency web sites, CERT pages, and archived campaign repositories for the period 2015–2025. This initial search yielded 38 candidate campaigns across multiple regions. In the screening stage, we excluded campaigns that did not satisfy the inclusion criteria; this removed 14 campaigns. During the

eligibility assessment, campaigns were further evaluated for explicit references to AI-enabled threats. Campaigns that did not explicitly address these risks were excluded, reducing the pool by an additional 13 campaigns.

The final included sample comprised 11 national campaigns spanning five global regions (USA, Europe, Canada, Oceania and Asia), representing a mix of high-income, middle-income, and digitally advanced nations. These campaigns provided publicly accessible materials (web pages, posters, videos, social media tool kits) and explicitly incorporated messaging about AI-driven threats.

2.2 Analysis approach

The analysis employed a qualitative thematic analysis of campaign messaging content. Materials were examined for recurring patterns in language, framing, and recommended actions, with particular attention to how campaigns conceptualised people's behaviour, responsibility, and capability in the context of AI-enabled threats. This cross-country comparison enabled identification of persistent design patterns as well as emerging shifts in awareness messaging. The campaigns and an initial meta-analysis are presented in Table 1.

Table 1: Public cyber security awareness campaigns and guidance incorporating AI and deepfake threats

Campaign / Initiative (short)	**Country / Agency**	**Year(s) with AI / deepfake mention**	**Type of AI threat integrated**
AI Fools Week / AI Fools: Stay Sharp! [24]	USA — National Cybersecurity Alliance (NCA)	2025	Dedicated one-week awareness toolkit focused on AI-enabled scams (voice cloning, deepfakes), includes tip sheets, posters, social assets.
Cybersecurity Awareness Month — "Secure Our World" (resources & tips) [12]	USA — CISA Cybersecurity Awareness Month 2025	2025	Materials & tip sheets explicitly call out AI-powered phishing, deepfakes and voice cloning as emergent threats.
European Cybersecurity Month (ECSM) — campaign theme and materials [17]	EU / ENISA (EU-wide)	2024–2025	ENISA promoted member-state materials on "scams, AI & deepfakes" as part of ECSM activities and created AI/deepfake resources.
NCSC – Assessments & public guidance on AI threats [23]	UK — NCSC	2024–2025	NCSC published assessments on the near-term impact of AI on the cyber threat and updated public guidance referencing AI-enabled threats.
Get Cyber Safe — New technology / AI guidance [6]	Canada — Get Cyber Safe / Canadian Centre for Cyber Security	2024–2025	Public blog/guidance pages describe how AI personalises phishing, automates fake accounts and recommend mitigations for AI-driven scams.
Own Your Online — Deepfake awareness pages [8]	New Zealand — CERT NZ / OwnYourOnline	2024–2025	Dedicated deepfake explainer page (what deepfakes are, risks, detection & reporting advice) aimed at public awareness.

Campaign / Initiative (short)	Country / Agency	Year(s) with AI / deepfake mention	Type of AI threat integrated
ACSC / Australian Gov — generative AI & content credentials guidance [2]	Australia — ACSC / Act Now. Stay Secure.	2024–2025	ACSC guidance on multimedia provenance and content credentials (joint info sheet) and campaign messaging referencing AI-driven scams.
BSI – Deepfakes guidance & public advice [18]	Germany — BSI / DsiN	2023–2024	BSI published guidance on deepfake dangers, detection and mitigation as part of national public advice.
CSA / Singapore resources & public survey [11]	Singapore — Cyber Security Agency (CSA)	2024–2025	CSA public materials and 2024/2025 survey results explicitly flag deepfakes/AI as an increasing risk and call for public awareness.
National cyber threat assessments referencing AI [7]	Canada — Cyber Centre (cyber.gc.ca)	2025	National threat assessment and guidance references AI-enabled attacks and points to public guidance (Get Cyber Safe).
ENISA / Member-state curated ECSM items (examples repeated across Europe) [17]	EU / Member states (e.g., UK/CH involvement)	2024–2025	ENISA/Member States curated materials & activities on social engineering including AI & deepfakes during ECSM events.

3 Findings

The analysis of 11 national-level cyber security awareness campaigns (2015–2025) reveals that while some initiatives are actively addressing AI-driven threats such as synthetic phishing and deepfakes, adoption remains uneven. For example, the U.S. Cybersecurity Awareness Month 2025 "Secure Our World" [12] materials explicitly acknowledge that "*threat actors increasingly exploit AI-driven phishing, social engineering, and deepfake-based impersonations*," integrating this content into weekly modules including an "*AI and Emerging Technology Security*" focus. Supporting tip sheets provide practical guidance for safe online AI usage [12]. Similarly, ENISA's European Cybersecurity Month activities incorporate AI-specific social engineering examples [17], while New Zealand's CERT NZ "*Own Your Online*" platform includes dedicated deepfake explainer pages [8].

The UK National Cyber Security Centre's phishing advice [23] continues to highlight indicators such as generic greetings and poor grammar, but only recently has it also begun to acknowledge that "*AI-generated messages that mimic real colleagues or services with alarming accuracy*" can be part of phishing attempts, reflecting an emerging but limited integration of AI threats into public guidance. Earlier campaigns, including those reviewed by Bada et al. [3], assumed that users could detect phishing through surface level cues and that vigilance alone could mitigate risk. Post-AI campaigns must contend with highly personalised, linguistically fluent, and contextually plausible messages generated by large language models, voice cloning, and synthetic media. While a subset of campaigns now explicitly reference AI-enabled scams, several continue to rely on

pre AI assumptions, widening the gap between awareness messaging and real-world threats. References to AI range from campaigns (e.g., NCA's AI Fools Week) [12] with dedicated AI tool kits to others (e.g., ENISA, CISA, NCSC) that integrate AI content into broader annual campaigns [12,17,23].

More broadly, our analysis identified six overarching themes relating to areas in campaigns that could be improved:

- **Theme 1: Overreliance on fear and risk amplification**. Many campaigns emphasise threats and catastrophic consequences of scams and online attacks, often through emotionally charged language. While such fear appeals may increase perceived threat, they are frequently not paired with clear, achievable actions that enhance users' sense of efficacy. Behavioural research suggests that fear without efficacy can lead to defensive avoidance or disengagement rather than sustained protective behaviour [3,20]. This suggests little progress since 2015 in replacing scare tactics with actionable guidance.
- **Theme 2: Knowledge focused content assuming rational actors**. Some campaigns still rely predominantly on information driven approaches, using checklists, infographics, and technical terminology. These approaches implicitly assume that individuals will rationally translate increased knowledge into secure behaviour [5]. These types of materials do well to enhance capability but may ignore motivation and opportunity, which are central to common behaviour change models such as COM-B [21].
- **Theme 3: Failure to address barriers**. Many campaigns instruct users to adopt specific practices such as enabling multi-factor authentication or updating software, without acknowledging constraints such as limited device control, organisational IT policies, usability issues, time pressure, or digital inequality. By failing to account for these barriers, campaigns risk reinforcing an intention-action gap, where users are aware of recommended behaviours but unable to perform them in practice [31,34].
- **Theme 4: Limited foundation in behavioural science mechanisms**. Only a minority of campaigns explicitly incorporate evidence based behaviour change techniques, such as timely prompts, social norm cues, implementation intentions, or strategies that support habit formation. Instead, many rely on broad urging to "be vigilant" or "stay alert," which place high cognitive demands on users. This limited engagement with behavioural science reflects patterns identified in prior work and suggests that insights from usable security and behaviour change research remain underutilised in campaign design [37,38].
- **Theme 5: Fatigue and desensitisation**. Several campaigns appear with increasing frequency across channels, particularly during annual awareness months or in response to emerging threats. While visibility may increase, repeated exposure to warnings may contribute to security fatigue [10]. Although aspects of fatigue were noted in earlier security research, this phenomenon was not widely recognised or systematically addressed in campaign

design in 2015. Our findings suggest that many campaigns still prioritise repetition over sustained engagement or behavioural reinforcement.

– **Theme 6: Failure to update messages for AI-assisted threats**. Pre-AI awareness campaigns implicitly assumed that phishing and social engineering could be detected through surface level cues such as poor grammar, generic greetings, or unusual formatting. In contrast, AI-enabled social engineering challenges this model by producing linguistically fluent, highly personalised, and contextually plausible messages at scale [19]. While a portion of recent national campaigns (e.g., NCA [12], ENISA-coordinated activities [17], and CERT NZ's [8] *Own Your Online*) acknowledge AI-generated threats, many others continue to promote legacy heuristics grounded in pre-AI assumptions. This misalignment between campaign messaging and attacker capabilities risks undermining credibility and may contribute to continued gaps between awareness and effective protective behaviour.

4 Discussion

Over the past decade, national cyber security awareness campaigns have made several improvements worthy of note. Campaigns now often feature more visual content, broader outreach channels, and attempts to employ positive framing or gamified elements [16]. There is greater recognition of human factors, security culture, and the limits of knowledge based approaches. However, fundamental issues still persist, such as: psychological and behavioural science grounding appears scarce, messaging continues to emphasise cognitive explanations of user behaviour rather than systemic and motivational factors, and attempts to evaluate and improve seem limited [3,33].

Several structural and institutional factors contribute to enduring limitations. For instance, there appears to be a preference for reuse of established templates and messaging formats, privileging visibility over effectiveness. Yet, there are no clear standardised metrics to evaluate behavioural change, leaving success measured by reach or engagement rather than meaningful adoption of protective practices [39]. Persistent reliance on awareness and vigilance as primary factors has several consequences. First, public confusion can arise from inconsistent or vague messaging across agencies and countries. Second, essential protective practices, such as MFA or independent verification, may see low adoption. Third, resources invested in campaigns may be wasted if intended behavioural outcomes are not achieved. Finally, users continue to bear disproportionate blame for breaches that reflect systemic or design shortcomings rather than individual negligence [5].

5 Implications for future campaigns: behavioural priorities for the next decade

The findings indicate that future campaigns must shift from awareness to behaviour support. Campaigns should explicitly embed behavioural frameworks

such as COM-B and the Behaviour Change Wheel [21] or MINDSPACE [14] to guide messaging design. Future campaigns should aim to shift responsibility solely from individuals, recognising that security is a sociotechnical challenge involving culture, technology, and incentives. Various parties are therefore involved including those who initiate the campaigns and those whose products and services are a core part of the public's online interactions.

Looking ahead to the next decade, the question is not whether awareness campaigns should address AI-enabled threats, but whether they can responsibly harness AI to support secure behaviour through design rather than persuasion alone. While AI has amplified social engineering threats, it also offers opportunities to improve cyber security awareness by enabling personalised, context aware, and behaviourally supportive interventions that reduce reliance on user vigilance alone [32]. Campaigns could use AI's adaptive potential to ethically support safe behaviour, tailoring guidance to user context, experience, and risk exposure. Messaging should also address digital fatigue, reframing awareness as enabling resilience rather than assigning blame. Behavioural research shows that interventions are more effective when they reduce cognitive load and align with existing workflows [21]. Campaigns should integrate with technical safeguards, such as automated warnings, protective defaults, and reporting mechanisms, thereby supporting secure behaviour through design rather than persuasion alone.

At the policy level, awareness campaigns should be considered as supportive rather than primary, limiting claims that user vigilance alone can prevent sophisticated attacks. Coordinated guidance is needed to ensure public messaging keeps pace with emerging AI threats, maintaining alignment between guidance and real-world attacker capabilities [12,16]. This would also help reduce the significant harms from these attacks [22]. Shifting responsibility from users to systems also has normative implications. By emphasising vigilance, campaigns implicitly position users as the final line of defence, reinforcing a culture of blame rather than resilience. A policy and design shift toward system level accountability would better reflect understandings of usable security, repositioning awareness as one component of layered defence rather than a substitute for robust system design.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Adams, A., Sasse, M.A.: Users are not the enemy. Communications of the ACM **42**(12), 40–46 (1999)
2. Australian Cyber Security Centre (ACSC): Content Credentials: Strengthening Multimedia Integrity in the Generative AI Era. https://www.cyber.gov.au/business-government/secure-design/artificial-intelligence/content-credentials-strengthening-multimedia-integrity-in-the-generative-ai-era (2024), accessed: 2025-12-12
3. Bada, M., Sasse, A.M., Nurse, J.R.C.: Cyber security awareness campaigns: Why do they fail to change behaviour? International Conference on Cyber Security for Sustainable Society (2015)

4. BBC News: What can I buy online at M&S since the hack? urlhttps://www.bbc.com/news/articles/c0el31nqnpvo (2025), accessed: 2025-12-23
5. Beautement, A., Sasse, M.A., Wonham, M.: The compliance budget: managing security behaviour in organisations. In: NSPW. pp. 47–58 (2008)
6. Canadian Centre for Cyber Security: New Technology - Get Cyber Safe. https://www.getcybersafe.gc.ca/en/blogs/new-technology (2024)
7. Canadian Centre for Cyber Security: National Cyber Threat Assessment 2025–2026. Tech. rep. (2025), https://www.cyber.gc.ca/en/guidance/national-cyber-threat-assessment-2025-2026
8. CERT NZ / Own Your Online: Deepfake Risks and Threats. https://www.ownyouronline.govt.nz/personal/know-the-risks/common-risks-and-threats/deepfake/ (2024), accessed: 2025-12-10
9. Chaudhary, S.: Driving behaviour change with cybersecurity awareness. Computers & Security **142**, 103858 (2024)
10. Cram, W.A., Proudfoot, J.G., D'Arcy, J.: When enough is enough: Investigating the antecedents and consequences of information security fatigue. Information Systems Journal **31**(4), 521–549 (2021)
11. Cyber Security Agency of Singapore (CSA): CSA's 2024 Cybersecurity Public Awareness Survey Shows an Improvement in the Adoption of Cyber Hygiene Practices. https://www.csa.gov.sg/news-events/press-releases/csa-s-2024-cybersecurity-public-awareness-survey-shows-an-improvement-in-the-adoption-of-cyber-hygiene-practices/ (2024), accessed: 2025-12-12
12. Cybersecurity and Infrastructure Security Agency (CISA): Secure Our World: Cybersecurity Awareness Month Resources. https://www.cisa.gov/resources-tools/resources/secure-our-world-cybersecurity-awareness-month-resources (2025), accessed: 2025-12-10
13. Delso-Vicente, A.T., Diaz-Marcos, L., Aguado-Tevar, O., de Blanes-Sebastián, M.G.: Factors influencing employee compliance with information security policies: a systematic literature review of behavioral and technological aspects in cybersecurity. Future Business Journal **11**(1), 28 (2025)
14. Dolan, P., Hallsworth, M., Halpern, D., King, D., Vlaev, I.: MINDSPACE: Influencing Behaviour Through Public Policy. Policy report, Institute for Government & UK Cabinet Office (2010), accessed: 2025-12-13
15. Egelman, S., Peer, E.: Scaling the security wall: Developing a security behavior intentions scale (sebis). In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. pp. 2873–2882 (2015)
16. European Union Agency for Cybersecurity (ENISA): European Cybersecurity Month 2022 Campaign Report. Technical report, European Union Agency for Cybersecurity (ENISA) (2023), https://www.enisa.europa.eu/publications/european-cybersecurity-month-2022-campaign-report, accessed: 2025-12-13
17. European Union Agency for Cybersecurity (ENISA): Promoting Security in the Digital World during the European Cybersecurity Month. https://www.enisa.europa.eu/news/promoting-security-in-the-digital-world-during-the-european-cybersecurity-month (2024), accessed: 2025-12-10
18. Federal Office for Information Security (BSI): Deepfakes – Information and Recommendations. https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/Deepfakes/deepfakes_node.html (2024), accessed: 2025-12-12
19. Jabir, R., Le, J., Nguyen, C.: Phishing attacks in the age of generative artificial intelligence: A systematic review of human factors. AI **6**(8) (2025)

20. Marett, K., Vedadi, A., Durcikova, A.: A quantitative textual analysis of three types of threat communication and subsequent maladaptive responses. Computers & Security (2019)
21. Michie, S., Van Stralen, M.M., West, R.: The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation science **6**(1), 42 (2011)
22. Mott, G., Turner, S., Nurse, J.R.C., Pattnaik, N., MacColl, J., Huesch, P., Sullivan, J.: 'There was a bit of PTSD every time I walked through the office door': Ransomware harms and the factors that influence the victim organization's experience. Journal of Cybersecurity **10**(1) (2024)
23. National Cyber Security Centre (NCSC): Impact of AI on the Cyber Threat. https://www.ncsc.gov.uk/report/impact-of-ai-on-cyber-threat (2024)
24. National Cybersecurity Alliance (NCA): AI Fools: Stay Sharp! https://www.staysafeonline.org/aifools (2025), accessed: 2025-12-13
25. NCSC: Impact of AI on cyber threat from now to 2027. urlhttps://www.ncsc.gov.uk/report/impact-ai-cyber-threat-now-2027 (2025)
26. Nurse, J.R.C.: Cybercrime and you: How criminals attack and the human factors that they seek to exploit. In: The Oxford Handbook of Cyberpsychology (2019)
27. Nurse, J.R.C., Williams, N., Collins, E., Panteli, N., Blythe, J., Koppelman, B.: Remote working pre-and post-covid-19: an analysis of new threats and risks to security and privacy. In: International Conference on Human-Computer Interaction. pp. 583–590. Springer (2021)
28. Page, M.J. et al.: Prisma 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ (2021)
29. Pattnaik, N., Nurse, J.R.C., Turner, S., Mott, G., MacColl, J., Huesch, P., Sullivan, J.: It's more than just money: the real-world harms from ransomware attacks. In: International Symposium on Human Aspects of Information Security and Assurance. pp. 261–274. Springer (2023)
30. Phipps, A., Nurse, J.R.C.: Inside ransomware groups: An analysis of their origins, structures, and dynamics. Computers & Security p. 104705 (2026)
31. Prümmer, J., van Steen, T., van den Berg, B.: A systematic review of current cybersecurity training methods. Computers & Security **136**, 103585 (2024)
32. Rana, S., Chicone, R.: AI-Driven Personalized Learning in Cybersecurity Training (2025)
33. Sasse, A.M., Brostoff, S., Weirich, D.: Transforming the "weakest link": A human/computer interaction approach to usable and effective security. BT Technology Journal **19**(3) (2016)
34. Sheeran, P.: Intention—behavior relations: A conceptual and empirical review. European Review of Social Psychology **12**(1), 1–36 (2002)
35. Silic, M., Silic, D., Kind-Trüller, K.: From shadow it to shadow ai–threats, risks and opportunities for organizations. Strategic Change (2025)
36. The World Economic Forum: Why humans are the weakest link, and other cybersecurity news (2025), https://www.weforum.org/stories/2025/07/hackers-employees-cybersupport-and-other-cybersecurity-news, accessed: 2025-12-23
37. Uchendu, B., Nurse, J.R.C., Bada, M., Furnell, S.: Developing a cyber security culture: Current practices and future needs. Computers & Security **109**, 102387 (2021)
38. Van Steen, T., Norris, E., Atha, K., Joinson, A.: What (if any) behaviour change techniques do government-led cybersecurity awareness campaigns use? Journal of Cybersecurity **6**(1) (2020)

39. Wash, R., Rader, E.: Too much knowledge? security beliefs and protective behaviors among united states internet users. In: SOUPs (2015)
40. Whitten, A., Tygar, J.D.: Why johnny can't encrypt: A usability evaluation of pgp 5.0. In: USENIX security symposium. vol. 348, pp. 169–184 (1999)