RAIS Journal for Social Sciences | Vol. 10, No. 2, 2026
ISSN 2574-0245 (Print) | ISSN 2574-1179 (Online) I DOI:10.5281/zenodo.24453640

Beyond the Weakest Link: A Critical Integrative Review of Human Factors in Cybersecurity, 2000-2025

Calvin Nobles

School of Cybersecurity and Information Technology, University of Maryland Global Campus, Maryland, United States, cn8972@gmail.com

Abstract: Background: The weakest-link metaphor treats security failure as an individual deficit and underweights the work systems that shape performance. **Purpose:** This review examined how human factors scholarship in cybersecurity developed from 2000 through July 2025, which mechanisms recur across evidence traditions, and which operational practices remain robust under evidence-role sensitivity testing. **Method:** The study conducted a bounded retrospective integrative review of 69 preserved evidence sources. The author reconciled every included source against the reference list, separated 40 primary empirical studies from 29 secondary, conceptual, methodological, and contextual sources, applied a design-sensitive appraisal, and used chronological, domain, tension, and practice coding. Four sensitivity analyses tested the findings using only primary empirical sources, only peer-reviewed sources, substantial primary evidence, and a corpus that excluded author-involved sources. The study treats retained search strings as provenance rather than as proof of an exhaustive historical search. **Results:** The corpus widened its unit of analysis from the individual at an interface to the employee within policy, the workforce within a sociotechnical system, and the analyst within a human-AI team. Five interacting domains and four recurring tensions organized the evidence. Control-work fit, behaviorally grounded role-specific learning, and layered social-engineering resilience retained strong convergence. Work-system risk and adaptive resilience retained moderate convergence. Cumulative demand and reporting systems remained moderate-emerging, and AI reliance calibration remained emerging. Excluding non-peer-reviewed and author-involved sources did not change the practice ratings. **Conclusion:** The Human Factors Cybersecurity Systems Framework links system levels, human performance mechanisms, practices, moderators, outcomes, and feedback. The framework offers an auditable explanatory structure and an evaluation agenda, but it does not constitute a validated predictive model or universal maturity scale.

Keywords: Human Factors, Cybersecurity, Human-Centered Cybersecurity, Usable Security, Security Fatigue, Sociotechnical Resilience, Human-AI Teaming, Integrative Review, Sensitivity Analysis

1. Introduction

Cybersecurity has traditionally concentrated on technical controls, vulnerability management, cryptography, identity and access management, network defense, and regulatory compliance. These elements remain indispensable, but security outcomes emerge through interactions among people, tasks, tools, policies, incentives, workloads, organizational routines, and adversarial manipulation. In this article, human factors refers to the cognitive, behavioral, affective, organizational, and interaction conditions that influence security performance. It does not function as a synonym for human error.

The familiar description of people as the weakest link compresses this system into an individual-deficit explanation. Sasse et al. (2001) showed that many apparent user failures originate in security mechanisms that conflict with cognition and work. Beautement et al. (2008) described a finite compliance budget, and Herley (2009) explained why users may rationally reject advice whose immediate cost appears greater than its expected benefit. These findings do not excuse negligent, reckless, or malicious conduct. They demonstrate that control design, work demands, and accountability jointly shape secure action.

Human factors research in cybersecurity subsequently developed across partially separated traditions. Usable-security research examined authentication, warnings, and interaction burden. Behavioral information-security research examined motivation, policy compliance, habit, fear, efficacy, and deterrence. Phishing research examined cue recognition, trust, persuasion, and susceptibility. Organizational and workforce research examined culture, fatigue, staffing, alert burden, reporting, and analyst burnout. More recent scholarship examined automation bias, skill erosion, explainability, and calibrated reliance in human-AI teams. Adjacent reviews clarified important parts of this landscape. Furnell and Clarke (2012) documented the growing recognition of human aspects in security. Desolda et al. (2021) synthesized human factors in phishing, Nifakos et al. (2021) focused on healthcare cybersecurity, Tilbury and Flowerday (2024) mapped human-automation relationships in security operations centers, and Khadka and Ullah (2025) proposed a contemporary interdisciplinary framework. These reviews differ in period, scope, method, and unit of analysis. Collectively, they do not provide a single evidence- and role-sensitive account of the developmental path from interface usability and employee behavior to workforce resilience and human-AI teaming across the 2000-2025 period.

Table 1. Positioning relative to adjacent reviews

Review	Scope	Method and evidence base	Unit of analysis	Distinct contribution or remaining gap
Furnell & Clarke (2012)	Human aspects of security; period not bounded	Narrative and conceptual review; source count not reported	End user	Established the importance of human considerations but did not map cross-period evidence or practice strength.
Desolda et al. (2021)	Phishing, 2001-May 2019	Systematic literature review; 52 studies	Phishing target	Deep synthesis of susceptibility and defense within one threat class.
Nifakos et al. (2021)	Healthcare cybersecurity through 2021	Systematic review; 70 sources	Clinical and healthcare work	Strong sector-specific account of behavior and organizational conditions.
Tilbury & Flowerday (2024)	Contemporary SOC automation	Scoping review; 49 sources	Security analyst and automation	Mapped augmentation and automation risks in security operations.
Khadka & Ullah (2025)	Interdisciplinary human factors, 2011-2024	Structured review; 41 sources	Behavior, culture, training, and ethical AI	Connected contemporary domains but did not trace the full 2000-2025 developmental arc.
Present review	Human factors in cybersecurity, 2000-July 2025	Bounded retrospective integrative review; 69-source audited corpus; evidence-role and sensitivity analyses	Interface, employee, organization, workforce, and human-AI team	Integrates chronology, domains, tensions, source roles, robustness tests, and eight source-traceable practices.

Note. The comparison identifies differences in review design and scope. It does not treat a broader scope as inherently superior to a specialized review.

The present article makes five linked contributions. First, it uses a bounded integrative-review identity that matches the heterogeneous corpus and limits inference to evidence that the surviving record can audit. Second, it reconciles every included source, reports a complete source-level matrix, and removes unverified historical selection counts from the analytic method. Third, it separates primary empirical evidence from secondary, conceptual, and

contextual contributions and tests the stability of practice ratings under four sensitivity analyses. Fourth, it identifies four recurring tensions that explain why apparently reasonable security interventions can fail. Fifth, it advances a multilevel Human Factors Cybersecurity Systems Framework that connects system conditions, performance mechanisms, operational practices, contextual moderators, outcomes, and feedback. The framework remains review-derived and provisional; it does not constitute a validated predictive model or maturity scale.

2. Review Purpose and Analytic Questions

The review sought to explain how human factors scholarship in cybersecurity developed from 2000 through July 2025, identify cross-level mechanisms and tensions, determine which operational practices exhibit defensible evidentiary convergence, and test whether those conclusions remain stable under stricter subsets of evidence. Four questions guided the synthesis:

RQ1: How did the unit of analysis and dominant human factors concerns evolve within the evidence corpus from 2000 through July 2025?

RQ2: Which cross-level mechanisms and tensions recur across usable-security, behavioral, organizational, workforce, and human-AI evidence traditions?

RQ3: Which operational practices receive strong, moderate, moderate-emerging, or emerging evidentiary convergence, and what boundary conditions constrain those practices?

RQ4: Do the practice ratings remain stable when the analysis uses primary empirical sources only, peer-reviewed sources only, substantial primary evidence, or a corpus that excludes author-involved sources?

3. Method

3.1 Review Design and Epistemic Scope

The study employed a bounded, retrospective, critical, integrative review. Integrative reviews can combine experimental, observational, qualitative, review, conceptual, and contextual sources when the purpose involves concept development, critical synthesis, and theory-building rather than pooled intervention-effect estimation (Snyder, 2019; Torraco, 2005, 2016; Whittemore & Knafl, 2005). The review followed the core logic of problem identification, literature provenance, data evaluation, data analysis, and presentation. The retrospective designation reflects the reanalysis of a preserved evidence corpus rather than the prospective execution of a newly registered systematic-review protocol.

The unit of inference is the preserved 69-source corpus. The study does not estimate the prevalence of themes across all cybersecurity scholarship, claim to have conducted an exhaustive retrieval, or infer pooled intervention effects. PRISMA 2020 and PRISMA-S informed the disclosure of available and unavailable search information, while selected SWiM principles informed transparent grouping and synthesis without meta-analysis (Campbell et al., 2020; Page et al., 2021; Rethlefsen et al., 2021). The article does not claim full compliance with guidelines designed for prospectively documented systematic reviews.

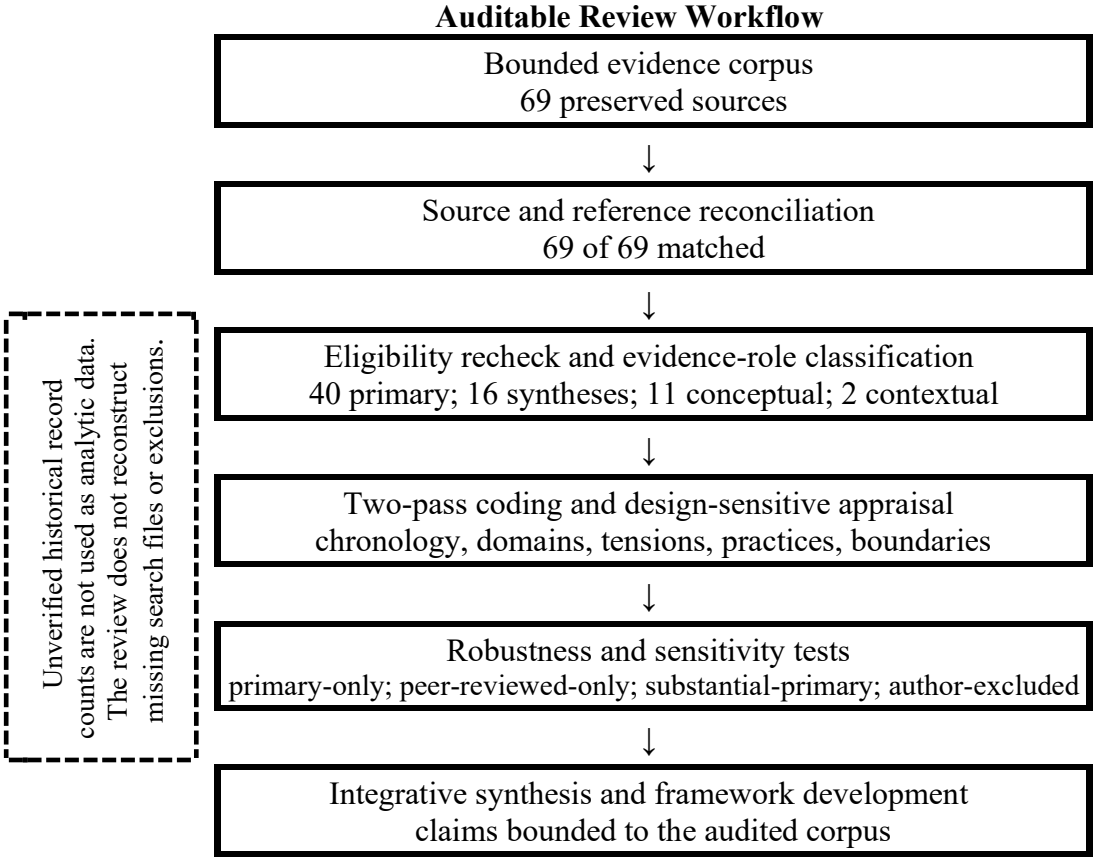
3.2 Corpus Provenance and Claim Boundaries

The complete evidence matrix and the matching reference list define the auditable corpus. The retained archive did not contain the record-level retrieval exports, the duplicate list, the excluded-title file, the database-specific execution records, or an independent reviewer file needed to verify a historical PRISMA selection flow. The revised method therefore begins with the 69 preserved included sources and does not use unverified aggregate retrieval or exclusion counts as analytic data. This decision removes false precision rather than reconstructing information that no longer exists.

The author reconciled each source against the reference list and checked author-year identity, publication period, design or source type, evidence role, publication status, weight category, domain coding, human factors focus, and practice linkage. Stable source identifiers

1 through 69 support source-level traceability. A deterministic reconciliation regenerated the period, evidence-role, weight, primary-domain, and practice counts directly from Appendix A. Figure 1 depicts only procedures that the preserved record can support.

Figure 1. Auditable included-corpus verification and synthesis workflow



Note. The workflow starts with the preserved included-source corpus. It does not present a reconstructed PRISMA identification flow.

3.3 Eligibility Criteria

The eligibility criteria bounded the corpus while allowing evidence from human-computer interaction, behavioral information security, organizational studies, workforce research, healthcare, cyber risk, and human-AI teaming. Table 2 states the criteria applied during the included-corpus eligibility recheck.

Table 2. Eligibility criteria

Decision	Criterion
Include	Publication dated from January 2000 through July 31, 2025 and available in English.
Include	Explicit relevance to cybersecurity, information security, phishing, social engineering, insider threat, usable security, security culture, cyber risk, cyber resilience, or security operations.
Include	Examination of a cognitive, behavioral, affective, organizational, workforce, interaction, usability, or sociotechnical factor.
Include	Empirical findings, a validated measure, a transparent evidence synthesis, a theoretical contribution, or a defensible practice implication.
Include	Peer-reviewed journal or conference publication, or a limited nonjournal source retained for a clearly stated foundational, empirical, or institutional reason.
Include	Sufficient methodological or conceptual detail for classification, weighting, and synthesis.
Include	A non-peer-reviewed source only when it supplied unique foundational, empirical, institutional, or scholarly value and its publication status remained explicit.

Decision	Criterion
Exclude	Exclusive focus on technical mechanisms without a human, behavioral, organizational, workforce, or usability dimension.
Exclude	Vendor marketing, blog commentary, opinion writing, or another source that could not support scholarly extraction.
Exclude	General privacy, computing, or information-technology discussion without explicit cybersecurity relevance.
Exclude	Duplicate record, non-English source, or insufficient information for source verification and coding.

3.4 Information Sources and Discovery Provenance

The retained review record named ten information sources. Table 3 groups them by function rather than treating publisher platforms, bibliographic indexes, and supplementary discovery as equivalent databases. The table describes the architecture represented in the preserved corpus; it does not certify exact historical execution, hit counts, or database-level contribution counts.

Table 3. Retained discovery architecture

Search function	Sources	Role in corpus development
Bibliographic, indexing, and disciplinary digital libraries	IEEE Xplore; ACM Digital Library; Scopus; Web of Science; PubMed	Structured disciplinary and cross-disciplinary discovery
Publisher-platform discovery	ScienceDirect; SpringerLink; Taylor & Francis Online; Wiley Online Library	Supplementary full-text and publisher-level discovery
Supplementary discovery	Google Scholar	Broad discovery, cited-by searching, and source verification
Citation tracing	Backward reference review and forward cited-by searching	Expansion from relevant anchor sources

The archive retained 18 conceptual Boolean strings spanning human factors, usable security, phishing, social engineering, awareness, behavior, fatigue, culture, cyber risk, insider threat, resilience, security operations, and human-AI teaming. Appendix B reproduces those strings as provenance. Appendix D converts the conceptual architecture into a prospective replication protocol that requires exact platform names, interfaces, execution dates, complete syntax, returned counts, exports, deduplication records, and reviewer decisions.

3.5 Data Extraction, Coding, and Reconciliation

The source-level extraction matrix recorded source identifier, author and year, publication period, design or source type, evidence role and publication status, weight category, primary and secondary domains, human factors focus, and practice linkage. Appendix C provides operational definitions and decision rules. The first analytic pass assigned one primary domain to each source so that chronological counts would not double-count multidisciplinary studies. The second pass retained secondary domains, identified mechanisms and cross-cutting tensions, recorded boundary or contrary evidence, and rechecked practice mappings.

The five domain codes were cognitive, behavioral, affective, organizational, and interaction and usability. The interaction and usability label identifies human-system interface conditions rather than technical controls. The final consistency audit checked all 69 rows, reconciled every evidence reference, and regenerated all descriptive counts from the matrix. One author completed the eligibility recheck, extraction audit, weighting, and coding, so the study does not report interrater agreement. Section 3.7 describes the safeguards used to expose dependence on subjective judgments.

3.6 Design-Sensitive Critical Appraisal

The corpus combines research designs and evidence functions that a single universal risk-of-bias score cannot evaluate on equivalent terms. The review therefore used a source-class appraisal crosswalk that combined Gough's review-specific weight-of-evidence logic with design-sensitive questions reflected in MMAT, ROBIS, and AMSTAR 2 (Gough, 2007; Hong et al., 2018; Shea et al., 2017; Whiting et al., 2016). The crosswalk did not convert unlike designs into a common numerical score. Instead, it judged whether each source could credibly perform the evidentiary role assigned to it.

Table 4. Source-class appraisal and evidentiary restrictions

Evidence class	Core appraisal questions	Permitted evidentiary function	Restriction
Primary empirical	Does the design fit the question? Are sampling or context, measures or data collection, analysis, and inferential limits sufficiently transparent?	Direct empirical contribution to mechanisms, conditions, outcomes, or practice claims.	A weak or indirect design cannot achieve a strong practice rating based on source count alone.
Secondary synthesis	Are the review question, search and selection approach, appraisal, synthesis, and limitations transparent and coherent?	Corroboration, field mapping, and identification of convergence or gaps.	A secondary review does not count as an independent primary confirmation and may overlap included primary studies.
Conceptual, theoretical, or methodological	Are constructs clear, reasoning coherent, propositions grounded, and implications bounded?	Theory development, construct clarification, methodological framing, and hypothesis generation.	Conceptual evidence cannot establish intervention effectiveness or causal certainty.
Contextual, gray, or practitioner	Is provenance identifiable, purpose transparent, and institutional or foundational relevance clear?	Historical, institutional, or practice context when peer-reviewed substitutes do not serve the same function.	Contextual evidence cannot independently elevate a practice above emerging support.

The combined judgment retained three categories. Substantial evidence indicated strong method-purpose coherence, high relevance, and a direct contribution to one or more review questions. Contributory evidence indicated adequate coherence and relevance but meaningful limits in design, sampling, transferability, directness, or generalizability. Contextual evidence supplied theory, historical influence, conceptual framing, or institutional context and did not function as direct causal support. The contextual label identifies evidentiary function rather than poor scholarly quality.

3.7 Bias Safeguards and Sensitivity Analyses

The review used six safeguards. First, evidence-role separation prevented secondary, conceptual, and contextual sources from functioning as interchangeable replications of primary studies. Second, the primary-only analysis repeated practice counts and rating judgments using the 40 primary empirical sources. Third, the peer-reviewed-only analysis removed the five sources that did not appear in peer-reviewed journals or confirmed peer-reviewed conference proceedings. Fourth, the author-involvement analysis removed two publications that included the present author. Fifth, the substantial-primary analysis examined whether each practice retained direct evidence among primary sources assigned substantial weight. Sixth, negative-case analysis retained contrary findings, boundary conditions, and contexts in which a practice could fail.

A practice rating remained robust when the peer-reviewed-only and author-excluded analyses preserved the rating and the primary-only analysis retained a defensible empirical base. When a practice depended mainly on secondary or conceptual evidence, the review retained or reduced the rating rather than allowing source volume to create an appearance of certainty. These analyses assess robustness within the bounded corpus. They do not estimate

publication bias statistically because the evidence lacks a common effect measure and comparable intervention outcome.

3.8 Synthesis and Evidentiary-Convergence Rules

The synthesis used four complementary lenses. Chronological synthesis examined the changing unit of analysis across five periods. Domain synthesis organized primary and secondary human factors codes. Critical synthesis identified recurrent tensions that connected otherwise separate literatures. Practice synthesis translated mechanisms and boundary conditions into eight operational practices. The review did not conduct a meta-analysis because the sources differed substantially in design, population, measures, outcomes, and theoretical frameworks.

Practice translation required a modifiable system or work target, an explicit human-performance mechanism, source-level traceability, documented boundary conditions, and a defensible primary empirical base for any rating above emerging. Strong convergent evidence required direct primary evidence across multiple designs and periods, corroboration from non-primary sources, and no major contradiction to the practice core. Moderate evidence indicated meaningful empirical convergence with narrower contexts, limited intervention testing, or an inferential step from mechanism to practice. Moderate-to-emerging evidence indicated a defensible organizational practice supported by several sources, but with little substantial direct intervention evidence. Emerging evidence indicated a small, recent, or predominantly conceptual and secondary evidence base. Ratings describe evidentiary convergence within the 69-source corpus; they do not estimate universal effect size or causal certainty.

4. Results

4.1 Included-Corpus Audit and Evidence Profile

The audit reconciled all 69 evidence rows with the evidence reference list and found complete values for source identity, period, design or source type, role and status, weight, domains, focus, and practice linkage. Forty sources were primary empirical studies, 16 were secondary syntheses, 11 were conceptual, theoretical, or methodological contributions, and 2 were contextual, gray, or practitioner sources. Sixty-four sources appeared in peer-reviewed journals or confirmed peer-reviewed conference proceedings. Five sources remained transparently labeled as non-peer-reviewed or nonjournal evidence. Two author-involved sources carried contextual weight and could not be used to determine a practice rating independently.

Table 5. Included-corpus audit and verification outcomes

Audit element	Result	Interpretation
Evidence-matrix and reference agreement	69 of 69 evidence rows matched	Every included source has a stable, traceable reference.
Required source-level fields	69 of 69 complete	All rows contain period, design, role/status, weight, domain, focus, and practice data.
Primary empirical evidence	40 of 69	Practice ratings could be tested without non-primary sources.
Peer-reviewed evidence	64 of 69	Five sources required explicit justification as foundational, empirical, institutional, or scholarly.
Author-involved evidence	2 of 69; both contextual weight	Removal tests possible self-citation dependence.
Cross-table count reconciliation	Complete	Period, role, weight, domain, and practice counts regenerate from Appendix A.

Table 6. Distribution of the 69 evidence sources

Dimension	Category	n	Share
Publication period	2000-2005	5	7.2%
Publication period	2006-2010	12	17.4%
Publication period	2011-2015	18	26.1%
Publication period	2016-2020	18	26.1%
Publication period	2021-July 2025	16	23.2%
Evidence role	Primary empirical	40	58.0%
Evidence role	Secondary synthesis	16	23.2%
Evidence role	Conceptual, theoretical, or methodological	11	15.9%
Evidence role	Contextual, gray, or practitioner	2	2.9%
Weight of evidence	Substantial	26	37.7%
Weight of evidence	Contributory	31	44.9%
Weight of evidence	Contextual	12	17.4%

The final period contains publications through July 31, 2025 rather than a complete five-year interval. The period count therefore cannot demonstrate a decline in scholarly activity. All temporal findings describe the composition and developmental logic of this corpus, not the prevalence of field-wide publications.

4.2 Sensitivity and Robustness Results

Table 7. Practice-level sensitivity analysis

Practice	Full corpus n	Primary n	Substantial primary n	Peer-reviewed n	Author-excluded n	Final convergence
P1	17	5	3	16	16	Strong convergent
P2	46	32	17	44	46	Strong convergent
P3	18	12	5	16	17	Moderate
P4	10	4	0	10	9	Moderate-emerging
P5	10	6	1	10	9	Moderate-emerging
P6	19	11	6	16	19	Strong convergent
P7	22	11	6	21	20	Moderate
P8	8	1	1	7	7	Emerging

The peer-reviewed-only and author-excluded analyses left every final rating unchanged. The primary-only analysis preserved a strong empirical base for P1, P2, and P6 and a moderate base for P3 and P7. P5 retained several primary sources but little substantial direct intervention evidence, so it remained moderate-emerging. P4 included four primary studies, but none received substantial weight; the stricter analysis therefore reduced P4 from moderate to moderate-emerging. P8 remained emerging because only one primary empirical study directly addressed skill erosion, while the broader claim relied heavily on human-automation theory, conceptual work, and secondary syntheses. These findings answer RQ4 and identify where additional intervention studies would have the greatest impact on the framework.

4.3 Developmental Trajectory

Within this corpus, the unit of analysis widened rather than simply moving from one topic to another. Early research concentrated on the person at an authentication or security interface. The 2006-2010 literature placed the employee inside policy, motivation, and mental models. The 2011-2020 literature expanded toward organizational conditions and the security

workforce. The 2021-July 2025 literature gave greater attention to organizational risk, human-AI teaming, automation, and expertise preservation. Earlier concerns did not disappear; interface design and behavior remained relevant as the field added higher system levels.

Table 8. Chronological trajectory and primary-domain profile

Period	n	Primary-domain profile	Dominant unit of analysis	Representative concerns
2000-2005	5	U 3; C 1; B 1	Individual at a security interface	Authentication usability, password restrictions, graphical authentication, social engineering, and behavior taxonomy
2006-2010	12	B 5; C 5; O 1; A 1	Employee within policy and organizational expectations	Compliance burden, efficacy, mental models, deterrence, habit, and automation complacency
2011-2015	18	B 8; O 4; C 3; A 2; U 1	Behavioral-security user and cybersecurity worker	Measurement instruments, phishing vulnerability, organizational barriers, and SOC analyst burnout
2016-2020	18	C 5; B 5; O 4; U 2; A 2	Sociotechnical system and security workforce	Security fatigue, cyber hygiene, organizational usability, compliance models, and SOC structures
2021-July 2025	16	O 7; C 6; A 2; U 1	Adaptive organization and human-AI team	Human cyber risk, AI-enabled phishing, trust calibration, skill erosion, and HFE capability

Note. C = cognitive; B = behavioral; A = affective; O = organizational; U = interaction and usability. Each source received one primary-domain code for counting and could receive secondary codes in Appendix A.

4.4 Five Interacting Human Factors Domains

The primary-domain recoding assigned 20 sources to cognitive factors, 19 to behavioral factors, 16 to organizational factors, 7 to affective factors, and 7 to interaction and usability. These counts support corpus description, not prevalence estimates for the broader field. The secondary codes in Appendix A show that most sources crossed domain boundaries.

Table 9. Human factors domains identified in the review

Domain	Primary n	Representative constructs	Cybersecurity mechanism	Representative sources
Cognitive	20	Attention, mental models, risk perception, cue recognition, cognitive bias, and judgment under uncertainty	Shapes recognition, interpretation, prioritization, and challenge	Canfield et al. (2016); Wash (2010); West (2008)
Behavioral	19	Clicking, reporting, passwords, workarounds, compliance, habit, misuse, and protective action	Translates beliefs, constraints, incentives, and routines into observable action	D'Arcy et al. (2009); Moody et al. (2018); Vance et al. (2012)
Affective	7	Fear, stress, fatigue, trust, confidence, frustration, and self-efficacy	Changes motivation, vigilance, disengagement, and reliance	Boss et al. (2015); Reeves et al. (2021); Stanton et al. (2016)
Organizational	16	Leadership, staffing, workload, policy clarity, culture, training quality, reporting, and learning	Creates conditions that enable or suppress secure action, escalation, and adaptation	Caputo et al. (2016); Sundaramurthy et al. (2015); Yeo & Banfield (2022)
Interaction and usability	7	Interface design, authentication burden, warnings, tool integration, access complexity, and recoverability	Determines whether controls fit human capabilities, work, and error recovery	De Angeli et al. (2005); Proctor et al. (2002); Sasse et al. (2001)

The domains function as an interacting system. A poorly designed warning can increase cognitive load, undermine confidence, encourage workaround behavior, and reveal

an organizational failure to test controls under realistic work conditions. Similarly, a punitive reporting process can transform an ordinary mistake into a delayed incident by adding affective threat and organizational friction. The evidence, therefore, resists interventions that isolate a single domain while leaving the surrounding system unchanged.

4.5 Four Cross-Cutting Tensions

Critical synthesis revealed four recurring tensions that connect the five domains. These tensions explain why security programs can increase formal control without improving realized security performance.

Table 10. Cross-cutting tensions and boundary conditions

Tension	Core problem	Boundary condition	Representative sources
Protection-friction	Controls can increase theoretical protection while consuming attention, memory, time, and task capacity.	Targeted friction can improve judgment; unnecessary or cumulative friction can lead to omissions and workarounds.	Sasse et al. (2001); Beautement et al. (2008); Herley (2009); Caputo et al. (2016)
Awareness-action	Knowledge and intention do not reliably become secure behavior because habit, efficacy, social influence, workload, and context intervene.	Learning must match role and task, provide efficacy, and measure transfer rather than completion alone.	Puhakainen & Siponen (2010); Johnston & Warkentin (2010); Moody et al. (2018)
Measurement-individualization	Organizations need measurable human cyber risk, but person-level scores can obscure work-system causes and introduce risks to validity, privacy, and fairness.	Use human factors measures to diagnose control conditions and work systems before labeling individuals.	Parsons et al. (2014, 2017); Evans et al. (2016); Huang et al. (2025)
Augmentation-deskilling	Automation can increase speed and coverage while also producing complacency, automation bias, reduced situation awareness, and skill erosion.	Benefits depend on task allocation, uncertainty cues, feedback, challenge pathways, and skill-maintenance design.	Parasuraman & Manzey (2010); Rinta-Kahila et al. (2023); Tilbury & Flowerday (2024)

The protection-friction and awareness-action tensions explain why stronger policies and more training do not automatically improve behavior. The measurement-individualization tension explains why risk quantification can reproduce the weakest-link problem under a more technical label. The augmentation-deskilling tension extends established human-automation theory into security operations and clarifies why AI governance must evaluate expertise and recovery, not only throughput.

4.6 Direct Answers to the Analytic Questions

RQ1 asked how the unit of analysis and dominant concerns evolved. The corpus shifted from a predominantly interface-centered focus to a focus on behavior in organizational contexts, workforce conditions, and human-AI collaboration. The movement represented an expansion of analytic scale rather than the replacement of earlier concerns.

RQ2 asked which mechanisms and tensions recurred across evidence traditions. Attention, mental models, efficacy, habit, cumulative demand, reporting and feedback, and trust calibration appeared repeatedly. Four tensions organized those mechanisms: protection-friction, awareness-action, measurement-individualization, and augmentation-deskilling.

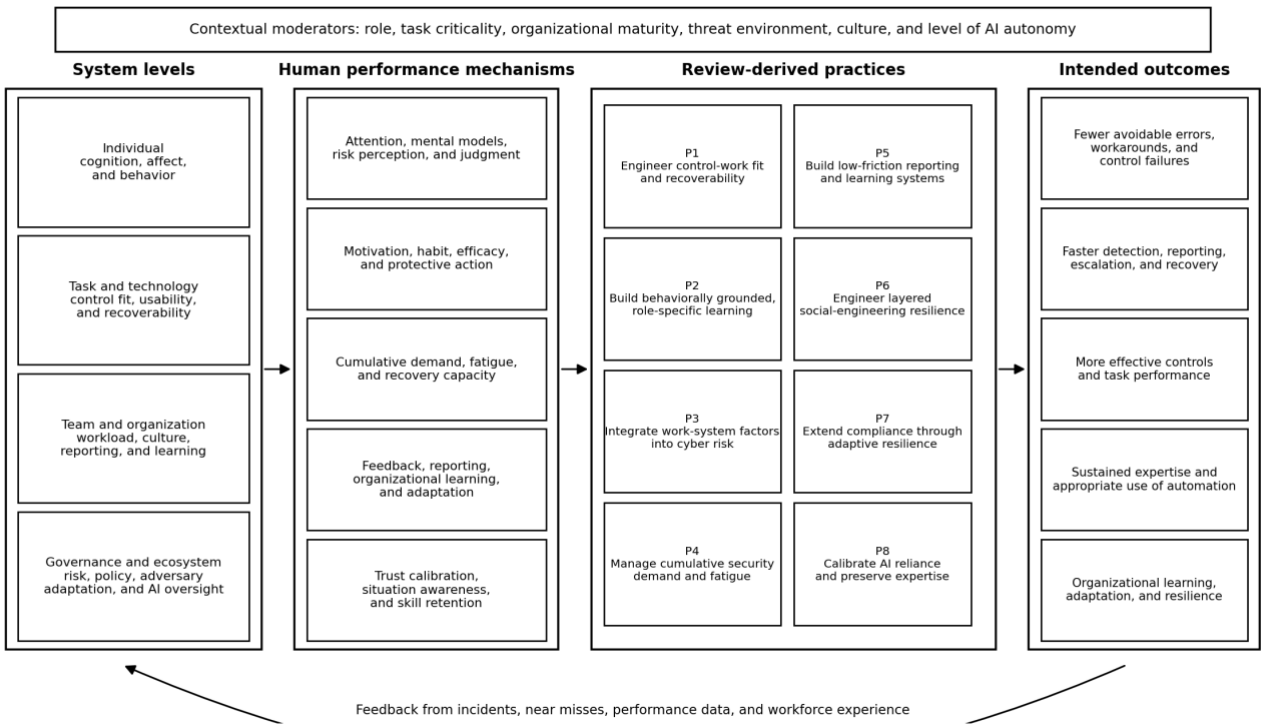
RQ3 asked which practices receive defensible convergence. Control-work fit, behaviorally grounded role-specific learning, and layered social-engineering resilience received strong convergent evidence. Integrating work-system factors into cyber risk and extending compliance through adaptive resilience received moderate evidence. Managing cumulative security demand and building low-friction reporting and learning systems received moderate-emerging evidence. AI reliance calibration and expertise preservation remained in the emerging stage.

RQ4 asked whether the practice ratings remained stable under stricter evidence subsets. Peer-reviewed-only and author-excluded analyses preserved all ratings. Primary-only and substantial-primary analyses confirmed the strongest practices but revealed greater uncertainty regarding cumulative demand, reporting systems, and the preservation of AI expertise. The review incorporated that uncertainty by downgrading P4 and retaining cautious ratings for P5 and P8.

5. Human Factors Cybersecurity Systems Framework

The synthesis supports a provisional explanatory framework rather than a list of isolated recommendations. The framework begins with four system levels: the individual; the task and technology; the team and organization; and governance and ecosystem conditions. Human performance mechanisms connect these levels to eight operational practices. Contextual moderators influence the proposed relationships, while feedback from incidents, near misses, performance data, and workforce experience supports adaptation. Intended outcomes include fewer avoidable errors and workarounds, faster reporting and recovery, more effective control performance, sustained expertise, and organizational learning.

Figure 2. Human Factors Cybersecurity Systems Framework



Note. The arrows represent review-derived and testable influence pathways. The framework is not an externally validated predictive model or maturity scale.

Table 11. Evidence traceability for the eight review-derived practices

Practice	Sources; primary	Direct and convergent evidence	Boundary or contrary evidence	Support
P1. Engineer control-work fit and recoverability	17; 5 primary	Authentication and graphical-interface experiments; organizational usability cases; cross-period usable-security evidence	No universal friction level; some friction is purposeful	Strong convergen t

Practice	Sources; primary	Direct and convergent evidence	Boundary or contrary evidence	Support
P2. Build behaviorally grounded, role-specific learning	46; 32 primary	Action research, experiments, surveys, instruments, field exercises, and behavior models	Knowledge and intention do not ensure action; evidence for algorithmic personalization remains limited	Strong convergent
P3. Integrate work-system factors into cyber risk	18; 12 primary	Validated awareness and intention instruments; profiling studies; organizational and risk research	No common enterprise standard; person-level scoring creates validity, privacy, and fairness risks	Moderate
P4. Manage cumulative security demand and fatigue	10; 4 primary	Authentication-burden studies; qualitative fatigue research; SOC workload and automation evidence	Intervention evidence remains thin; some security demand is necessary	Moderate-emerging
P5. Build low-friction reporting and learning systems	10; 6 primary	Reporting, adherence, organizational case, healthcare, and workforce evidence	Direct cybersecurity tests of psychological-safety interventions were absent	Moderate-emerging
P6. Engineer layered social-engineering resilience	19; 11 primary	Experiments, eye-tracking, field studies, surveys, reviews, and simulation evidence	Predictors vary by context; training alone does not remove susceptibility	Strong convergent
P7. Extend compliance through adaptive resilience	22; 11 primary	Compliance models, action research, field studies, organizational cases, and resilience frameworks	Resilience complements rather than replaces minimum controls and accountability	Moderate
P8. Calibrate AI reliance and preserve expertise	8; 1 primary	Automation theory, human-AI concepts, reviews, SOC evidence, and one empirical skill-erosion case	Automation can improve throughput; effects depend on task allocation, feedback, and calibration	Emerging

Table 12. Candidate operational indicators for framework implementation

Practice	Leading indicators	Intended performance outcome	Interpretive caution
P1	Task completion, comprehension, workaround frequency, support burden, recovery from error	Realized control performance and fewer avoidable failures	Do not interpret lower friction as automatically safer.
P2	Scenario performance, response efficacy, transfer to work, reporting behavior, retention over time	Observed behavior change rather than completion alone	Separate knowledge, intention, and behavior.
P3	Workload, staffing, policy comprehension, repeated workarounds, reporting latency, control usability	More diagnostic risk registers and control assessments	Avoid opaque employee risk scores.
P4	Alert volume and quality, authentication steps, omission, response time, disengagement, recovery	Preserved attention and lower fatigue-related failure	Use local baselines and intervention testing.
P5	Time to report, near-miss volume and quality, help seeking, closure time, climate perceptions	Faster escalation and organizational learning	Maintain proportionate accountability for reckless or malicious conduct.
P6	Verification-path use, compromise-to-report time, containment time, simulation behavior	Lower probability that one misjudgment becomes an uncontrolled incident	Do not treat simulated click rate as the sole outcome.

Practice	Leading indicators	Intended performance outcome	Interpretive caution
P7	Near-miss review, recurrence, corrective-action closure, recovery performance, adaptation quality	Learning, recovery, and reduced recurrence	Compliance remains a baseline, not the endpoint.
P8	Challenge and override quality, confidence calibration, manual diagnostic performance, recovery without automation	Appropriate reliance, preserved situation awareness, and sustained expertise	High throughput can conceal deskilling.

Note. These indicators translate the review into an evaluation agenda. They do not constitute a universally validated measurement instrument.

5.1 Control-Work Fit and Behaviorally Grounded Learning

P1 asks organizations to treat task analysis, workflow observation, usability testing, error analysis, and recovery design as security engineering. Controls should undergo evaluation under representative time pressure, device constraints, interruptions, and task switching. The design objective is not frictionless security. It is deliberate friction where it improves judgment, and minimal, unavoidable burden where it does not.

P2 replaces generic awareness with behaviorally grounded, role-specific learning. The evidence supports authentic scenarios, repeated exposure, rapid feedback, response efficacy, and measures of transfer. It supports tailoring by role, authority, exposure, and threat context more strongly than it supports automated personalization at the individual level. The practice therefore avoids claiming an adaptive capability that the corpus did not directly validate.

5.2 Work-System Risk, Cumulative Demand, and Reporting

P3 incorporates workload, staffing, policy comprehension, reporting latency, fatigue, repeated workarounds, and control usability into cyber-risk analysis. Validated instruments can improve measurement, but organizations should use them to diagnose working conditions and control performance rather than to produce opaque labels of personal defect.

P4 treats security demand as cumulative across applications, alerts, authentication steps, policy requirements, and competing operational priorities. Alert rationalization, phishing-resistant authentication, coherent access design, and clear escalation pathways can preserve attention for consequential decisions. Because intervention evidence remains limited, organizations should implement these changes with local baselines, prospective measures, and explicit counterfactuals.

P5 focuses on low-friction reporting and learning systems. Reporting should be fast, visible, easy to initiate, and explicitly valued by leaders. Accountability should distinguish ordinary error, at-risk conduct, reckless action, and malicious behavior. The evidence supports this system logic, but the corpus did not directly test psychological-safety interventions in cybersecurity. The revised practice name, therefore, states what the evidence can support without implying a stronger causal claim.

5.3 Layered Social-Engineering Resilience and Adaptive Compliance

P6 assumes that attackers exploit trust, authority, urgency, familiarity, identity, and context. Layered resilience combines verification pathways, phishing-resistant authentication, contextual warnings, realistic exercises, rapid reporting, and technical containment. No single training format removed susceptibility, and individual predictors varied across contexts. The appropriate design objective is to prevent a single moment of misjudgment from escalating into an uncontrolled incident.

P7 extends compliance through adaptive resilience. Compliance establishes minimum expectations and control baselines, but it cannot anticipate every interaction among people, technology, workflow, and adversaries. Resilient organizations examine near misses, recurring workarounds, reporting delays, and incident adaptations to redesign systems and

governance. This formulation avoids a false choice between compliance and resilience; the latter strengthens the former through learning and recovery.

5.4 Calibrated AI Reliance and Expertise Preservation

P8 remains emerging. AI-enabled security tools should expose confidence, uncertainty, provenance, and limitations in forms that support analyst judgment. Organizations should measure appropriate challenge and override, preserve manual and diagnostic practice, and test recovery when automated support becomes unavailable or wrong. Automation can improve speed and coverage. The design objective is calibrated reliance that preserves situation awareness and expertise, not generalized distrust of AI.

5.5 Framework Robustness and Intended Use

The peer-reviewed-only and author-excluded analyses preserved the framework's five domains, four tensions, eight practices, and overall pattern of ratings. The primary-only analysis also retained all eight practices, but it exposed weaker direct evidence for cumulative demand and AI-enabled expertise preservation. This result supports internal structural robustness and source traceability. It does not establish content validity through expert consensus, construct validity, predictive accuracy, implementation effectiveness, or transfer across sectors. Organizations should use the framework as a hypothesis map for measurement and intervention design rather than as a scoring instrument.

6. Discussion

6.1 Theoretical Contribution

The central developmental movement in the corpus is not a simple shift from the person to the system. It is a shift from single-level attribution toward multilevel causation. Individual judgment still matters, but the evidence increasingly locates judgment inside tasks, interfaces, teams, organizations, governance arrangements, and automated systems. This interpretation preserves individual agency and accountability while rejecting the claim that individual characteristics alone explain security performance.

The four tensions provide the review's primary theoretical synthesis. Protection-friction links usable security to workload and fatigue. Awareness-action connects behavior models to training transfer and organizational context. Measurement-individualization links cyber-risk quantification to validity, fairness, and system attribution. Augmentation-deskilling connects classic human-automation theory to emerging security operations. Together, these tensions explain why apparently stronger controls can reduce realized protection when they degrade comprehension, attention, reporting, adaptation, or expertise.

The Human Factors Cybersecurity Systems Framework converts those tensions into a provisional relational logic. System levels shape human performance mechanisms; practices target those mechanisms; contextual moderators alter likely effects; and feedback supports adaptation. This structure makes the framework more testable than a static catalog while avoiding a claim that the review established causal coefficients or predictive validity.

6.2 Methodological Contribution and Robustness

The review also demonstrates a defensible response to an incomplete historical archive. It does not reconstruct unavailable searches, excluded titles, or reviewer records. Instead, it narrows the unit of inference to a fully enumerated included corpus, assigns stable source identifiers, separates evidence roles, formalizes appraisal restrictions, regenerates every descriptive count from the source matrix, and reports sensitivity analyses that expose dependence on source class and author involvement. This approach cannot make the review exhaustive, but it makes the surviving synthesis auditable and prevents unsupported procedural claims.

The robustness results sharpen the substantive conclusions. Excluding non-peer-reviewed sources did not change any ratings, indicating that the main conclusions do not depend on the five transparently retained nonjournal or unconfirmed sources. Excluding the two author-involved publications also left every rating unchanged. The primary-only and substantial-primary tests identified a more consequential distinction: control-work fit, role-specific learning, and layered social-engineering resilience retain direct empirical depth, whereas cumulative demand and AI expertise preservation require more intervention and longitudinal evidence.

6.3 Evidence Strength and Boundary Conditions

Convergence ratings depended on directness, design, temporal spread, and inferential distance rather than source volume alone. P1 achieved strong convergence with five primary studies because direct usability evidence, organizational cases, and cross-period corroboration supported the control-work-fit mechanism. P4 moved to moderate-emerging because none of its four primary sources received substantial weight and intervention evidence remained thin. P5 remained moderate-emerging despite six primary sources because the specific reporting-climate intervention claim lacked direct testing. P7 remained moderate despite 11 primary sources, because the shift from compliance evidence to adaptive learning and recovery entailed a partly inferential step. P8 remained emerging because cyber-specific primary evidence was sparse.

The evidence also rejects universal prescriptions. Security friction can be useful, fear can motivate when efficacy is present, automation can improve throughput, and compliance can establish necessary baselines. The operational question is not whether an intervention is inherently human-centered. It is whether the intervention fits the task, population, threat, organizational maturity, and level of automation, and whether measured outcomes show improved realized security performance.

6.4 Implications for Leadership, Security Operations, and Education

Cybersecurity leaders should expand governance dashboards beyond training completion and phishing click rates. More diagnostic indicators include reporting latency, report quality, near-miss learning, workload, alert quality, policy comprehension, control usability, repeated workarounds, recovery performance, and appropriate challenge of automation. Governance should assign ownership of these system conditions rather than place all responsibility on end users. Security operations centers should integrate human factors into staffing, shift design, alert engineering, tool procurement, escalation, and incident review. Automation metrics should include situation awareness, override appropriateness, skill retention, and recovery when automated support fails, not only speed, coverage, and false-positive rates.

Cybersecurity education should include usable security, human factors engineering, cyberpsychology, organizational behavior, workforce well-being, social engineering resilience, and human-centered AI. Students should learn to analyze work and systems, test controls under realistic conditions, and evaluate how automation redistributes cognitive work. Workforce development should treat the preservation of expertise as an operational requirement as AI assistance expands.

6.5 Testable Propositions and Validation Roadmap

The framework supports falsifiable propositions rather than only broad recommendations. Table 13 identifies five propositions and research designs that would move the field from descriptive convergence toward intervention evidence. These propositions are review-derived hypotheses, not findings established by the present study.

Table 13. Review-derived propositions and recommended research designs

Proposition	Testable statement	Recommended design
1. Control-work fit	Control-work fit mediates the relationship between formal control strength and realized security performance; excessive unvalidated friction increases workarounds and error-recovery costs.	Field experiments or interrupted time-series studies that compare control variants under representative work conditions.
2. Learning transfer	Role-specific scenario learning improves observed secure behavior when it increases response efficacy and closely matches actual task conditions.	Randomized or stepped-wedge trials with observed behavior and delayed transfer measures.
3. Cumulative demand	Cumulative security demand predicts omissions and workarounds through fatigue and attentional depletion.	Longitudinal workload studies that combine telemetry, experience sampling, behavior, and recovery measures.
4. Reporting and learning	Low-friction, proportionate reporting systems reduce reporting latency and increase near-miss learning relative to cumbersome or punitive systems.	Cluster trials or natural experiments that measure latency, report quality, help seeking, and corrective-action closure.
5. AI calibration and skill	AI assistance improves throughput without degrading judgment only when calibration cues, challenge pathways, and skill-maintenance mechanisms preserve situation awareness.	Longitudinal human-AI experiments that vary uncertainty display, task allocation, feedback, and periods of manual operation.

Table 14. Prospective validation roadmap for the framework

Phase	Validation objective	Recommended method	Decision product
1. Content validity	Test whether system levels, mechanisms, practices, moderators, outcomes, and indicators adequately represent expert knowledge.	Multidisciplinary Delphi panel with cybersecurity practitioners, human factors specialists, behavioral scientists, educators, and governance leaders.	Revised construct definitions, content-validity indices, and documented disagreements.
2. Measurement development	Operationalize constructs and determine whether indicators are understandable, feasible, and nonredundant.	Cognitive interviews, expert review, pilot measurement, and item reduction across roles and sectors.	Candidate scales, indicator specifications, and scoring rules that remain separate from the conceptual framework until validated.
3. Construct and reliability testing	Assess internal structure, convergent and discriminant validity, temporal stability, and fairness.	Multi-organization field study with confirmatory factor analysis, reliability testing, measurement invariance, and subgroup analysis.	Validated measures and evidence on where comparisons are or are not legitimate.
4. Intervention and predictive testing	Determine whether practices improve security performance and whether relationships among frameworks predict outcomes.	Cluster trials, stepped-wedge designs, interrupted time series, and longitudinal human-AI experiments.	Effect estimates, boundary conditions, implementation costs, and predictive performance.
5. External replication	Test transfer across sectors, cultures, organizational maturity levels, and automation contexts.	Independent replications with preregistered protocols and harmonized core outcomes.	Evidence of generalizability, necessary adaptations, and conditions that falsify or constrain the framework.

Additional priorities include cross-cultural and underserved populations, ethically governed human cyber-risk measures, intervention trials for security fatigue, direct tests of reporting and learning climates, and comparative studies of AI-enabled deskilling across novice and expert analysts. A prospective update should register a protocol, preserve complete search exports, retain record-level decisions, use independent reviewers, apply design-specific appraisal tools, and archive the extraction matrix and analytic code.

7. Residual Scope Boundaries and Limitations

The first boundary concerns coverage. The review analyzes a bounded English-language corpus published from 2000 through July 31, 2025, and does not claim to be exhaustive in identifying the field. The revision resolves the original archive problem by removing unverified historical selection counts and limiting inference to the 69 preserved sources. It

does not recreate records that no longer exist, so relevant omitted scholarship remains possible. Appendix D provides the requirements for a prospectively documented update.

The second boundary concerns reviewer independence. One author completed the included-source eligibility recheck, coding audit, appraisal, and synthesis. The operational codebook, complete source matrix, deterministic count reconciliation, negative-case analysis, and four sensitivity analyses reduce inconsistency and expose where conclusions depend on evidence class. They do not substitute for independent human screening, coding, disagreement resolution, or interrater agreement. A prospective replication with at least two reviewers remains necessary before the study can claim independent verification.

The third boundary concerns evidence heterogeneity and dependency. Experiments, cross-sectional surveys, qualitative studies, organizational cases, instruments, reviews, conceptual contributions, and contextual sources answer different questions and support different inferences. The source-class appraisal and evidence-role restrictions prevent direct equivalence, while the primary-only analysis anchors practice ratings in empirical evidence. The review still cannot estimate pooled effect sizes or universal causal certainty. Some secondary syntheses may also include primary studies that are represented separately in the corpus; the analysis therefore treats reviews as corroboration rather than as independent primary replications.

The fourth boundary concerns publication and reporting bias. The absence of a common intervention effect and comparable outcomes prevents meaningful funnel plot or small-study effect analyses. Peer-reviewed-only, substantial-primary, and negative-case analyses reduce dependence on weak or selectively favorable evidence, but they cannot rule out unpublished null findings or selective outcome reporting.

The fifth boundary concerns framework validation. Source traceability and sensitivity analyses provide internal robustness, not external validity. The framework has not yet undergone multidisciplinary Delphi validation, psychometric testing, implementation trials, predictive evaluation, or independent replication. Organizations should use it to organize hypotheses, measures, and controlled interventions rather than as a universal score or certified maturity model.

8. Conclusion

Cybersecurity outcomes emerge from cognition, behavior, affect, organizational conditions, interaction design, workload, culture, staffing, and automation. The reviewed evidence does not support the view that the human factor can be located solely within an individual. It supports a multilevel account in which people, tasks, controls, teams, governance, and adversaries jointly determine whether performance is secure or insecure.

The strongest convergent evidence supports engineering control-work fit and recoverability, building behaviorally grounded role-specific learning, and deploying layered social-engineering resilience. Moderate evidence supports incorporating work-system factors into cyber risk and extending compliance through adaptive resilience. Managing cumulative security demand and building low-friction reporting and learning systems remain moderate-emerging, as direct evidence of intervention is limited. Calibrated reliance on AI while preserving expertise remains an urgent, emerging area of research.

The Human Factors Cybersecurity Systems Framework advances the field beyond the weakest-link metaphor by specifying system levels, performance mechanisms, practices, moderators, outcomes, and feedback. The revised review strengthens that contribution through a fully reconciled source matrix, explicit evidentiary restrictions, and sensitivity analyses that show which claims remain stable and which require caution. The framework's value lies not in providing a universal score, but in making human factors claims measurable, contestable, and testable within real cybersecurity work.

Funding

The author received no external funding for this study.

Competing Interests

The author declared no competing interests.

Data and Materials Availability

Appendix A provides the complete source-level evidence, role, weight, domain, and practice matrix. Appendix B provides the retained conceptual search strings and a provenance statement that distinguishes available from unavailable archival material. Appendix C provides the operational codebook, appraisal rules, and practice-rating criteria. Appendix D provides a prospective replication and validation protocol. All descriptive counts reported in the article were regenerated from Appendix A. The article does not claim access to database-level records or record-level exclusion files that were not preserved.

References

Aleroud, A., & Zhou, L. (2017). Phishing environments, techniques, and countermeasures: A survey. *Computers & Security, 68*, 160–196. https://doi.org/10.1016/j.cose.2017.04.006

Alsharnouby, M., Alaca, F., & Chiasson, S. (2015). Why phishing still works: User strategies for combating phishing attacks. *International Journal of Human-Computer Studies, 82*, 69–82. https://doi.org/10.1016/j.ijhcs.2015.05.005

Anwar, M., He, W., Ash, I., Yuan, X., Li, L., & Xu, L. (2017). Gender difference and employees' cybersecurity behaviors. *Computers in Human Behavior, 69*, 437–443. https://doi.org/10.1016/j.chb.2016.12.040

Bada, M., Sasse, A. M., & Nurse, J. R. C. (2015). Cyber security awareness campaigns: Why do they fail to change behaviour? In *International Conference on Cyber Security for Sustainable Society: Conference proceedings* (pp. 118–131).

Beautement, A., Sasse, M. A., & Wonham, M. (2008). The compliance budget: Managing security behavior in organizations. In *Proceedings of the 2008 New Security Paradigms Workshop* (pp. 47–58). Association for Computing Machinery. https://doi.org/10.1145/1595676.1595684

Blythe, J. M., Coventry, L., & Little, L. (2015). Unpacking security policy compliance: The motivators and barriers of employees' security behaviors. In L. F. Cranor, R. Biddle, & S. Consolvo (Eds.), *SOUPS 2015: Proceedings of the Eleventh Symposium on Usable Privacy and Security* (pp. 103–122). USENIX Association.

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., & Polak, P. (2015). What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly, 39*(4), 837–864. https://doi.org/10.25300/MISQ/2015/39.4.5

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly, 34*(3), 523–548. https://doi.org/10.2307/25750690

Cain, A. A., Edwards, M. E., & Still, J. D. (2018). An exploratory study of cyber hygiene behaviors and knowledge. *Journal of Information Security and Applications, 42*, 36–45. https://doi.org/10.1016/j.jisa.2018.08.002

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ, 368*, l6890. https://doi.org/10.1136/bmj.l6890

Canfield, C., Fischhoff, B., & Davis, A. (2016). Quantifying phishing susceptibility for detection and behavior decisions. *Human Factors, 58*(8), 1158–1172. https://doi.org/10.1177/0018720816665025

Caputo, D. D., Pfleeger, S. L., Sasse, M. A., Ammann, P., Offutt, J., & Deng, L. (2016). Barriers to usable security? Three organizational case studies. *IEEE Security & Privacy, 14*(5), 22–32. https://doi.org/10.1109/MSP.2016.95

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security, 32*, 90–101. https://doi.org/10.1016/j.cose.2012.09.010

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research, 20*(1), 79–98. https://doi.org/10.1287/isre.1070.0160

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies, 63*(1–2), 128–152. https://doi.org/10.1016/j.ijhcs.2005.04.020

Desolda, G., Ferro, L. S., Marrella, A., Catarci, T., & Costabile, M. F. (2021). Human factors in phishing attacks: A systematic literature review. *ACM Computing Surveys, 54*(8), Article 173. https://doi.org/10.1145/3469886

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 581–590). Association for Computing Machinery. https://doi.org/10.1145/1124772.1124861

Downs, J. S., Holbrook, M., & Cranor, L. F. (2006). Decision strategies and susceptibility to phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security* (pp. 79–90). Association for Computing Machinery. https://doi.org/10.1145/1143120.1143131

Egelman, S., & Peer, E. (2015). Scaling the security wall: Developing a security behavior intentions scale. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2873–2882). Association for Computing Machinery. https://doi.org/10.1145/2702123.2702249

Endsley, M. R. (2023). Supporting human-AI teams: Transparency, explainability, and situation awareness. *Computers in Human Behavior, 140*, Article 107574. https://doi.org/10.1016/j.chb.2022.107574

Evans, M., Maglaras, L. A., He, Y., & Janicke, H. (2016). Human behaviour as an aspect of cyber security assurance. *Security and Communication Networks, 9*(17), 4667–4679. https://doi.org/10.1002/sec.1657

Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security, 31*(8), 983–988. https://doi.org/10.1016/j.cose.2012.08.004

Furnell, S., Khern-am-nuai, W., Esmael, R., Yang, W., & Li, N. (2018). Enhancing security behavior by supporting the user. *Computers & Security, 75*, 1–9. https://doi.org/10.1016/j.cose.2018.01.016

Gomez, S. R., Mancuso, V., & Staheli, D. (2019). Considerations for human-machine teaming in cybersecurity. In D. D. Schmorrow & C. M. Fidopiastis (Eds.), *Augmented cognition: 13th International Conference, AC 2019* (pp. 153–168). Springer. https://doi.org/10.1007/978-3-030-22419-6_12

Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213–228. https://doi.org/10.1080/02671520701296189

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., & Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. *Computers & Security, 73*, 345–358. https://doi.org/10.1016/j.cose.2017.11.015

Greavu-Şerban, V., Constantin, F., & Necula, S.-C. (2025). Exploring heuristics and biases in cybersecurity: A factor analysis of social engineering vulnerabilities. *Systems, 13*(4), Article 280. https://doi.org/10.3390/systems13040280

Grobler, M., Gaire, R., & Nepal, S. (2021). User, usage and usability: Redefining human centric cyber security. *Frontiers in Big Data, 4*, Article 583723. https://doi.org/10.3389/fdata.2021.583723

Hadlington, L. (2017). Human factors in cybersecurity: Examining the link between internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviors. *Heliyon, 3*(7), Article e00346. https://doi.org/10.1016/j.heliyon.2017.e00346

Halevi, T., Memon, N., & Nov, O. (2015). Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2544742

Herley, C. (2009). So long, and no thanks for the externalities: The rational rejection of security advice by users. In *Proceedings of the New Security Paradigms Workshop* (pp. 133–144). Association for Computing Machinery. https://doi.org/10.1145/1719030.1719050

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221

Huang, W., Romanosky, S., & Uchill, J. (2025). Beyond technicalities: Assessing cyber risk by incorporating human factors (Research Report No. RR-A3841–1). RAND Corporation. https://doi.org/10.7249/RRA3841–1

Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security, 31*(1), 83–95. https://doi.org/10.1016/j.cose.2011.10.007

Iuga, C., Nurse, J. R. C., & Erola, A. (2016). Baiting the hook: Factors impacting susceptibility to phishing attacks. *Human-centric Computing and Information Sciences, 6*, Article 8. https://doi.org/10.1186/s13673-016-0065-2

Jabir, R., Le, J., & Nguyen, C. (2025). Phishing attacks in the age of generative artificial intelligence: A systematic review of human factors. *AI, 6*(8), Article 174. https://doi.org/10.3390/ai6080174

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS Quarterly, 34*(3), 549–566. https://doi.org/10.2307/25750691

Khadka, K., & Ullah, A. B. (2025). Human factors in cybersecurity: An interdisciplinary review and framework proposal. *International Journal of Information Security, 24*, Article 119. https://doi.org/10.1007/s10207–025–01032–0

Legg, P. A., Buckley, O., Goldsmith, M., & Creese, S. (2017). Automated insider threat detection system using user and role-based profile assessment. *IEEE Systems Journal, 11*(2), 503–512. https://doi.org/10.1109/JSYST.2015.2438442

McCormac, A., Zwaans, T., Parsons, K., Calic, D., Butavicius, M., & Pattinson, M. (2017). Individual differences and information security awareness. *Computers in Human Behavior, 69*, 151–156. https://doi.org/10.1016/j.chb.2016.11.065

Mitnick, K. D., & Simon, W. L. (2002). *The art of deception: Controlling the human element of security*. Wiley.

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a unified model of information security policy compliance. *MIS Quarterly, 42*(1), 285–311. https://doi.org/10.25300/MISQ/2018/13853

Mutlutürk, M., Wynn, M. G., & Metin, B. (2024). Phishing and the human factor: Insights from a bibliometric analysis. *Information, 15*(10), Article 643. https://doi.org/10.3390/info15100643

Nifakos, S., Chandramouli, K., Nikolaou, C. K., Papachristou, P., Koch, S., Panaousis, E., & Bonacina, S. (2021). Influence of human factors on cyber security within healthcare organisations: A systematic review. *Sensors, 21*(15), Article 5119. https://doi.org/10.3390/s21155119

Nobles, C. (2022). Stress, burnout, and security fatigue in cybersecurity: A human factors problem. *HOLISTICA – Journal of Business and Public Administration, 13*(1), 49–72. https://doi.org/10.2478/hjbpa-2022–0003

Nobles, C., Rangarajan, A., & Burrell, D. N. (2025). Human factors engineering-as-a-service in cybersecurity. In D. Burrell & C. Nguyen (Eds.), *New horizons in leadership: Inclusive explorations in health, technology, and education* (pp. 315–342). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-6437-6.ch014

Nurse, J. R. C., Creese, S., Goldsmith, M., & Lamberts, K. (2011). Guidelines for usable cybersecurity: Past and present. In *Proceedings of the Third International Workshop on Cyberspace Safety and Security* (pp. 21–26). IEEE. https://doi.org/10.1109/CSS.2011.6058566

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Parsons, K., Calic, D., Pattinson, M., Butavicius, M., McCormac, A., & Zwaans, T. (2017). The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies. *Computers & Security, 66*, 40–51. https://doi.org/10.1016/j.cose.2017.01.004

Parsons, K., McCormac, A., Butavicius, M., Pattinson, M., & Jerram, C. (2014). Determining employee awareness using the Human Aspects of Information Security Questionnaire. *Computers & Security, 42*, 165–176. https://doi.org/10.1016/j.cose.2013.12.003

Pfleeger, S. L., & Caputo, D. D. (2012). Leveraging behavioral science to mitigate cyber security risk. *Computers & Security, 31*(4), 597–611. https://doi.org/10.1016/j.cose.2011.12.010

Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F., & Guerri, D. (2022). Leveraging human factors in cybersecurity: An integrated methodological approach. *Cognition, Technology & Work, 24*(2), 371–390. https://doi.org/10.1007/s10111–021–00683-y

Proctor, R. W., Lien, M.-C., Vu, K.-P. L., Schultz, E. E., & Salvendy, G. (2002). Improving computer security for authentication of users: Influence of proactive password restrictions. *Behavior Research Methods, Instruments, & Computers, 34*(2), 163–169. https://doi.org/10.3758/BF03195438

Puhakainen, M., & Siponen, M. (2010). Improving employees' compliance through information systems security training: An action research study. *MIS Quarterly, 34*(4), 757–778. https://doi.org/10.2307/25750704

Reeves, A., Delfabbro, P., & Calic, D. (2021). Encouraging employee engagement with cybersecurity: How to tackle cyber fatigue. *SAGE Open, 11*(1), Article 21582440211000049. https://doi.org/10.1177/21582440211000049

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA Statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, Article 39. https://doi.org/10.1186/s13643–020–01542-z

Rinta-Kahila, T., Penttinen, E., Salovaara, A., Soliman, W., & Ruissalo, J. (2023). The vicious circles of skill erosion: A case study of cognitive automation. *Journal of the Association for Information Systems, 24*(5), 1378–1412. https://doi.org/10.17705/1jais.00829

Sasse, M. A., Brostoff, S., & Weirich, D. (2001). Transforming the weakest link: A human-computer interaction approach to usable and effective security. *BT Technology Journal, 19*(3), 122–131. https://doi.org/10.1023/A:1011902718709

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*, j4008. https://doi.org/10.1136/bmj.j4008

Siponen, M., & Vance, A. (2010). Neutralization: New insights into the problem of employee information systems security policy violations. *MIS Quarterly, 34*(3), 487–502. https://doi.org/10.2307/25750688

Siponen, M., Pahnila, S., & Mahmood, M. A. (2014). Employees' adherence to information security policies: An exploratory field study. *Information & Management, 51*(2), 217–224. https://doi.org/10.1016/j.im.2013.08.006

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stanton, B. C., Theofanos, M. F., Prettyman, S. S., & Furman, S. (2016). Security fatigue. *IT Professional, 18*(5), 26–32. https://doi.org/10.1109/MITP.2016.84

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers & Security, 24*(2), 124–133. https://doi.org/10.1016/j.cose.2004.07.001

Sundaramurthy, S. C., Bardas, A. G., Case, J., Ou, X., Wesch, M., McHugh, J., & Rajagopalan, S. R. (2015). A human capital model for mitigating security analyst burnout. In L. F. Cranor, R. Biddle, & S. Consolvo (Eds.), *SOUPS 2015: Proceedings of the Eleventh Symposium on Usable Privacy and Security* (pp. 347–359). USENIX Association.

Tilbury, J., & Flowerday, S. (2024). Humans and automation: Augmenting security operation centers. *Journal of Cybersecurity and Privacy, 4*(3), 388–409. https://doi.org/10.3390/jcp4030020

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review, 4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review, 15*(4), 404–428. https://doi.org/10.1177/1534484316671606

Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from habit and protection motivation theory. *Information & Management, 49*(3–4), 190–198. https://doi.org/10.1016/j.im.2012.04.002

Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security operations center: A systematic study and open challenges. *IEEE Access, 8*, 227756–227779. https://doi.org/10.1109/ACCESS.2020.3045514

Vishwanath, A. (2015). Habitual Facebook use and its impact on getting deceived on social media. *Journal of Computer-Mediated Communication, 20*(1), 83–98. https://doi.org/10.1111/jcc4.12100

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated information processing model. *Decision Support Systems, 51*(3), 576–586. https://doi.org/10.1016/j.dss.2011.03.002

Wash, R. (2010). Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/1837110.1837125

Wash, R., & Rader, E. (2011). Influencing mental models of security: A research agenda. In *Proceedings of the 2011 New Security Paradigms Workshop* (pp. 57–66). Association for Computing Machinery. https://doi.org/10.1145/2073276.2073283

West, R. (2008). The psychology of security. *Communications of the ACM, 51*(4), 34–40. https://doi.org/10.1145/1330311.1330320

Whiting, P., Savović, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., & Churchill, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology, 69*, 225–234. https://doi.org/10.1016/j.jclinepi.2015.06.005

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing, 52*(5), 546–553. https://doi.org/10.1111/j.1365–2648.2005.03621.x

Williams, E. J., Hinds, J., & Joinson, A. N. (2018). Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies, 120*, 1–13. https://doi.org/10.1016/j.ijhcs.2018.06.004

Yeo, L. H., & Banfield, J. (2022). Human factors in electronic health records cybersecurity breach: An exploratory analysis. *Perspectives in Health Information Management, 19*(2), Article 1i. https://pmc.ncbi.nlm.nih.gov/articles/PMC9123525/

Appendix A. Complete Evidence, Weight, Domain, and Practice Matrix

Table A1 lists all 69 evidence sources used in the synthesis. Evidence roles: PE = primary empirical; SS = secondary synthesis; CT = conceptual, theoretical, or methodological; CG = contextual, gray, or practitioner. Publication status: PJ = peer-reviewed journal; PC = peer-reviewed conference proceedings; CU = conference paper with peer-review status not independently confirmed; WP = working paper; PB = practitioner book; RR = institutional research report; BC = scholarly book chapter. Weight: S = substantial; C = contributory; X = contextual. Domains: C = cognitive; B = behavioral; A = affective; O = organizational; U = interaction and usability. The first domain code is primary; later codes are secondary.

Table A1. All 69 evidence sources with evidence role, weight, domain coding, and practice linkage

No.	Source	Period	Design/source type	Role/status	Weight	Domains	Human factors focus	Practice
1	De Angeli et al. (2005)	2000-2005	Controlled experiment	PE/PJ	S	U; C, B	Graphical authentication usability and memorability	P1
2	Mitnick and Simon (2002)	2000-2005	Foundational practitioner book	CG/PB	X	C; B, O	Social engineering as exploitation of trust and context	P6
3	Proctor et al. (2002)	2000-2005	Controlled experiment	PE/PJ	C	U; C, B	Password restrictions, recall, and authentication burden	P1, P4
4	Sasse et al. (2001)	2000-2005	HCI framework	CT/PJ	X	U; C, O	Weakest-link framing recast as a design failure	P1, P4
5	Stanton et al. (2005)	2000-2005	Survey and taxonomy	PE/PJ	C	B	Taxonomy of end-user security behavior	P2, P3
6	Beautement et al. (2008)	2006-2010	Conference paper	CT/PC	C	O; C, B, U	Finite compliance budget and security friction	P1, P4, P7
7	Bulgurcu et al. (2010)	2006-2010	Organizational survey	PE/PJ	S	B; C, A, O	Attitudes, awareness, rational beliefs, and policy compliance	P2, P7
8	D'Arcy et al. (2009)	2006-2010	Empirical study	PE/PJ	S	B; C, O	Deterrence, awareness, and information-systems misuse	P2, P7
9	Dhamija et al. (2006)	2006-2010	Controlled experiment	PE/PC	S	C; U, B	Visual and cognitive cues in phishing	P1, P6
10	Downs et al. (2006)	2006-2010	Empirical study	PE/PC	C	C; B	Decision strategies and phishing susceptibility	P2, P6
11	Herley (2009)	2006-2010	Conference paper	CT/PC	X	B; C, O, U	Rational rejection of costly security advice	P1, P4, P7
12	Johnston and Warkentin (2010)	2006-2010	Experimental survey	PE/PJ	S	A; C, B	Fear appeals, efficacy, and protective behavior	P2
13	Parasuraman and Manzey (2010)	2006-2010	Theoretical and empirical review	SS/PJ	C	C; A, U	Automation complacency and bias	P8
14	Puhakainen and Siponen (2010)	2006-2010	Action research	PE/PJ	S	B; C, O	Training design and policy compliance	P2
15	Siponen and Vance (2010)	2006-2010	Organizational survey	PE/PJ	S	B; C, O	Neutralization and policy violations	P2, P7
16	Wash (2010)	2006-2010	Empirical study	PE/PC	C	C; B	Folk models of home computer security	P2, P6
17	West (2008)	2006-2010	Conceptual review	SS/PJ	X	C	Cognitive biases in security decisions	P2, P6
18	Alsharnouby et al. (2015)	2011-2015	Eye-tracking study	PE/PJ	S	C; U, B	Phishing detection strategies and cue attention	P1, P2, P6
19	Bada et al. (2015)	2011-2015	Review and conference paper	SS/CU	X	B; C, O	Why awareness campaigns fail to change behavior	P2, P6
20	Blythe et al. (2015)	2011-2015	Empirical study	PE/PC	C	O; B, A, U	Motivators and barriers in policy compliance	P2, P5, P7
21	Boss et al. (2015)	2011-2015	Empirical study	PE/PJ	S	A; C, B	Fear appeals and protective behavior	P2
22	Crossler et al. (2013)	2011-2015	Literature review	SS/PJ	C	B; C, O	Behavioral information-security research agenda	P2, P3, P7
23	Egelman and Peer (2015)	2011-2015	Instrument development	PE/PC	S	B; C	Security behavior intentions scale	P2, P3

No.	Source	Period	Design/source type	Role/status	Weight	Domains	Human factors focus	Practice
24	Furnell and Clarke (2012)	2011-2015	Review article	SS/PJ	C	O; C, B, U	Recognition of human aspects in security	P1, P2, P7
25	Halevi et al. (2015)	2011-2015	Working paper and field study	PE/WP	X	A; C, B	Personality, self-efficacy, and spear-phishing vulnerability	P2, P6
26	Ifinedo (2012)	2011-2015	Empirical study	PE/PJ	S	B; C, A, O	Planned behavior and protection motivation in compliance	P2, P7
27	Nurse et al. (2011)	2011-2015	Conference paper	CT/PC	C	U; C, O	Guidelines for usable cybersecurity	P1, P2
28	Parsons et al. (2014)	2011-2015	Instrument development	PE/PJ	S	B; C, A	Human Aspects of Information Security Questionnaire	P2, P3
29	Pfleeger and Caputo (2012)	2011-2015	Applied review	SS/PJ	C	O; C, B	Behavioral science for cybersecurity risk	P1, P2, P3
30	Siponen et al. (2014)	2011-2015	Field study	PE/PJ	C	B; O, C	Employee adherence to information-security policies	P2, P5, P7
31	Sundaramurthy et al. (2015)	2011-2015	Ethnographic study	PE/PC	C	O; A, C, U	SOC analyst burnout and human capital	P4, P5, P7
32	Vance et al. (2012)	2011-2015	Empirical study	PE/PJ	S	B; C, A, O	Habit, protection motivation, and compliance	P2, P7
33	Vishwanath (2015)	2011-2015	Empirical study	PE/PJ	C	B; C	Habitual social-media use and deception	P2, P6
34	Vishwanath et al. (2011)	2011-2015	Empirical study	PE/PJ	S	C; B, A	Individual differences in phishing vulnerability	P2, P6
35	Wash and Rader (2011)	2011-2015	Research agenda	CT/PC	X	C; B	Security mental models and communication	P2, P6
36	Aleroud and Zhou (2017)	2016-2020	Survey	SS/PJ	C	C; B, U	Phishing techniques and countermeasures	P2, P6
37	Anwar et al. (2017)	2016-2020	Empirical study	PE/PJ	C	B; C, A	Gender differences in cybersecurity behavior	P2, P3
38	Cain et al. (2018)	2016-2020	Empirical study	PE/PJ	C	B; C	Cyber hygiene knowledge and behavior	P2, P3
39	Canfield et al. (2016)	2016-2020	Empirical study	PE/PJ	S	C; B	Quantification of phishing susceptibility	P2, P3, P6
40	Caputo et al. (2016)	2016-2020	Multi-organization case study	PE/PJ	C	O; U, C, B	Organizational barriers to usable security	P1, P5, P7
41	Evans et al. (2016)	2016-2020	Literature-based study	SS/PJ	C	O; B, C	Human behavior in cybersecurity assurance	P3, P7
42	Furnell et al. (2018)	2016-2020	Review and conceptual study	SS/PJ	C	U; B, C, O	User-centered support for secure behavior	P1, P2
43	Gomez et al. (2019)	2016-2020	Conceptual study	CT/PC	X	U; C, O	Human-machine teaming in cybersecurity	P8
44	Gratian et al. (2018)	2016-2020	Empirical study	PE/PJ	C	B; C, A	Human traits and cybersecurity intentions	P2, P3
45	Hadlington (2017)	2016-2020	Empirical study	PE/PJ	C	A; C, B	Impulsivity, internet addiction, and risky behavior	P2, P3
46	Iuga et al. (2016)	2016-2020	Empirical study	PE/PJ	C	C; B, A, U	Factors affecting phishing susceptibility	P2, P6
47	Legg et al. (2017)	2016-2020	Systems study	PE/PJ	C	O; B, C	User and role profiling for insider threat	P3
48	McCormac et al. (2017)	2016-2020	Empirical study	PE/PJ	C	C; B, A	Individual differences in information-security awareness	P2, P3
49	Moody et al. (2018)	2016-2020	Theory-building study	PE/PJ	S	B; C, A, O	Unified model of policy compliance	P2, P5, P7
50	Parsons et al. (2017)	2016-2020	Validation study	PE/PJ	S	B; C, A	Further validation of the HAIS-Q	P2, P3
51	Stanton et al. (2016)	2016-2020	Qualitative study	PE/PJ	C	A; C, B, O, U	Security fatigue as disengagement from security demands	P4
52	Vielberth et al. (2020)	2016-2020	Systematic study	SS/PJ	S	O; U, C, A	SOC structures, workload, role clarity, and tool integration	P1, P4, P5, P8

No.	Source	Period	Design/source type	Role/status	Weight	Domains	Human factors focus	Practice
53	Williams et al. (2018)	2016-2020	Empirical study	PE/PJ	S	C; B, A, O	Workplace phishing susceptibility	P2, P6
54	Desolda et al. (2021)	2021-2025	Systematic review	SS/PJ	S	C; B, A, U	Human factors in phishing	P2, P6
55	Endsley (2023)	2021-2025	Theoretical study	CT/PJ	X	C; U, O	Transparency, explainability, and situation awareness in human-AI teams	P8
56	Greavu-Șerban et al. (2025)	2021-2025	Empirical factor analysis	PE/PJ	S	C; B, A	Heuristics and biases in social-engineering vulnerability	P2, P3, P6
57	Grobler et al. (2021)	2021-2025	Conceptual study	CT/PJ	X	U; O, C, B	Users, usage, and usability in human-centric cybersecurity	P1, P7
58	Huang et al. (2025)	2021-2025	Research report	CG/RR	C	O; B, C, A, U	Human factors in enterprise cyber-risk assessment	P3
59	Jabir et al. (2025)	2021-2025	Systematic review	SS/PJ	S	C; B, U, O	Human factors and generative-AI-enabled phishing	P2, P6
60	Khadka and Ullah (2025)	2021-2025	Interdisciplinary review	SS/PJ	C	O; C, B, A, U	Human-centric framework spanning behavior, culture, and ethical AI	P2, P5, P7, P8
61	Mutluturk et al. (2024)	2021-2025	Bibliometric review	SS/PJ	S	C; B	Phishing and human-factor research trends	P2, P6
62	Nifakos et al. (2021)	2021-2025	Systematic review	SS/PJ	S	O; B, C, U, A	Human factors in healthcare cybersecurity	P1, P2, P5, P7
63	Nobles (2022)	2021-2025	Conceptual and applied study	CT/PJ	X	A; C, O	Stress, burnout, and security fatigue	P4, P5, P7
64	Nobles et al. (2025)	2021-2025	Applied conceptual chapter	CT/BC	X	O; U, C	Human factors engineering as an organizational service capability	P1, P3, P7, P8
65	Pollini et al. (2022)	2021-2025	Integrated methodological study	CT/PJ	C	O; U, C, B	Integrated human factors methodology for cybersecurity	P1, P3, P7
66	Reeves et al. (2021)	2021-2025	Empirical study	PE/PJ	C	A; B, O, C	Cyber fatigue and employee engagement	P2, P4
67	Rinta-Kahila et al. (2023)	2021-2025	Case study	PE/PJ	S	C; U, O	Skill erosion under cognitive automation	P8
68	Tilbury and Flowerday (2024)	2021-2025	Scoping review	SS/PJ	S	O; U, C, A	Automation bias and augmentation in SOCs	P4, P8
69	Yeo and Banfield (2022)	2021-2025	Exploratory study	PE/PJ	C	O; B, U, C	Human causes of electronic health record breaches	P2, P5, P7

Appendix B. Discovery Concepts and Archival Provenance

The appendix distinguishes retained historical information from unavailable information. The conceptual search strings are reproduced as retained. They are not presented as exact database-specific syntax, proof of original keystrokes, or a reproducible search log.

Table B1. Retained conceptual search strings

ID	Conceptual query
S1	"human factors" AND (cybersecurity OR "cyber security" OR "information security")
S2	"human-centered cybersecurity" OR "human-centred cybersecurity"
S3	"usable security" AND (cybersecurity OR "information security")
S4	phishing AND "human factors"
S5	"social engineering" AND "human factors"
S6	"security awareness" AND (cybersecurity OR "information security")
S7	"information security behavior" OR "information security behaviour"
S8	"security fatigue" AND (cybersecurity OR "information security" OR cyber)
S9	"cybersecurity culture" OR "information security culture"

ID	Conceptual query
S10	"cyber risk" AND "human factors"
S11	"human vulnerability" AND (cybersecurity OR "information security")
S12	"healthcare cybersecurity" AND "human factors"
S13	"security policy compliance" AND (behavior OR behaviour)
S14	"cyber resilience" AND "human factors"
S15	"insider threat" AND "human behavior"
S16	"security education training awareness" OR (SETA AND cybersecurity)
S17	"security operations center" AND (burnout OR workload OR analyst)
S18	("automation bias" OR complacency OR "human-AI teaming") AND ("security operations" OR SOC OR cybersecurity)

Table B2. Archival provenance and permitted use

Archival element	Status	Use in the present article
Complete included-source matrix	Available	Defines the 69-source analytic corpus and supports all descriptive counts.
Matching evidence reference list	Available	Supports source-level identity and traceability.
Conceptual search strings	Available	Reported as discovery provenance and as input to the prospective replication protocol.
Database-specific exact syntax, interfaces, and execution dates	Unavailable	Not reconstructed and not used to claim reproducibility.
Database hit counts, exports, and contribution counts	Unavailable	Removed from the analytic method and from any PRISMA flow claim.
Record-level duplicate and exclusion files	Unavailable	No record-level exclusion analysis or interrater screening claim is made.
Included-source codebook and appraisal rules	Available in Appendix C	Supports replication of the included-corpus analysis.
Prospective update protocol	Available in Appendix D	Specifies how a future review can eliminate the current archival gaps.

Appendix C. Operational Codebook and Decision Rules

Table C1. Source-level codebook

Variable	Operational definition	Decision rule
Source ID	Stable integer assigned to each included source.	IDs 1-69 remain fixed across tables and narrative traceability.
Publication period	One of five chronological bands: 2000-2005, 2006-2010, 2011-2015, 2016-2020, or 2021-July 2025.	Assign by publication year; the final band ends July 31, 2025.
Design/source type	The method or scholarly form that generated the source's contribution.	Use the source's actual design rather than its topic label.
Evidence role	PE, SS, CT, or CG according to whether the source supplies primary data, synthesis, conceptual/methodological reasoning, or context.	A source receives one primary evidence role.
Publication status	PJ, PC, CU, WP, PB, RR, or BC.	Report uncertain or nonjournal status explicitly; do not silently treat it as peer reviewed.
Weight	Review-specific judgment of method-purpose coherence, relevance, and directness.	S = substantial; C = contributory; X = contextual. Do not interpret X as poor quality when the source serves a contextual function.
Primary domain	The dominant human factors mechanism or level needed to explain the source's main contribution.	Assign one primary code for mutually exclusive counting.
Secondary domains	Additional cognitive, behavioral, affective, organizational, or interaction/usability mechanisms materially present.	Assign zero or more secondary codes; do not use them in primary-domain totals.

Variable	Operational definition	Decision rule
Human factors focus	Concise statement of the source's main human-performance contribution.	Describe the mechanism or condition, not merely the cybersecurity topic.
Practice linkage	One or more P1-P8 practices informed by the source.	Require a traceable mechanism or boundary relationship; topic similarity alone is insufficient.
Boundary or contrary evidence	Evidence that limits transfer, identifies tradeoffs, or contradicts an unrestricted practice claim.	Retain in synthesis and use it to calibrate the final convergence rating.

Table C2. Evidentiary-convergence rating rules

Rating	Minimum evidentiary logic	Required caution
Strong convergent	Direct primary evidence across multiple designs and periods; corroboration from other evidence roles; no major contradiction to the practice core.	Does not imply a universal effect size, identical implementation, or absence of boundary conditions.
Moderate	Meaningful empirical convergence, but contexts remain narrow, intervention testing is limited, or translation from mechanism to practice requires an inferential step.	Implementation requires local measurement and explicit evaluation.
Moderate-emerging	Several sources support the system logic, but substantial direct intervention evidence remains sparse or absent.	Treat as a testable organizational practice, not an established causal intervention.
Emerging	The evidence base is small, recent, or predominantly conceptual and secondary, with limited cyber-specific primary research.	Use for hypothesis generation and carefully controlled pilot testing only.

Appendix D. Prospective Replication and Validation Protocol

A future update can eliminate the present archival and independence limitations only through new prospective work. Table D1 specifies the minimum record that a replication team should create before claiming systematic or independently verified coverage.

Table D1. Minimum protocol for a prospectively auditable update

Step	Required action	Required retained artifact
1. Protocol	Define review type, questions, coverage dates, eligible evidence classes, planned analyses, and deviations before searching.	Time-stamped protocol and, where appropriate, public registration.
2. Search design	Translate each conceptual string into database-specific syntax and peer review the strategy.	Full strategy for every platform, interface, field code, limit, and date range.
3. Search execution	Run each search on a recorded date and save exactly what the platform returned.	Execution log, hit count, export file, and screenshot or system record for each source.
4. Deduplication	Import all records into named software and preserve automated and manual duplicate decisions.	Raw exports, deduplicated library, software/version record, and duplicate log.
5. Screening	Use at least two independent reviewers for title/abstract and full text, with piloted criteria.	Reviewer-specific decisions, reasons, conflict log, resolution record, and agreement statistic.
6. Extraction	Use a piloted extraction form and verify critical fields independently.	Versioned extraction matrix, data dictionary, and audit trail of changes.
7. Appraisal	Apply design-specific tools to empirical studies and appropriate tools to syntheses and textual evidence.	Item-level appraisal decisions, reviewer judgments, disagreements, and source-level summaries.
8. Synthesis	Predefine grouping, evidence-dependency safeguards, sensitivity analyses, and rules for translating evidence into practice claims.	Analytic code, decision rules, negative cases, and all intermediate tables.
9. Reporting and archive	Report deviations and deposit nonrestricted materials in a persistent repository.	PRISMA-S checklist where applicable, data and code repository, and complete record-level exclusion list.
10. Framework validation	Conduct expert content validation, measurement development, field testing, and independent replication.	Delphi records, instruments, psychometric results, intervention data, and replication reports.