

BEYOND AUTOMATION: GENERATIVE AI'S INFLUENCE ON HUMAN DECISION-MAKING, TRUST, AND ACCOUNTABILITY IN INTELLIGENT CYBERSECURITY SYSTEMS

Prof. Dr. Sajida Parveen^{*1}, Muhammad Zunair Khan², Muhammad Farman³, Nazia Azim⁴,
Anees Kumar Abbasi⁵

^{1*}Director, Quality Enhancement Cell, Karachi Institute of Economics and Technology, Karachi, Pakistan

²Karachi Institute of Economics and Technology, Karachi, Pakistan

³Department of Chemistry, University of Okara, Okara, 56300, Pakistan

⁴Lecturer department of computer science Abdul Wali Khan University Mardan

⁵Department of Computer Sciences & Information Technology, Women University of AJ&K, Bagh.

¹drsajidaparveen90@gmail.com, ²zunair@kiet.edu.pk, ³farmanrafique@uo.edu.pk, ⁴n.azim@awkum.edu.pk,
⁵aneesqumar@wuajk.edu.pk

DOI: <https://doi.org/10.5281/zenodo.23261482>

Keywords

Generative Artificial Intelligence;
Cybersecurity; Human Decision-
Making; Calibrated Trust;
Automation Bias; Accountability.

Article History

Received: 25 June 2026

Accepted: 16 September 2026

Published: 02 October 2026

Copyright @Author

Corresponding Author: *

Prof. Dr. Sajida Parveen

Abstract

Generative Artificial Intelligence is getting involved in more and more cybersecurity decision-making processes. However, the efficiency is just as important, but this needs to be balanced against issues including dependence, measurement of trust, and responsibility. The study examines how the use of GenAI-supported cybersecurity systems affects 412 cybersecurity professionals, IT security professionals, security operations managers, security managers, and related professionals from their perspectives of human decision-making, calibrated trust, automation bias, and transitioning to an accountable state. A structured questionnaire on five-point Likert scale was used for the study design which was cross sectional quantitative. The analysis of data was done by descriptive statistics, Cronbach's alpha, Pearson correlation and multiple linear regression. There was a strong relationship between GenAI reliance and perceived decision efficiency ($r = .52$, $p\text{-value} < .001$), and calibrated trust ($r = .47$, $p\text{-value} < .001$), and also showed a significant association with automation bias ($r = .39$, $p\text{-value} < .001$). At the multivariable level, reliance on GenAI was the greatest determinant of decision efficiency ($\beta = .41$, $p < .001$), explainability was the greatest among the determinants of calibrated trust ($\beta = .34$, $p < .001$), and human oversight was negatively associated with automation bias ($\beta = -.29$, $p < .001$). The strongest predictor of the diffusion of responsibility was indicated as responsibility ambiguity ($p < .001$, $\beta = .31$). When combined with explainable outputs, active human participation and well-defined accountability, the findings indicate that GenAI can enhance the cybersecurity decision processes. The findings confirm that the role of AI in the workplace must be focused on humans; it should never be a replacement for the essential cybersecurity decision-making.

INTRODUCTION

The rapid advancement of generative artificial intelligence (GenAI) is transforming cybersecurity operations. Large language models and related generative systems can process large volumes of security data, interpret threat intelligence, support vulnerability analysis, assist with incident investigations, generate response recommendations, and facilitate communication between technical systems and human analysts. Systematic reviews demonstrate that large language models are increasingly utilized across cybersecurity functions, including threat detection, malware analysis, vulnerability assessment, security operations, and defensive decision support (Hasanov et al., 2024). Recent evidence further indicates that their role is evolving beyond isolated analytical tasks toward more integrated cyber-defense environments, where AI systems participate directly in security assessment and response processes (Zhang et al., 2025).

This transition extends beyond conventional automation. Traditional security automation typically executes predefined rules or repetitive procedures, whereas GenAI systems are able to interpret unstructured information, generate context-sensitive responses, summarize complex evidence, and recommend courses of action. As a result, cybersecurity professionals are increasingly likely to interact with AI as a decision-support partner rather than merely as an automated tool. Recent research on Security Operations Centers (SOCs) reflects this shift toward human-AI collaboration, where AI assists with alert triage, investigation, and response, while human analysts retain supervisory and intervention responsibilities (Kumar et al., 2026). Likewise, contemporary frameworks for augmented cybersecurity operations emphasize differentiated roles for AI agents while maintaining human supervision over critical security functions (Yaich et al., 2025).

The increasing involvement of AI in decision-making processes raises important questions about the evolving role of human judgment. AI systems can rapidly process information that would

otherwise impose significant cognitive demands on security professionals, potentially enhancing the speed and efficiency of decision-making. Nevertheless, human expertise remains essential in situations characterized by uncertainty, contextual complexity, novel attack vectors, competing organizational priorities, and incomplete information. The concept of *suggestive AI* (artificial intelligence's *sis*) suggests that artificial intelligence arises not from replacing human decision-makers, but from integrating computational capabilities with human contextual reasoning and judgment (Jarrahi, 2018). This perspective is particularly pertinent to cybersecurity, where similar technical indicators may require different responses depending on operational context, organizational risk tolerance, and potential consequences.

Trust is central to effective human-AI collaboration. It determines whether users accept, verify, reject, or disregard automated recommendations. Lee and See (2004) argue that automation design should not aim to maximize trust, but to foster appropriate reliance, where users' trust aligns with the system's actual capabilities and limitations. Excessive trust can result in inappropriate reliance, while insufficient trust may cause users to overlook valuable technological assistance. Research on human trust in artificial intelligence similarly demonstrates that system reliability, transparency, technical performance, and representation influence the development of both cognitive and affective trust in AI systems (Glikson & Woolley, 2020). In cybersecurity contexts, where erroneous decisions can have significant operational consequences, calibrating trust is therefore more critical than trust alone.

A related issue is automation bias, which arises when users place excessive weight on automated recommendations and fail to adequately consider contradictory information or exercise independent judgment. This problem predates *showed AI*. Parasuraman and Riley (1997) showed that automation bias can occur when operators rely too heavily on technological systems, resulting in failures in monitoring and

decision-making. AI systems often communicate recommendations in fluent, confident, human-like language, which can make erroneous outputs seem credible. Experimental studies of AI-assisted decision-making indicate that users may accept incorrect AI recommendations even when they possess information that should prompt independent scrutiny. Notably, interventions that require users to engage cognitively with AI advice can mitigate such overreliance (Bućinca et al., 2021). Recent syntheses of human-AI collaboration research further identify automation bias as a significant challenge and emphasize the importance of active verification rather than passive acceptance of AI recommendations (Romeo & Conti, 2026).an important role in addressing this problem. When users can understand why an AI system generated a particular recommendation, they are better positioned to assess whether its reasoning is relevant, credible, and consistent with available evidence. Explainability can therefore influence trust, acceptance, and the quality of human engagement with AI-supported decisions. Shin (2021) found that explainability and perceived causability were associated with users' perceptions and trust toward AI systems. Nevertheless, explainability should not be interpreted as automatically preventing overreliance. Explanations that are excessively complex, misleading, or accepted without critical evaluation may create an illusion of understanding rather than genuine scrutiny. For intelligent cybersecurity systems, explainability must therefore support calibrated trust—a level of trust that allows professionals to benefit from AI recommendations while preserving the willingness and ability to question them.

Human oversight is integral to effective AI-assisted cybersecurity. Such environments can provide substantial operational support while preserving meaningful human control over critical decisions. Human oversight encompasses reviewing AI-generated evidence, validating recommendations against other security information, identifying contextual errors, intervening when automated outputs appear unreliable, and maintaining

authority over consequential actions. This distinction is essential because the value of technical automation does not require eliminating human judgment; emerging approaches to human-AI cybersecurity collaboration increasingly prioritize adaptive autonomy, calibrated trust, and supervisory structures rather than unrestricted delegation to AI systems (Kumar et al., 2026; Yaich et al., 2025).

Beyond decision quality and trust, GenAI introduces significant governance challenges related to accountability. Cybersecurity decisions increasingly result from interactions among analysts, managers, software developers, organizational policies, automated systems, and AI-generated recommendations. When an AI-supported security decision leads to an adverse outcome, assigning responsibility can be difficult. Novelli et al. (2024) conceptualize AI accountability as a sociotechnical relationship involving answerability, oversight, standards, processes, and consequences. They further emphasize the challenge of allocating responsibility when multiple human and technological actors contribute to an AI-mediated outcome. In cybersecurity, this ambiguity can lead to accountability diffusion, where responsibility is so widely distributed that no individual actor clearly assumes responsibility for the final decision.

This accountability issue matters because GenAI systems operate within organizational structures. Decisions regarding people and institutions make decisions selection, data access, deployment conditions, system permissions, validation procedures, escalation requirements, and human review are made by humans, attributing an adverse security decision solely to “the AI” can obscure the organizational processes that contributed to the outcome. Effective AI governance therefore requires clearly defined responsibilities for system developers, organizations, cybersecurity teams, and final decision-makers. Human oversight is thus essential not only for technical accuracy but also for maintaining identifiable responsibility for decisions made with AI assistance (Novelli et al., 2024).

Despite the rapid expansion of research on large language model (LLM) applications in cybersecurity and a substantial literature on trust, automation, explainability, and AI governance, these domains are often examined in isolation. Cybersecurity research typically focuses on technical performance and threat-detection capabilities, while human-AI research investigates trust and reliance, and governance research addresses accountability. These dimensions still need to be examined to integrate these dimensions within GenAI-assisted cybersecurity environments. Specifically, it is important to assess whether increased reliance on GenAI improves decision efficiency while also increasing automation bias, whether explainability fosters calibrated rather than unconditional trust, whether human oversight mitigates excessive reliance, and whether ambiguity in responsibility contributes to accountability diffusion. Accordingly, this examines how assisted cybersecurity systems influence decision-making, trust, automation bias, and perceived accountability. It specifically investigates the relationships among GenAI reliance, decision efficiency, calibrated trust, system explainability, human oversight, automation bias, responsibility ambiguity, and accountability diffusion. By integrating behavioral, technological, and governance dimensions into a single analytical framework, the study aims to move the discussion beyond whether cybersecurity processes should be automated. Instead, it focuses on identifying the conditions under which GenAI can enhance security decision-making while preserving critical human judgment, appropriately calibrated trust, meaningful oversight, and clearly defined accountability.

Literature Review

Integrating generative artificial intelligence into cybersecurity marks a shift from rule-based automation tools that can interpret and synthesize complex data to tools that assist decision-making. Large language models (LLMs), which underpin many contemporary GenAI applications, have demonstrated potential in vulnerability

assessment, malware analysis, threat intelligence, phishing detection, security monitoring, and incident response. Hasanov et al. (2024), through a systematic review of LLM applications in cybersecurity, identified a rapidly expanding range of defensive uses, while also highlighting limitations related to reliability, privacy, adversarial manipulation, and the generation of inaccurate information.

Recent evidence indicates that LLMs' cybersecurity implications extend beyond simple task automation. Zhang et al. (2025), in a review of over 300 studies across diverse cybersecurity scenarios, demonstrated that LLMs increasingly serve as adaptable components within security workflows rather than as isolated analytical tools. These systems can interpret natural-language security information, generate code, provide recommendations, and facilitate interaction between technical security infrastructure and human users. However, their probabilistic nature introduces risks such as hallucination, inconsistent outputs, adversarial prompting, and inappropriate reliance. This duality underscores the importance of human interaction with GenAI, as technical capability alone does not guarantee effective AI-assisted security decisions.

The expanding use of GenAI signals a shift from automation to augmentation. While conventional automation delegates predefined tasks to machines, augmented decision-making involves ongoing interaction between human professionals and intelligent systems. Jarrahi (2018) argues that AI and human decision-makers have complementary strengths: computational systems process large volumes of information rapidly, while humans provide contextual interpretation, intuition, ethical reasoning, and adaptability to ambiguity. This perspective is especially pertinent to cybersecurity, where analysts often face incomplete information, time constraints, uncertainty, and rapidly evolving threats. Therefore, GenAI's value lies not only in generating recommendations but also in how effectively those recommendations integrate with human professional judgment.

Decision efficiency is a primary rationale for adopting AI-assisted cybersecurity technologies. Security operations often require analysts to process substantial volumes of alerts, logs, vulnerabilities, indicators of compromise, and threat intelligence. GenAI has the potential to alleviate this cognitive burden by summarizing evidence, prioritizing information, explaining technical patterns, and generating response strategies. Reviews of cybersecurity-focused LLMs indicate that these capabilities can accelerate analysis and enhance access to complex technical information, especially when AI systems are integrated into broader security workflows (Hasanov et al., 2024; Zhang et al., 2025).

However, increased efficiency does not automatically translate to improved decision quality. The advantages of human-AI collaboration depend on users' ability to critically integrate AI-generated information rather than substituting machine recommendations for independent judgment. Jarrahi (2018) emphasized that AI augmentation is most effective when it combines technological and human strengths. This distinction is crucial in cybersecurity: while rapid recommendations may enhance operational efficiency, uncritical acceptance of incorrect recommendations can introduce new vulnerabilities. As a result, reliance on GenAI is likely to yield both beneficial and potentially adverse behavioral outcomes. Users decide whether to accept or reject AI-generated information. Glikson and Woolley (2020), in their review of empirical research on human trust in AI, demonstrated that trust is influenced by factors including system performance, reliability, transparency, representation, and the characteristics of the human user. Trust is therefore not simply an individual attitude; it emerges through interaction between technological capabilities and users' perceptions of those capabilities.

The literature increasingly differentiates between trust and calibrated trust. The goal of trustworthy human-AI interaction is not to maximize confidence in the system. Excessive trust can result in overreliance, while insufficient trust may lead

users to ignore valuable AI assistance. In their review of trust-in-AI literature, Henrique and Santos (2024) concluded that effective AI-supported decision-making requires balancing users' expectations with system capabilities. Their analysis identifies trust calibration as a critical research direction, as appropriate reliance depends on users' ability to discern when AI recommendations are useful and when further verification is necessary.

This distinction is particularly significant in cybersecurity. If an analyst distrusts a reliable AI system, valuable information may be overlooked, delaying threat detection or response. Conversely, excessive confidence in GenAI outputs can allow incorrect recommendations to influence operational decisions without sufficient scrutiny. Therefore, calibrated trust is a more appropriate objective than unconditional acceptance of intelligent cybersecurity systems.

Explainability is often proposed as a means to enhance human understanding of AI-assisted decisions. Rather than merely providing outputs, explainable systems offer information that clarifies how recommendations are generated and what evidence supports them. Shin (2021) found that explainability and causability significantly affect users' perceptions, trust, and acceptance of artificial intelligence. When users comprehend the reasoning behind an AI recommendation, they are better equipped to assess its appropriateness for the decision context.

In cybersecurity, explainability is especially important because recommendations may prompt consequential actions, such as isolating endpoints, blocking network traffic, escalating incidents, modifying access privileges, or identifying activities as malicious. An analyst who receives an interpretable explanation can compare the system's reasoning with other available evidence, rather than relying solely on the recommendation. However, explainability does not inherently ensure effective oversight. Bućinca et al. (2021) demonstrated experimentally that users may still over-rely on incorrect AI recommendations even when explanations are provided. Their study found that cognitive forcing interventions, which

require users to engage actively with recommendations, were more effective in reducing overreliance than simply presenting explanations. This finding suggests that explainability is beneficial only when it promotes meaningful evaluation rather than passive acceptance.

Consequently, explainability may support calibrated trust, but this relationship is not automatic. Its effectiveness depends on whether explanations enable users to critically assess AI recommendations. This distinction underscores the need to examine explainability alongside human oversight and automation bias, rather than as an isolated technological feature.

Automation bias is a significant behavioral risk in AI-assisted decision-making. It broadly refers to the tendency of users to place excessive weight on automated recommendations, potentially overlooking or inadequately evaluating contradictory evidence. With GenAI, this risk may be heightened because outputs are often presented in fluent and persuasive natural language.

Buçinca et al. (2021) found that individuals using AI-supported decision systems often accepted AI advice even when recommendations were incorrect. Their results indicate that reducing overreliance requires active cognitive engagement, rather than simply increasing the amount of information provided by the system.

Romeo and Conti (2026), in their review of automation bias in human-AI collaboration, also identify excessive reliance on algorithmic recommendations as a persistent challenge in high-stakes decision environments. They note that explainable AI can improve human-AI interaction, but explanations alone cannot eliminate automation bias. Effective human-AI collaboration design must encourage independent evaluation, verification, and appropriate human involvement.

These findings are particularly relevant to cybersecurity. Analysts experiencing alert fatigue, information overload, or time pressure may increasingly use GenAI recommendations as cognitive shortcuts. While such reliance can enhance efficiency when AI outputs are accurate,

it can also be problematic if the model generates inaccurate, incomplete, or misleading security recommendations. Automation bias thus serves as an important counterbalance to assertions that increased GenAI reliance inherently improves cybersecurity decision-making.

Human oversight is often proposed as a means to manage risks associated with AI-supported decision-making. However, simply assigning a person to be "in the loop" does not ensure meaningful supervision. Effective oversight requires that humans have sufficient information, expertise, authority, and opportunity to identify questionable outputs and intervene when necessary.

Langer et al. (2025) conceptualized effective human oversight as decision-makers' capacity to reliably detect inaccurate or unfair AI output; analysis indicates that oversight effectiveness depends on users' capacity to distinguish problematic recommendations from valid ones. Human supervision should thus be regarded as an active cognitive and organizational process, rather than merely the presence of a human operator.

Tsamados et al. (2025) further distinguish between supervisory control and human-AI teaming. They argue that more capable foundation models necessitate sophisticated forms of human control, as their behavior can be probabilistic, flexible, and difficult to predict. A collaborative approach emphasizes assigning appropriate roles to humans and machines, rather than viewing human intervention as a final procedural checkpoint.

This argument is directly relevant to cybersecurity. Security analysts must retain the ability to inspect, challenge, modify, or reject GenAI-generated recommendations, especially for high-impact decisions. Robust human oversight can help mitigate automation bias by promoting ongoing professional scrutiny. This theoretical relationship aligns with the present study's investigation of whether increased human oversight is associated with reduced automation bias. This also complicates responsibility allocation. Traditional accountability structures generally assume that identifiable human actors make decisions and can be required to explain or justify them. In AI-

mediated environments, however, an outcome may result from interactions among model developers, system providers, organizational managers, technical administrators, security analysts, organizational policies, and algorithmic recommendations.

Novelli et al. (2024) argue that AI accountability should be understood as a sociotechnical relationship involving answerability, standards, oversight, and consequences. Accountability cannot be meaningfully assigned unless relevant actors, obligations, and review mechanisms are clearly defined. Their framework demonstrates that AI governance requires not only technical transparency but also organizational structures that specify responsibility for particular decisions and outcomes.

This issue is particularly challenging in cybersecurity. When an AI recommendation leads to incorrect incident classification, inappropriate access restriction, delayed response, or other adverse security outcomes, responsibility may be attributed to the analyst, the deploying organization, or the technology developers. If roles are not clearly defined, this distributed involvement can result in responsibility ambiguity and ultimately accountability diffusion.

Human oversight can help address this problem when decision authority is clearly allocated. However, oversight without meaningful control may create only the appearance of accountability. Tsamados et al. (2025) emphasize that effective human control requires appropriate interaction between system and human capabilities, highlighting the need to distinguish substantive oversight from merely symbolic human involvement.

The reviewed literature demonstrates that GenAI offers significant potential to support cybersecurity through rapid information processing, threat analysis, interpretation, and decision assistance. However, research on human-AI interaction shows that technological performance alone does not determine the effectiveness of intelligent decision-support systems. Increased reliance may enhance efficiency but also foster overreliance; trust can facilitate appropriate system use but may

become problematic if it exceeds system capability; explainability may improve understanding without necessarily eliminating automation bias; and human oversight can reduce AI-related risks only when it enables genuine independent evaluation and intervention.

A key limitation of the existing literature is that these factors are often examined within separate research traditions. Cybersecurity studies primarily address technical applications and security performance (Hasanov et al., 2024; Zhang et al., 2025), while human-AI studies focus on trust, explanation, and reliance (Bućinca et al., 2021; Glikson & Woolley, 2020; Shin, 2021). Other scholarship addresses oversight and AI governance (Langer et al., 2025; Novelli et al., 2024; Tsamados et al., 2025). There is relatively little empirical research that integrates these dimensions within a single framework specifically focused on GenAI-assisted cybersecurity decision-making.

The present study addresses this gap by examining GenAI reliance, decision efficiency, calibrated trust, system explainability, human oversight, automation bias, responsibility ambiguity, and accountability diffusion within a unified analytical framework. This integrated perspective facilitates assessment of both the benefits and behavioral risks of GenAI-supported cybersecurity, while also considering the organizational mechanisms necessary to preserve professional judgment and accountability.

Methodology

Research Design

A quantitative cross-sectional design was employed to investigate the influence of generative artificial intelligence on human decision-making, trust, automation bias, and accountability within intelligent cybersecurity systems. The research examined the relationships among GenAI reliance, decision efficiency, calibrated trust, system explainability, human oversight, automation bias, responsibility ambiguity, and accountability diffusion.

Study Population and Sampling

The study population comprised individuals with professional or practical experience in cybersecurity, information technology, security operations, incident response, digital risk management, or AI-supported decision environments. Eligibility criteria required participants to be at least 18 years old and to have prior exposure to generative AI or AI-assisted decision-support tools.

A purposive sampling technique was employed to recruit participants via professional networks, information technology and cybersecurity communities, institutional contacts, and online platforms. Following eligibility and completeness screening, 412 valid responses were retained for analysis. The sample consisted of cybersecurity analysts, IT professionals, security operations and incident-response personnel, managers or team leaders, and individuals from academic or research-related roles.

Data Collection Instrument

Data collection was conducted using a structured, self-administered questionnaire. The instrument included items on demographic and professional characteristics, as well as multi-item measures assessing the principal study constructs. Responses were recorded on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

GenAI reliance assessed the extent to which participants incorporated AI-generated recommendations into cybersecurity decisions. Decision efficiency measured perceived improvements in the speed and effectiveness of decision-making. Calibrated trust evaluated appropriate reliance on GenAI while maintaining the ability to verify or reject system recommendations. Explainability measured the clarity and interpretability of AI-generated outputs.

Human oversight measured the extent to which professionals retained supervisory control and final decision authority. Automation bias captured the tendency to accept AI-generated recommendations without sufficient independent verification. Responsibility ambiguity assessed

uncertainty regarding responsibility for AI-assisted decisions. Accountability diffusion measured the extent to which responsibility was perceived as distributed across multiple human and technological actors.

Professional information collected included sex, age group, occupational role, years of experience, and frequency of GenAI use.

Reliability and Data Preparation

Questionnaire items were reviewed for clarity and relevance to cybersecurity and AI-assisted decision-making. Internal consistency was evaluated using Cronbach's alpha, with values of .70 or higher considered acceptable. In the final analysis, reliability coefficients ranged from .81 to .88. Composite scores for each construct were calculated by averaging the relevant items, with higher scores indicating greater levels of the corresponding variable.

Statistical Analysis

Data analysis included both descriptive and inferential statistical procedures. Frequencies and percentages were calculated for categorical variables, while means and standard deviations were reported for continuous study. Pearson correlation analysis was conducted to examine bivariate relationships among GenAI reliance, decision efficiency, calibrated trust, explainability, human oversight, automation bias, responsibility ambiguity, and accountability diffusion.

Four multiple linear regression models were estimated. Decision efficiency was modeled using GenAI reliance, explainability, human oversight, professional experience, and automation bias as predictors. Calibrated trust was modeled with explainability, GenAI reliance, human oversight, responsibility ambiguity, and automation bias as predictors. Automation bias was examined in relation to GenAI reliance, human oversight, responsibility ambiguity, professional experience, and explainability. Accountability diffusion was modeled using responsibility ambiguity, automation bias, human oversight, explainability, and GenAI reliance as predictors.

Standardized regression coefficients (β), 95% confidence intervals, p-values, R^2 , and adjusted R^2 were reported. Statistical significance was defined as $p < .05$ using two-tailed tests. Standard regression assumptions, including linearity, homoscedasticity, normality of residuals, independence, and multicollinearity, were assessed prior to interpretation of the final models.

Ethical Considerations

Participation was voluntary, and informed consent was obtained prior to completion of the questionnaire. Participant confidentiality was maintained, and results were reported only in aggregated form. Ethical approval was obtained from the relevant institutional authority before data collection.

Results

Participant Characteristics

The analysis included 412 respondents from cybersecurity, information technology, security operations, managerial, and research-related roles. Most participants were aged 21 to 40 years, accounting for 72.9% of the sample. Cybersecurity analysts represented the largest professional group (30.6%), followed by information technology professionals (26.2%) and security operations or incident-response personnel (19.2%). In terms of professional experience, 31.8% of respondents had three to five years of experience, while 42.7% reported more than five years. GenAI use was prevalent, with 39.6% of participants reporting daily use and an additional 34.5% using GenAI-enabled tools several times per week.

Table 1. Demographic and Professional Characteristics of Participants (N = 412)

Characteristic	Category	n	%
Sex	Male	276	67.0
	Female	126	30.6
	Prefer not to say	10	2.4
Age group	21–30 years	144	35.0
	31–40 years	156	37.9
	41–50 years	78	18.9
	>50 years	34	8.3
Professional role	Cybersecurity analyst	126	30.6
	IT professional	108	26.2
	SOC/incident-response staff	79	19.2

	Manager/team lead	55	13.3
	Academic/research/other	44	10.7
Professional experience	<3 years	105	25.5
	3–5 years	131	31.8
	6–10 years	107	26.0
	>10 years	69	16.7
Frequency of GenAI use	Daily	163	39.6
	Several times per week	142	34.5
	Weekly or less	107	26.0

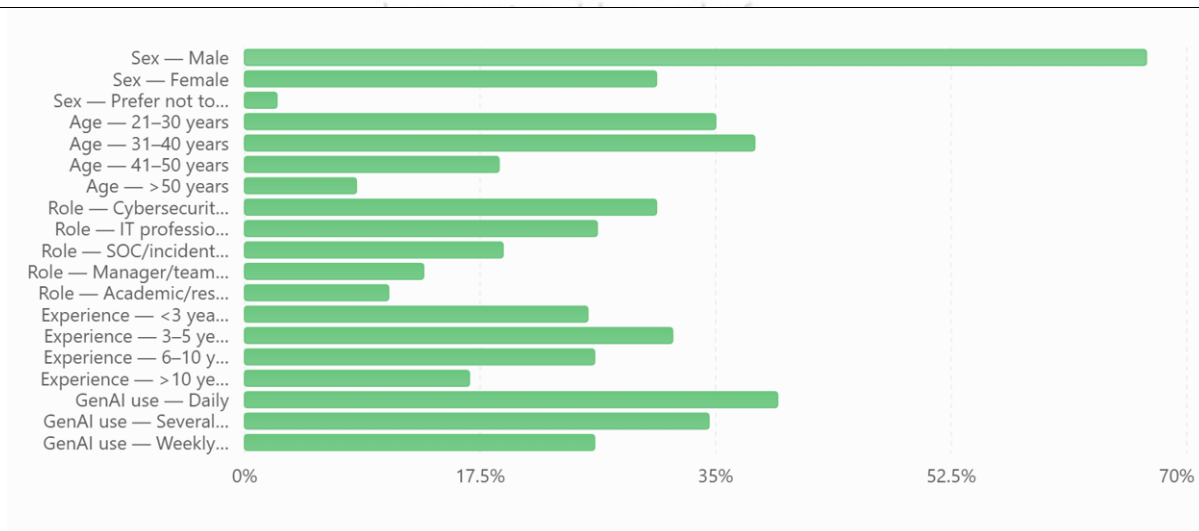


Figure 1 Demographic and Professional Characteristics of Participants

The participant profile included respondents with varying levels of cybersecurity experience and exposure to GenAI-assisted technologies, providing sufficient variation to examine differences in reliance, trust, oversight, automation bias, and perceptions of accountability.

Descriptive Statistics and Reliability

All study constructs were measured using multi-item, five-point Likert-type scales, with higher scores reflecting higher levels of the corresponding constructs. Internal consistency was satisfactory across all measures, with Cronbach's alpha coefficients ranging from .81 to .88.

The highest mean score was observed for human oversight ($M = 3.82$, $SD = 0.69$), followed by

perceived decision-making efficiency ($M = 3.79$, $SD = 0.68$). GenAI reliance was also relatively high mean score ($M = 3.61$, $SD = 0.71$), suggesting substantial integration of AI-generated

recommendations into cybersecurity decision processes. Automation bias had a comparatively moderate mean of 3.01 ($SD = 0.74$), while accountability diffusion averaged 2.89 ($SD = 0.80$).

Table 2. Descriptive Statistics and Internal Consistency of Study Variables

Construct	Mean	SD	Cronbach's α
GenAI reliance	3.61	0.71	.84
Decision efficiency	3.79	0.68	.86
Calibrated trust	3.54	0.72	.88
System explainability	3.47	0.76	.87
Human oversight	3.82	0.69	.83
Automation bias	3.01	0.74	.81
Responsibility ambiguity	3.08	0.77	.86
Accountability diffusion	2.89	0.80	.85

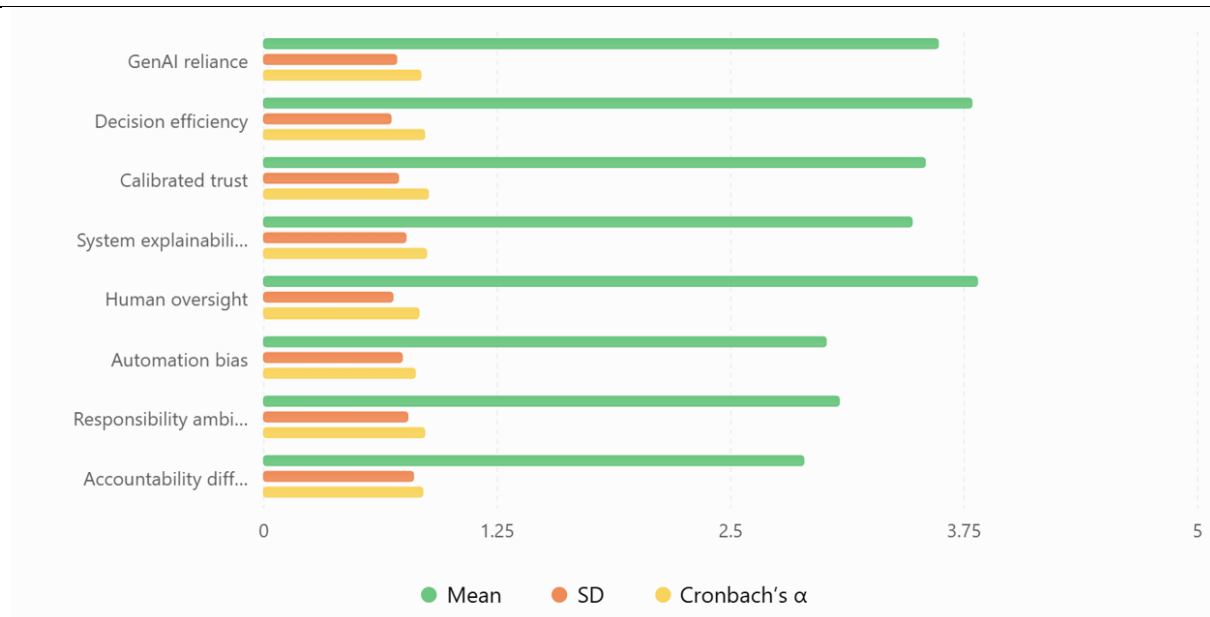


Figure 2 Descriptive Statistics and Internal Consistency of Study Variables

Reliability coefficients exceeded the conventional .70 threshold for all scales, indicating satisfactory internal consistency and supporting their use in subsequent correlational and regression analyses.

Bivariate Associations

Pearson correlation analysis identified several significant relationships among GenAI use, human-centered system characteristics, and cybersecurity decision outcomes.

GenAI reliance showed a strong positive association with perceived decision efficiency ($r = .52, p < .001$), indicating that respondents who relied more on GenAI generally perceived their cybersecurity decisions as faster and more efficient. GenAI reliance was also positively

associated with calibrated trust ($r = .47, p < .001$). However, greater reliance was also associated with higher automation bias ($r = .39, p < .001$), suggesting that increased dependence on AI recommendations may heighten the tendency to accept system outputs without sufficient independent verification.

System explainability showed a strong positive correlation with calibrated trust ($r = .56, p < .001$), whereas human oversight was negatively associated with automation bias ($r = -.41, p < .001$). Responsibility ambiguity was positively correlated with accountability diffusion ($r = .49, p < .001$), indicating that uncertainty about who was responsible for AI-assisted decisions was linked to weaker perceptions of individual accountability.

Table 3. Selected Pearson Correlations among Major Study Variables

Variables	r	p
GenAI reliance – Decision efficiency	.52	<.001
GenAI reliance – Calibrated trust	.47	<.001
GenAI reliance – Automation bias	.39	<.001
GenAI reliance – Accountability diffusion	.26	<.001
Explainability – Calibrated trust	.56	<.001
Explainability – Decision efficiency	.38	<.001
Explainability – Automation bias	-.17	.001
Human oversight – Decision efficiency	.29	<.001
Human oversight – Automation bias	-.41	<.001
Human oversight – Accountability diffusion	-.28	<.001
Responsibility ambiguity – Accountability diffusion	.49	<.001

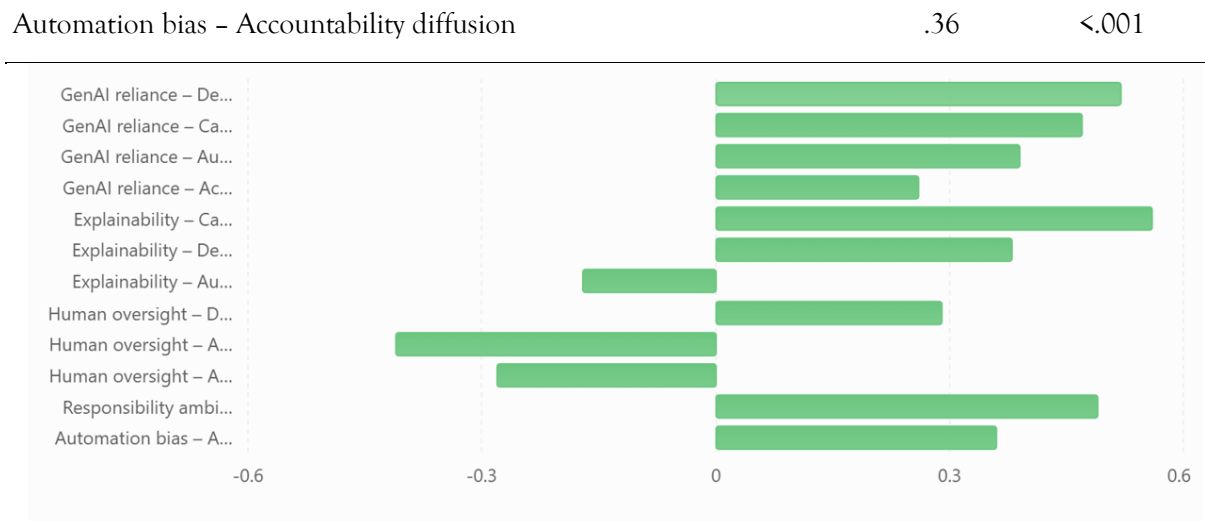


Figure 3 Selected Pearson Correlations among Major Study Variables

These findings demonstrate that the benefits and risks associated with GenAI were not mutually exclusive. Greater reliance was associated with improved perceived efficiency and trust, but also with increased susceptibility to automation bias. Conversely, explainability and human oversight appeared to operate as protective system characteristics by supporting calibrated trust and reducing uncritical dependence on AI recommendations.

Multivariable Regression Analysis

Estimated multiple linear regression models to determine whether GenAI reliance, explainability, human oversight, responsibility ambiguity, and related characteristics independently predicted decision efficiency, calibrated trust, automation bias, and accountability diffusion.

Table 4. Multivariable Regression Models for Cybersecurity Decision Outcomes

Outcome	Predictor	Standardized β	95% CI	p
Decision efficiency	GenAI reliance	.41	.32 to .50	<.001
	Explainability	.21	.12 to .30	<.001
	Human oversight	.18	.09 to .27	<.001
	Professional experience	.10	.02 to .18	.014
	Automation bias	.05	-.03 to .13	.214
Calibrated trust	Explainability	.34	.25 to .43	<.001
	GenAI reliance	.27	.18 to .36	<.001
	Human oversight	.16	.08 to .24	<.001
	Responsibility ambiguity	-.12	-.20 to -.04	.004
	Automation bias	.07	-.01 to .15	.083

Automation bias	GenAI reliance	.32	.23 to .41	<.001
	Human oversight	-.29	-.38 to -.20	<.001
	Responsibility ambiguity	.17	.09 to .25	<.001
	Professional experience	-.13	-.21 to -.05	.002
	Explainability	-.11	-.19 to -.03	.007
Accountability diffusion	Responsibility ambiguity	.31	.22 to .40	<.001
	Automation bias	.22	.13 to .31	<.001
	Human oversight	-.18	-.27 to -.09	<.001
	Explainability	-.10	-.18 to -.02	.014
	GenAI reliance	.08	-.01 to .17	.079

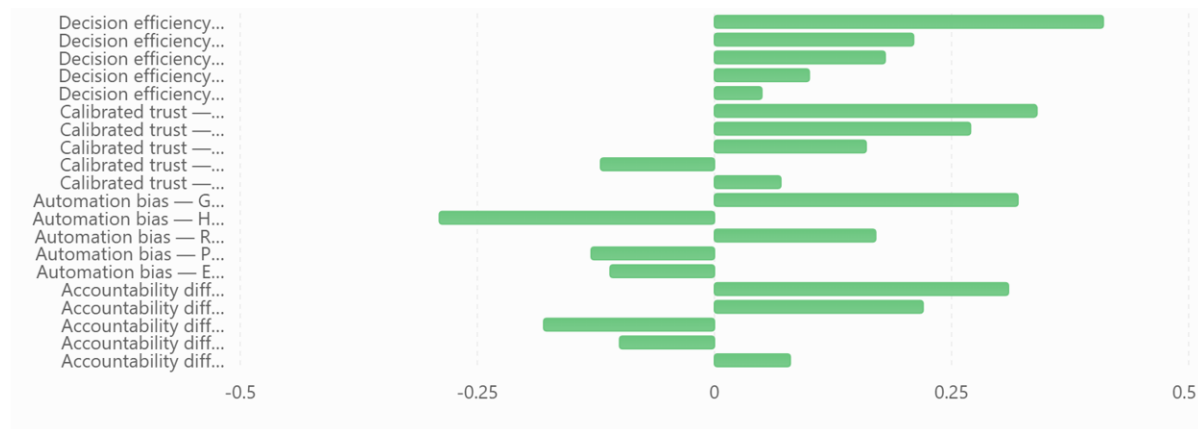


Figure 4 Multivariable Regression Models for Cybersecurity Decision Outcomes

Model statistics:

Decision efficiency: $R^2 = .38$, adjusted $R^2 = .37$, $p < .001$; calibrated trust: $R^2 = .43$, adjusted $R^2 = .42$, $p < .001$; automation bias: $R^2 = .31$, adjusted $R^2 = .30$, $p < .001$; accountability diffusion: $R^2 = .28$, adjusted $R^2 = .27$, $p < .001$.

GenAI and Decision Efficiency

The regression model predicting decision efficiency was statistically significant and accounted for 38% of the variance. GenAI reliance was the strongest independent predictor of perceived decision efficiency ($\beta = .41$, $p < .001$). Explainability ($\beta = .21$, $p < .001$) and human oversight ($\beta = .18$, $p < .001$) were also significant positive predictors.

These findings indicate that GenAI-supported cybersecurity environments were perceived as

more efficient when users actively incorporated AI-generated information into decision-making processes. However, efficiency was particularly evident when the system provided understandable outputs and maintained meaningful human oversight.

Automation bias did not independently predict decision efficiency after controlling for other variables ($\beta = .05$, $p = .214$). This suggests that efficiency gains were more closely tied to productive GenAI integration than to unquestioning acceptance of AI recommendations.

Predictors of Calibrated Trust

The calibrated-trust model explained 43% of the variance. Explainability was the strongest predictor ($\beta = .34$, $p < .001$). Participants who

reported greater understanding of how and why the GenAI system produced its recommendations exhibited higher levels of appropriately calibrated trust.

GenAI reliance also independently predicted calibrated trust ($\beta = .27, p < .001$), whereas human oversight showed a smaller but significant positive association ($\beta = .16, p < .001$). In contrast, responsibility ambiguity negatively predicted calibrated trust ($\beta = -.12, p = .004$).

The results indicate that trust in GenAI-assisted cybersecurity systems was not determined solely by exposure to or use of AI. Rather, trust was stronger when AI outputs were interpretable and when human users maintained an active supervisory role in the final decision-making process.

Predictors of Automation Bias

The model predicting automation bias explained 31% of the variance. GenAI reliance was positively associated with automation bias ($\beta = .32, p < .001$), confirming that greater dependence on AI-generated recommendations was linked to a stronger tendency to accept them without sufficient independent scrutiny.

Human oversight showed a significant inverse association with automation bias ($\beta = -.29, p < .001$). This was among the strongest effects in the model, indicating that maintaining human review and intervention substantially reduced the tendency to accept GenAI outputs uncritically.

Professional experience was also inversely associated with automation bias ($\beta = -.13, p = .002$), indicating that more experienced cybersecurity professionals were less likely to automatically defer to AI-generated recommendations. Explainability was similarly associated with lower automation bias ($\beta = -.11, p = .007$).

Collectively, these results indicate that automation bias poses a meaningful behavioral risk in GenAI-assisted cybersecurity, particularly when users become highly dependent on system recommendations without sufficient oversight or interpretability.

Predictors of Accountability Diffusion

The final regression model explained 28% of the variance in perceived diffusion of accountability. Responsibility ambiguity was the strongest predictor ($\beta = .31, p < .001$). Participants were therefore more likely to perceive responsibility as dispersed or unclear when organizational arrangements failed to specify whether accountability rested with the human operator, the system developer, the organization, or the AI-supported decision process.

Automation bias was also positively associated with accountability diffusion ($\beta = .22, p < .001$), whereas human oversight was negatively associated with accountability diffusion ($\beta = -.18, p < .001$). Explainability showed a smaller negative association ($\beta = -.10, p = .014$).

After controlling for these factors, GenAI reliance was not a statistically significant independent predictor of accountability diffusion ($\beta = .08, p = .079$). This finding matters because it suggests that using GenAI alone does not necessarily weaken accountability. Instead, accountability problems appear to arise primarily when responsibility structures are ambiguous, users defer excessively to automated recommendations, and meaningful human oversight is limited.

Overall Interpretation of Findings

The results reveal a dual effect of GenAI within intelligent cybersecurity systems. Greater reliance on GenAI was associated with substantially higher perceived decision efficiency and greater trust in system recommendations, supporting the operational value of AI-assisted cybersecurity decision-making. At the same time, greater reliance increased automation bias, demonstrating that efficiency benefits can coexist with behavioral risks.

Explainability emerged as a central determinant of calibrated trust, while human oversight played an important protective role against automation bias and accountability diffusion. Responsibility ambiguity was particularly important for understanding accountability, indicating that unclear allocation of responsibility contributed more directly to accountability diffusion than GenAI use itself.

Overall, the findings indicate that effective human-GenAI collaboration in cybersecurity depends not only on the technical capabilities of intelligent systems but also on how those systems are governed and integrated into professional decision-making processes. GenAI appears most beneficial when it serves as decision support rather than as an autonomous decision authority, with understandable recommendations, active human verification, and clearly defined responsibility for consequential cybersecurity decisions.

Discussion

This study investigated the impact of generative artificial intelligence on decision-making, trust, automation bias, and accountability within intelligent cybersecurity systems. The results indicate that GenAI enhances perceived decision efficiency and facilitates calibrated trust; however, increased reliance may also elevate automation bias. Explainability and human oversight were identified as significant protective factors, whereas responsibility ambiguity correlated with greater accountability diffusion.

Reliance on GenAI showed a strong association with decision efficiency ($r = .52, p < .001$) and remained the strongest predictor in the regression analysis ($\beta = .41, p < .001$). This result aligns with existing evidence that large language models facilitate cybersecurity operations by enabling rapid information processing, threat interpretation, and decision support (Hasanov et al., 2024). It further reinforces the concept of human-AI symbiosis, wherein AI provides computational efficiency and humans contribute contextual judgment and professional expertise (Jarrahi, 2018). Given the cross-sectional design, these findings reflect associations rather than causal relationships.

Explainability emerged as the strongest predictor of calibrated trust ($\beta = .34, p < .001$). This finding indicates that users are more likely to develop appropriate trust when they comprehend the mechanisms and rationale underlying AI-generated recommendations. Glikson and Woolley (2020) similarly identified transparency and system reliability as primary determinants of

trust in AI, while Shin (2021) demonstrated that explainability enhances users' trust and acceptance of intelligent systems. Within cybersecurity, explainable recommendations enable analysts to evaluate whether AI-generated outputs align with available evidence, rather than accepting them uncritically.

GenAI reliance was also positively associated with automation bias ($r = .39, p < .001$; $\beta = .32, p < .001$). This result suggests that increased dependence on AI recommendations may heighten the likelihood of accepting system outputs without adequate independent verification. Bućinca et al. (2021) similarly observed that users may over-rely on incorrect AI recommendations and demonstrated that active cognitive engagement can mitigate this behavior. Romeo and Conti (2026) identify automation bias as a significant challenge in human-AI collaboration. These findings are especially pertinent to cybersecurity, as AI-generated responses may appear technically credible even when erroneous.

Human oversight demonstrated a significant protective effect, exhibiting negative associations with both automation bias ($\beta = -.29, p < .001$) and accountability diffusion ($\beta = -.18, p < .001$). These results indicate that substantive human supervision can mitigate excessive reliance on AI and maintain responsibility for security decisions. Langer et al. (2025) emphasize that effective oversight necessitates that humans possess both the capability and authority to identify problematic AI outputs and intervene as needed. Consequently, human involvement should entail active review and decision-making authority, rather than serving as a formality.

Responsibility ambiguity emerged as the strongest predictor of accountability diffusion ($\beta = .31, p < .001$). Notably, GenAI reliance was not a significant independent predictor of accountability diffusion after controlling for other variables ($\beta = .08, p = .079$). These findings suggest that accountability challenges are primarily attributable to unclear organizational responsibility rather than AI utilization alone. Novelli et al. (2024) similarly contend that AI

accountability relies on clearly defined roles, answerability, oversight, and responsibility among both human and technological actors.

In summary, the findings illustrate a dual effect of GenAI in cybersecurity. Increased reliance can enhance perceived efficiency but may also elevate automation bias. Explainability supports calibrated trust, while substantive human oversight and clearly defined responsibility structures mitigate behavioral and governance risks. These results indicate that GenAI should serve as a decision-support tool rather than as an autonomous substitute for professional judgment. This study has several limitations. The cross-sectional design precludes causal inference, and reliance on self-reported measures may not accurately capture actual behavior during cybersecurity incidents. The purposive sampling approach further restricts generalizability. Future research should employ experimental and longitudinal designs, objective cybersecurity performance metrics, and more diverse professional samples.

In conclusion, GenAI can enhance cybersecurity decision-making when paired with transparent system design, active human oversight, and clearly defined accountability mechanisms. Effective human-AI collaboration depends not only on technological capabilities but also on the quality of human judgment and organizational governance.

Conclusion

This study investigated the impact of generative artificial intelligence on human decision-making, trust, automation bias, and accountability within intelligent cybersecurity systems. The results show that GenAI offers significant potential to enhance cybersecurity operations, particularly by increasing perceived decision efficiency and helping process complex security information. However, the study also highlights notable behavioral and governance risks associated with greater reliance on AI-generated recommendations.

Reliance on GenAI was positively correlated with decision efficiency and calibrated trust, suggesting that cybersecurity professionals benefit from the speed, information synthesis, and analytical

support provided by intelligent systems. However, increased reliance was also linked to higher automation bias, indicating that greater dependence on AI may diminish critical scrutiny when users accept automated recommendations without sufficient evaluation. This dual relationship demonstrates that the value of GenAI should not be assessed solely based on efficiency or technical capability.

Explainability was identified as a critical factor in fostering calibrated trust. Users demonstrated greater confidence in GenAI when system outputs were clear and interpretable. Human oversight served a key protective function by mitigating automation bias and restricting the diffusion of accountability. These findings underscore the necessity of maintaining substantive human involvement in significant cybersecurity decisions, rather than delegating full decision authority to intelligent systems.

The study further demonstrates that accountability challenges are significantly shaped by ambiguity in responsibility. When organizational roles and decision authority lack clarity, responsibility for AI-assisted outcomes may be distributed among analysts, managers, developers, and automated systems. Establishing clear accountability structures is therefore essential to ensure that GenAI strengthens, rather than undermines, responsible cybersecurity governance.

Overall, the findings indicate that GenAI should function as a decision-support tool rather than replace professional judgment. Effective human-AI collaboration depends on transparent system design, calibrated trust, active human verification, relevant professional expertise, and clearly defined responsibility for final decisions. Consequently, cybersecurity organizations should supplement technological investments with governance frameworks, user training, and oversight mechanisms that maintain human agency and accountability.

Future research should assess GenAI-assisted cybersecurity decisions through experimental and longitudinal designs, objective measures of decision accuracy, and diverse professional

samples. Such studies would clarify how trust and reliance develop over time and identify the circumstances in which GenAI yields sustainable improvements in cybersecurity performance. Ultimately, the effectiveness of intelligent cybersecurity systems will depend not only on technological advancement but also on the responsibility with which humans design, supervise, and utilize these systems.

REFERENCES

- Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1-21. DOI: 10.1145/3449287
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627-660. DOI: 10.5465/annals.2018.0057
- Hasanov, I., Virtanen, S., Hakkala, A., & Isoaho, J. (2024). Application of large language models in cybersecurity: A systematic literature review. *IEEE Access*, 12, 176751-176778. DOI: 10.1109/ACCESS.2024.3505983
- Henrique, B. M., & Santos, E., Jr. (2024). Trust in artificial intelligence: Literature review and main path analysis. *Computers in Human Behavior: Artificial Humans*, 2(1), 100043. <https://doi.org/10.1016/j.chbah.2024.100043>
- Ahmad, S. (2025). Statistical Analysis of Leadership Styles and Their Impact on Hierarchical Effectiveness in Organizations. *Global Journal of Sciences*.
- Ahmad, S., & Museera, S. (2024). The Strategic Influence of Cloud Computing on Contemporary Marketing and Management Practices. *Journal of Engineering and Computational Intelligence Review*, 2(2).
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577-586. DOI: 10.1016/j.bushor.2018.03.007
- Ahmad, S., Wong, W. K., Riaz, S., & Iqbal, A. (2024). The role of employee motivation and its impact on productivity in modern workplaces. *Arabian Journal of Business and Management Review*, 13(2).
- Kumar, M., Rani, R., Epiphaniou, G., & Maple, C. (2026). Adaptive trust-aware SOC human-AI teaming for resilient operations. *Computers & Security*, 169, 104971. DOI: 10.1016/j.cose.2026.104971
- Langer, M., Baum, K., & Schlicker, N. (2025). Effective human oversight of AI-based systems: A signal detection perspective on the detection of inaccurate and unfair outputs. *Minds and Machines*, 35, Article 1. <https://doi.org/10.1007/s11023-024-09701-0>
- Ahmad, S., Nisa, G., Shah, S. H. I., & Manzoor, H. (2025). Exploring the Interplay of Creativity and Proactive Personality: Mediating Role of Self-Efficacy. *Research Journal for Social Affairs*, 3(4).
- Ahmad, S., et al. (2025). An Investigation into the Relationship between Emotional Intelligence and Academic Success among University Students. *Pegem Journal of Education and Instruction*, 15(1).
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50-80. DOI: 10.1518/hfes.46.1.50_30392
- Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society*, 39, 1871-1882. <https://doi.org/10.1007/s00146-023-01635-y>

- Ahmad, S. (2025). Entrepreneurship and Sustainable Leadership Practices: How entrepreneurial leaders incorporate sustainability. *Journal of Entrepreneurship and Business Venturing*, 5(1).
- Ahmad, S. (2025). Statistical Analysis of Leadership Styles and Their Impact on Hierarchical Effectiveness in Organizations. *Global Journal of Sciences*.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. DOI: 10.1518/001872097778543886
- Romeo, G., & Conti, D. (2026). Exploring automation bias in human-AI collaboration: A review and implications for explainable AI. *AI & Society*, 41, 259–278. DOI: 10.1007/s00146-025-02422-7
- Romeo, G., & Conti, D. (2026). Exploring automation bias in human-AI collaboration: A review and implications for explainable AI. *AI & Society*, 41, 259–278. <https://doi.org/10.1007/s00146-025-02422-7>
- Ahmad, S. (2025). Entrepreneurship and Sustainable Leadership Practices: How entrepreneurial leaders incorporate sustainability. *Journal of Entrepreneurship and Business Venturing*, 5(1).
- Ahmad, S. (2024). The Impact of Decision-making by Charismatic Leadership in Conflicted and Tangled Circumstances. *KASBIT Journal*, 17(1).
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. DOI: 10.1016/j.ijhcs.2020.102551
- Tsamados, A., Floridi, L., & Taddeo, M. (2025). Human control of AI systems: From supervision to teaming. *AI and Ethics*, 5, 1535–1548. <https://doi.org/10.1007/s43681-024-00489-4>
- Yaich, R., Balondrade, A., Sicard, A., Fouquiau, C., Giraud, G., Amokrane-Ferka, K., & Arbaretier, E. (2025). Symbiotic human-AI collaboration for augmented cybersecurity operations. *Proceedings of the AAAI Symposium Series*, 6(1), 350–358. DOI: 10.1609/aaais.v6i1.36072
- Zhang, J., Bu, H., Wen, H., Liu, Y., Fei, H., Xi, R., Li, L., Yang, Y., Zhu, H., & Meng, D. (2025). When LLMs meet cybersecurity: A systematic literature review. *Cybersecurity*, 8, 55. DOI: 10.1186/s42400-025-00361-w
- Ahmad, S. (2026). Navigating Polycontextual Global Virtual Work in the AI Era: When Leadership Is the Hub A Resource Caravan Model of Institutional Navigation, AI-Mediated Communication, and Global Leadership: Associations with Trust, Performance, and Well-Being. *Journal of Intelligent Decision Making and Information Science*, 3(11s), 1942-1969.
- Ahmad, S., Smerat, A., Wong, D. W.-K., & Bilal, D. M. (2025). A Comparative Statistical Study on Social Media Usage and Its Impact on the Mental Well-Being of Young Employees. *Sigmae*, 14(4).
- Choudhry Shahid, N. T. D. A. Khan, S. A. Prof. Dr. M. B. O. A. S. A. (2026). Designing An Ai Governance Framework For Generative Ai In Higher Education: Integrating Large Language Models, Academic Integrity, Employability, And National Ai Policies. *Journal of Intelligent Decision Making and Information Science*, 3(12s).
- Inam, M. (2026). Extreme Heat Exposure and Acute Cardiovascular Events in Pakistan: Clinical Risk, Multisystem Complications and Health-System Resilience. *Natural Resources for Human Health*, 6(12s), 1212-1228.

Kamboh, M. A., Ullah, I., Zeb, S. S., Saleemi, M. W., Athar, F., Ahmad, S., & Bayram, M. (2026). Climate shocks and unequal educational opportunity in Pakistan: household economic strain and the political-institutional context. *Aposta: Revista De Ciencias Sociales*, 24(112), e2202. <https://doi.org/10.23900/ra.v24i112.2202>

