ACADEMIA International Journal for Social Sciences
Volume 5, Issue 1(a), 2026 ISSN-L (Online): 3006-6638

# A Comparative Study of Trust in AI-Driven Security Alert Systems and Human Cybersecurity Experts in IT Environments

**Marriyam Irshad**
marryam.irshad@akhuwat.edu.pk
Lecturer IT, Akhuwat College for Women Chakwal, Pakistan

**Syeda Maryam Haider Gardezi**
maryamhaider2004@gmail.com
BSCYS Scholar, Air University Multan Campus, Islamabad, Pakistan

**Umme Ruman**
rumiruman74@gmail.com
BSCYS Scholar, Air University Multan Campus, Islamabad, Pakistan

**Arhum Luqman**
arhumluqman786@gmail.com
BSCYS Scholar, Air University Multan Campus, Islamabad, Pakistan

**Dr. Muhammad Arfan Lodhi**
samaritan_as@hotmail.com
Higher Education Department, Punjab, Pakistan

**Corresponding Author: Dr. Muhammad Arfan Lodhi** samaritan_as@hotmail.com

**Received:** 24-11-2025 **Revised:** 08-12-2025 **Accepted:** 21-12-2025 **Published:** 05-01-2026

## ABSTRACT

*The proliferation of automated cyber security tools including antivirus software, browser security warnings, email phishing alert systems, and workplace security platforms has fundamentally altered how individuals encounter and respond to digital threats. Concurrently, human security experts continue to provide context-sensitive advisory services, raising critical questions about how end users comparatively trust these two advisory channels. Despite growing organizational reliance on automated security systems, limited empirical evidence exists regarding how non-expert and semi-technical users perceive, compare, and act upon automated security alerts versus advice from human security experts. Understanding this trust dynamic is essential for designing more effective cyber security systems and education programmes. This study employed a cross-sectional quantitative survey design. Data were collected from 31 valid respondents using a structured 5-point Likert-scale questionnaire organized into five sections covering demographics, automated trust, human expert trust, comparative preferences, and behavioral intentions. Participants were primarily undergraduate students aged 18–24 with varying technical backgrounds, recruited via Google Forms during April 2026.Findings reveal moderate trust in both automated alerts (Composite M = 3.50, SD = 0.73) and human experts (Composite M = 3.39, SD = 0.81). When sources conflicted, 77.4% of respondents preferred to seek additional information rather than defer exclusively to either advisory channel. In high-risk scenarios, 51.6% favoured relying on both sources equally, 32.3% preferred human experts, and only 16.1% preferred automated systems alone. Users occupy a deliberative middle ground, valuing both the speed and consistency of automated alerts and the contextual judgment of human experts. The findings carry implications for cybersecurity interface design, security education, and institutional advisory policy.*

***Keywords:*** *cyber security trust, automated security alerts, human security experts, algorithm aversion, security decision-making*

## INTRODUCTION

### Background of the Study

In the contemporary digital landscape, cybersecurity threats have grown exponentially in both frequency and sophistication. Cybercrime damages were estimated at $8.44 trillion USD globally in 2022 and are projected to reach $13.82 trillion by 2028 (ACM Computing Surveys, 2025). Phishing attacks skyrocketed by an estimated 4,151% since the widespread adoption of large language models beginning in late 2022, and the human element was implicated in 68% of all data breaches reported in 2024 (Keepnet, 2026; Proofpoint, 2024). These statistics underscore the urgency of understanding how users interact with the two primary advisory channels now available in cybersecurity: automated alert systems and human security experts.

Xie and Zhou (2025) conducted two controlled online experiments with a combined sample of 599 participants to examine how attributing emergency alerts to AI versus human experts shapes public perceptions of credibility and behavioral coping outcomes. They found that AI-generated alerts were perceived as significantly less credible than those issued by human experts, and that errors by AI sources produced greater credibility losses than equivalent errors by human sources, despite AI being assigned less moral blame. These asymmetries — greater credibility erosion for AI errors alongside lower blame attribution — suggest that users hold automated systems to a higher reliability standard than human advisors.

Horowitz and Kahn (2024) contributed a landmark preregistered experiment across 9,000 adults in nine countries, examining the relationship between AI knowledge, trust in AI, and automation bias in a threat-identification task modeled on national security scenarios. Their findings confirmed a Dunning-Kruger dynamic: participants with the lowest AI knowledge showed algorithm aversion, those with moderate knowledge exhibited automation bias, and only those with the highest knowledge showed calibrated reliance. This curvilinear pattern has significant implications for cybersecurity, where users range from complete novices to certified professionals with very different starting points for trust.

The phenomenon of automation bias the systematic tendency to accept automated recommendations uncritically was reviewed extensively by Kueper et al. (2025) across 96 empirical studies published between 2015 and 2025. They found that trust consistently accounted for up to 24.1% of the variance in reliant behavior, and that negative experiences with automated systems erode trust in ways resistant to easy recovery. Within Security Operations Centers (SOCs), IJCTECE (2025) demonstrated that moderate automation promotes healthy trust and improved decision accuracy, while high automation induces over-reliance and reduces independent alertness. From the Technology Acceptance Model (TAM) perspective (Davis, 1989), users' decisions to act upon technological systems are mediated by perceived usefulness, perceived ease of use, and in cybersecurity-extended versions trust. Springer (2025) found that cybersecurity awareness strongly mediates trust and perceived risk as determinants of behavioral intention toward AI-powered security tools. Alert fatigue, a direct consequence of high false-positive rates documented at up to 99% in some enterprise SOC deployments (ACM, 2025), is a primary mechanism by which trust in automated alerts erodes over time. The Threat Quotient (2023) annual automation adoption survey confirmed this: lack of trust in automation outcomes was identified as the single most cited challenge by cybersecurity leaders, ahead of budget and regulatory concerns.

The algorithm aversion-appreciation literature provides a further theoretical lens. Dietvorst et al. (2015) formally characterized algorithm aversion — users rejecting algorithmic advice even when it demonstrably outperforms human judgment — while Logg et al. (2019) documented the opposing phenomenon of algorithm appreciation in many domains. Jussupow et al. (2024) resolved this apparent contradiction in a MIS Quarterly integrative review, concluding that task domain, user expertise, and perceived decision stakes determine the direction of the effect. Schemmer et al. (2023) synthesized 96 empirical trust calibration studies and confirmed that trust is a dynamic, context-sensitive judgment rather than a stable attitude, supporting the present study's multi-dimensional approach to measuring trust across different advisory source types.

KPMG (2024) surveyed senior organizational leaders on AI in security and found that 38% identified trusting AI reliability as their top operational concern and 32% reported difficulty determining threat severity even with automated tools deployed. Ebert et al. (2022) demonstrated that security warning salience significantly affects initial compliance, but that habituation rapidly erodes effectiveness under repeated exposure — establishing an important ceiling on automated alert trust. Korka et al. (2023) extended this to argue that design-level interventions centred on contextual relevance and graduated urgency are necessary to combat habituation. Proof point (2024) documented a consistent gap between security teams' perceptions of employee awareness and actual end-user risk behaviors, while MDPI (2025) found moderate-to-high trust in AI-based phishing detection tools among university students, with trust positively associated with cybersecurity education exposure. Taken together, these twenty studies establish the theoretical and empirical foundation motivating the present investigation.

**Statement of the Problem**

The central challenge in contemporary cybersecurity is not merely the technical detection of threats, but the effective communication of those threats to end users in a manner that produces appropriate, timely responses. Automated systems offer speed, scalability, and rule-based consistency. Human security experts offer contextual judgment, nuanced interpretation, and interpersonal trust. However, users frequently lack the technical expertise to evaluate the reliability of either source independently, and the literature reveals substantial heterogeneity in how different populations trust each channel. A critical gap exists as there is an absence of empirical studies directly comparing end-user trust in automated security alerts versus human expert advice among non-professional, semi-technical populations. Existing research predominantly focuses on professional SOC analysts (IJCTECE, 2025; Lim et al., 2025), organizational leaders (KPMG, 2024), or broad domain-agnostic samples (Horowitz & Kahn, 2024). The perceptions and behavioral intentions of university students and young adults who are frequent targets of phishing and social engineering and who increasingly encounter automated security systems in academic and professional contexts remain empirically underexplored. Furthermore, most existing studies examine trust in either automated systems or human experts in isolation rather than in direct comparison. When both sources conflict, users face a meta-cognitive trust resolution problem: which channel is more reliable, and under what circumstances? This comparative dimension has not been systematically addressed for non-expert populations in cybersecurity. The present study directly addresses this gap. The rationale for focusing on undergraduate students is both theoretical — they span a meaningful range of technical competencies — and practical: the security behaviors formed during university years tend to persist into professional life, making this population an important target for intervention.

**Research Questions**

This study is guided by the following five research questions:

**RQ1:** To what extent do undergraduate students trust automated security alerts compared to human security experts, and how do trust levels differ across the dimensions of reliability, accuracy, confidence, risk reduction, and contextual understanding?

**RQ2:** What behavioral intentions do students report when automated security alerts and human expert advice are in conflict, and how do technical background and prior expert consultation experience moderate these intentions?

**RQ3:** In high-risk security scenarios, do students prefer to rely on automated systems, human experts, or a combination of both, and what individual characteristics are associated with each preference?

**RQ4:** How does students' frequency of encountering automated security alerts relate to their trust levels in automated systems versus human expert advisory channels?

**RQ5:** To what extent do students' perceptions of automated system consistency and human expert contextual understanding predict their stated preference for each advisory source in low-risk versus high-risk security scenarios?

**Significance of the Study**

This study makes a direct empirical contribution to the cybersecurity and human-computer interaction literature by providing one of the first systematic comparisons of trust in automated security alerts versus human security experts among a non-professional student population. Existing reviews have repeatedly called for more empirical work on non-expert populations in cybersecurity trust contexts (Schemmer et al., 2023; Kueper et al., 2025), and this study responds to that call by generating quantitative, instrument-based data that can be benchmarked against future investigations and used to track longitudinal shifts in trust dispositions as AI security systems continue to mature. Second, the findings carry practical implications of immediate relevance to cybersecurity education designers, university IT security departments, and organizational security policy architects. If users overwhelmingly prefer to seek additional information when advisory sources conflict — as 77.4% of this sample did — then the current dominant design philosophy of binary, comply-or-dismiss security alert interfaces is fundamentally misaligned with users' natural decision-making preferences. Cybersecurity training programmes should accordingly shift focus from mandating compliance with automated alert protocols toward building users' capacity to critically evaluate, compare, and synthesize information from multiple advisory sources. This reorientation toward deliberative triangulation as a security competency has practical implications for curriculum design, security awareness campaigns, and the design of human-AI collaborative interfaces. Third, the study generates evidence with direct policy implications at institutional and national levels. At the national policy level, the evidence supports incorporating human-AI collaborative decision-making training into cybersecurity certification programs and calls for regulatory frameworks governing automated security system certification to include precision thresholds as a trust-preservation standard, directly addressing the alert fatigue crisis documented by ACM (2025) and Threat Quotient (2023).

## REVIEW OF RELATED LITERATURE

### Theoretical Foundations

### Automated Security Alert Systems: An Introduction

Automated security alert systems are software-driven mechanisms that continuously monitor digital environments — networks, endpoints, email servers, web browsers, and enterprise platforms — to detect, classify, and communicate potential cybersecurity threats to users in real time. These systems encompass a broad taxonomy of tools: antivirus and anti-malware software that scans files and processes against known threat signatures, browser security warning systems that flag suspicious URLs, phishing detection platforms that analyze email headers and embedded links for known attack patterns, Security Information and Event Management (SIEM) systems that aggregate and correlate alerts across an enterprise network, and Intrusion Detection and Prevention Systems (IDS/IPS) that monitor network traffic for anomalous behavior.

The operational logic of automated security alerts is fundamentally probabilistic: these systems generate alerts based on rule sets, machine learning classification models, or heuristic pattern matching. This probabilistic nature means that no automated system achieves perfect precision — false positives (legitimate activities flagged as threats) are an inherent feature of the technology rather than an implementation failure. The ACM Computing Surveys (2025) systematic review documented false-positive rates approaching 99% in some enterprise SOC environments, meaning that the overwhelming majority of alerts processed by security analysts in these environments are false alarms. This structural characteristic has profound implications for user trust: repeated exposure to false alarms induces alert fatigue, a well-documented desensitization phenomenon in which users learn to dismiss or ignore alerts as a habituated response to their perceived unreliability.

Despite these limitations, automated security alert systems offer advantages that human experts cannot replicate at scale: they can process billions of data points in milliseconds, operate continuously without degradation from fatigue, and apply consistent decision rules across an entire enterprise estate. The ThreatQuotient (2023) survey found that automation adoption for alert triage grew from 18% to 30% of organizations between 2022 and 2023, and phishing analysis was the most common automated use case (31%). These adoption rates reflect a growing organizational consensus that automated systems are necessary if not sufficient for effective cybersecurity at scale — a consensus that makes the question of user trust in these systems both practically urgent and theoretically important.

### Technology Acceptance Model (TAM)

The Technology Acceptance Model (Davis, 1989) posits that perceived usefulness (PU) and perceived ease of use (PEOU) shape users' attitudes toward technology, which in turn predict behavioral intention and actual use. Cybersecurity-specific extensions of TAM have incorporated trust as a third mediating construct. Springer (2025) demonstrated that cybersecurity awareness strongly influences trust and perceived risk, which emerged as the dominant mediators of behavioral intention toward AI-powered security tools. For this study, TAM provides the theoretical logic for understanding why users with different technical backgrounds may differentially trust automated versus human advisory channels: users with higher cybersecurity awareness are expected to hold more calibrated trust dispositions, while non-technical users may rely on heuristic cues — such as the perceived human origin of advice — as substitutes for technical evaluation.

**Trust Calibration Theory and Automation Bias**

Trust calibration refers to the alignment between a user's subjective trust in a system and that system's objective reliability. Holland et al. (2024) demonstrated experimentally that trust adapts dynamically to reliability changes, but that high initial trust is asymmetrically persistent — resisting reduction even after significant performance drops. This asymmetry is particularly relevant in cybersecurity, where users who develop initial confidence in an automated alert system may continue to trust it even after observing false positives. Kueper et al.'s (2025) systematic review identified trust as accounting for up to 24.1% of the variance in reliant behavior across 96 empirical studies, while automation bias — the tendency to uncritically accept automated recommendations — was confirmed across multiple high-stakes domains including aviation, medicine, and national security (Skitka et al., 1999). The algorithm aversion-appreciation continuum (Dietvorst et al., 2015; Logg et al., 2019) and its resolution by Jussupow et al. (2024) and Horowitz and Kahn (2024) further establish that the direction and magnitude of trust in automated versus human sources depends critically on task domain, expertise level, and perceived decision stakes.

**Review of Empirical Studies**

The empirical literature on comparative trust in automated versus human advisory systems spans disaster communication, national security, healthcare, and cybersecurity. Xie and Zhou (2025) established that AI-generated alerts are inherently less credible than human expert alerts, and that AI errors produce disproportionate trust losses — a pattern attributable to higher baseline reliability expectations for algorithmic systems. Horowitz and Kahn (2024) demonstrated the Dunning-Kruger dynamic in AI trust across a nine-country sample, establishing that moderate AI knowledge is uniquely associated with automation bias. KPMG (2024) documented that 38% of senior security leaders cite trusting AI reliability as their top concern, while ThreatQuotient (2023) confirmed that trust barriers are the leading obstacle to automation adoption across cybersecurity organizations.

Schemmer et al. (2023) synthesized 96 trust calibration studies and concluded that trust is a dynamic, context-sensitive judgment moderated by task risk, automation degree, and user expertise. MDPI (2025), in a study methodologically parallel to the present investigation, found moderate-to-high trust in AI phishing detection tools among university students, positively associated with cybersecurity education. Ebert et al. (2022) established that warning saliency influences initial compliance but habituation rapidly degrades effectiveness. The HAT-Lab (2025) study documented that even technically sophisticated users exhibit high susceptibility to AI manipulation in agentic systems due to cognitive tunneling. Proofpoint (2024) established the human element in 68% of all breaches, and ACM (2025) confirmed that alert fatigue resulting from near-100% false-positive rates in some SOC environments is a primary trust-erosion mechanism. Table 1 provides a structured meta-synthesis of these and related studies.

Table 1: Meta-Synthesis of Key Prior Studies

| Authors | Year | Topic | Findings | Gap |
|---|---|---|---|---|
| Xie & Zhou | 2025 | AI vs. human expert disaster alerts and public trust | AI alerts rated less credible; AI errors produce greater trust loss | No cybersecurity-specific context; disaster-domain only |

| | | | | |
|---|---|---|---|---|
| Horowitz & Kahn | 2024 | Automation bias in AI-based national security decisions | Dunning-Kruger effect; moderate AI knowledge drives automation bias | National security focus; everyday user perceptions excluded |
| Holland et al. | 2024 | Trust calibration in variable-reliability automation | Trust adapts dynamically; high initial trust asymmetrically persists | Low-stakes task; not security-specific |
| Kueper et al. | 2025 | Automation bias review in human-AI collaboration | Trust accounts for 24.1% of reliant behavior variance; 96 studies synthesized | Cybersecurity domain under-explored |
| IJCTECE | 2025 | Human-AI collaboration in SOC environments | Moderate automation promotes healthy trust and decision accuracy | Experimental; professional analysts only |
| Lim et al. | 2025 | Explain ability in AI-driven SOC interfaces | Context-agnostic AI explanations produce miscalibrated trust across analyst tiers | SOC professionals only; non-experts not studied |
| Springer TAM Extension | 2025 | TAM applied to AI-powered cybersecurity tool adoption | Cybersecurity awareness strongly mediates trust and behavioral intention | Does not directly compare AI vs. human expert trust |
| Ebert et al. | 2022 | Saliency of security warnings and protection behavior | Warning salience influences compliance; habituation rapidly erodes effectiveness | Does not compare automated vs. human advisory trust |
| ThreatQuotient | 2023 | State of cybersecurity automation adoption | 30% use automation for alert triage; lack of trust is the top adoption barrier | Industry report; self-selected respondents; no controlled sample |
| KPMG | 2024 | Leader perspectives on AI in Security Operations Centers | 38% cite trusting AI reliability as top concern; 32% struggle with threat severity | Executive survey only; end-user trust not measured |
| Proofpoint | 2024 | Human behavior and phishing; security culture gaps | Human element in 68% of breaches; gap between security teams and end-user perceptions | Breach statistics; no direct trust comparison |

| | | | | |
|---|---|---|---|---|
| Schemmer et al. | 2023 | Measuring trust calibration for automated systems (CHI) | Risk, expertise, and automation degree all moderate trust; trust is dynamic | No dedicated cybersecurity domain focus |
| HAT-Lab | 2025 | Human susceptibility to AI manipulation in LLM agentic systems | Only 8.6% perceive AI manipulation; IT experts show higher susceptibility | Adversarial manipulation focus; not routine security alert trust |
| MDPI | 2025 | AI-based phishing detection and student cybersecurity awareness | Moderate-to-high trust in AI phishing detection; education improves calibration | Limited to students; no comparison with human expert preferences |
| ACM | 2025 | Alert fatigue in Security Operations Centers | False positives near 99% in some SOCs; fatigue reduces trust in automated alerts | SOC-professional focus; general public not addressed |
| Korka et al. | 2023 | Design of privacy notices and security warnings (CHI Workshop) | User fatigue causes most warnings to be bypassed; design improvements reduce risk | Position paper; no quantitative trust comparison |
| Logg et al. | 2019 | Algorithm appreciation: preference for algorithmic vs. human judgment | People often prefer algorithmic advice; effect varies by task domain and stakes | Pre-dates current AI security tools; not security-domain specific |
| Jussupow et al. | 2024 | Integrative perspective on algorithm aversion and appreciation | Task domain, expertise, and stakes determine direction of trust effect | Broad domain review; cybersecurity not examined specifically |
| ArXiv / LLM Cyber | 2025 | LLMs bridging expertise gaps in cybersecurity human-AI teaming | LLM definitiveness influences trust; non-experts over-rely on confident AI | Laboratory setting; non-representative sample of general users |
| Dietvorst et al. | 2015 | Algorithm aversion: people erroneously avoid algorithms after errors | Users reject algorithmic advice even when it demonstrably outperforms human judgment | Economic/sports forecasting domain; not cybersecurity-specific |

*Note. Studies selected for recency (2015–2025), relevance to trust in automated vs. human security advisory systems, and empirical quality.*

**Conceptual Framework**

**Narrative Description**

The conceptual framework integrates the Technology Acceptance Model, Trust Calibration Theory, and the Automation Bias–Algorithm Aversion continuum into a three-variable model. The independent variable (IV) is user characteristics: technical background, prior expert consultation experience, alert frequency, gender, and education level. These characteristics shape trust dispositions through the mechanisms identified in TAM and trust calibration research. The dependent variable (DV) is comparative trust, operationalized across two five-item Likert subscales: Automated Alert Trust (M = 3.50, SD = 0.73) and Human Expert Trust (M = 3.39, SD = 0.81). The outcome variable (OV) is behavioral intention, captured through three items. The framework predicts that: (1) IT/ Cybersecurity users will show higher automated trust due to greater perceived usefulness and familiarity with automated tools; (2) users with prior expert consultation experience will show higher human expert trust due to direct trust-updating through experience; and (3) the dominant behavioral intention will be information-seeking across all subgroups, reflecting deliberative triangulation rather than exclusive deference to either advisory channel.

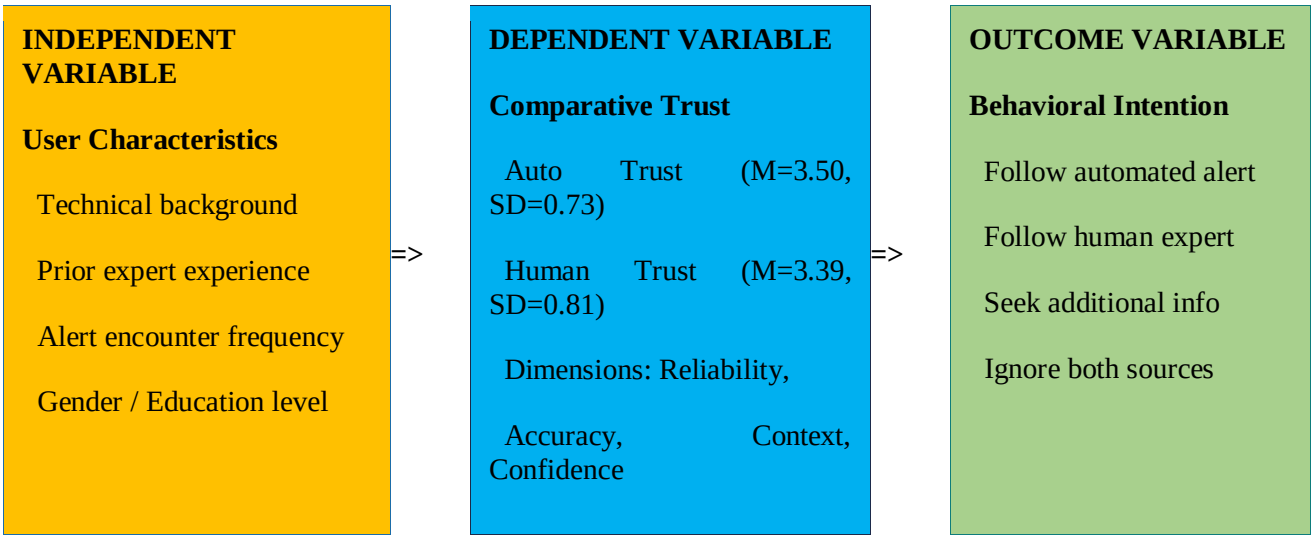


*Figure 1. Conceptual framework illustrating directional relationships among user characteristics (IV), comparative trust (DV), and behavioral intention (OV). Arrows denote hypothesized directional pathways.*

**RESEARCH METHODOLOGY**

**Research Design**

This study employed a cross-sectional, descriptive quantitative research design. Cross-sectional designs collect data at a single point in time, enabling characterization of prevailing attitudes and behavioral intentions without longitudinal resource requirements. The quantitative approach was selected because the research questions are inherently comparative, requiring measurement of trust constructs on a common numerical scale. Three variable categories structure the design. The independent variable (IV) comprises user characteristics: technical background, prior expert consultation, gender, education, and alert frequency.

The dependent variable (DV) is comparative trust, operationalized as mean composite scores on two five-item Likert subscales (Automated Alert Trust; Human Expert Trust). The outcome variable (OV) is behavioral intention, captured through three measures. The hypothesized relationship is directional: user characteristics shape trust dispositions (IV => DV), which determine behavioral intentions (DV => OV), with technical background moderating the IV => DV pathway.


*Figure 2. Conceptual map & meta sunthesis*

**Population and Sampling**

The target population comprised university students and recent graduates with varying technical knowledge, representing a demographic that routinely encounters automated security alerts but typically lacks professional cybersecurity expertise. Convenience sampling was employed, with participants recruited through electronic circulation of the Google Forms survey link within university networks during April 2026. Sampling criteria required: (1) current student or recent graduate status; (2) provision of informed consent at the survey outset; and (3) complete responses to all items. One response indicating non-consent was excluded, yielding N = 31.

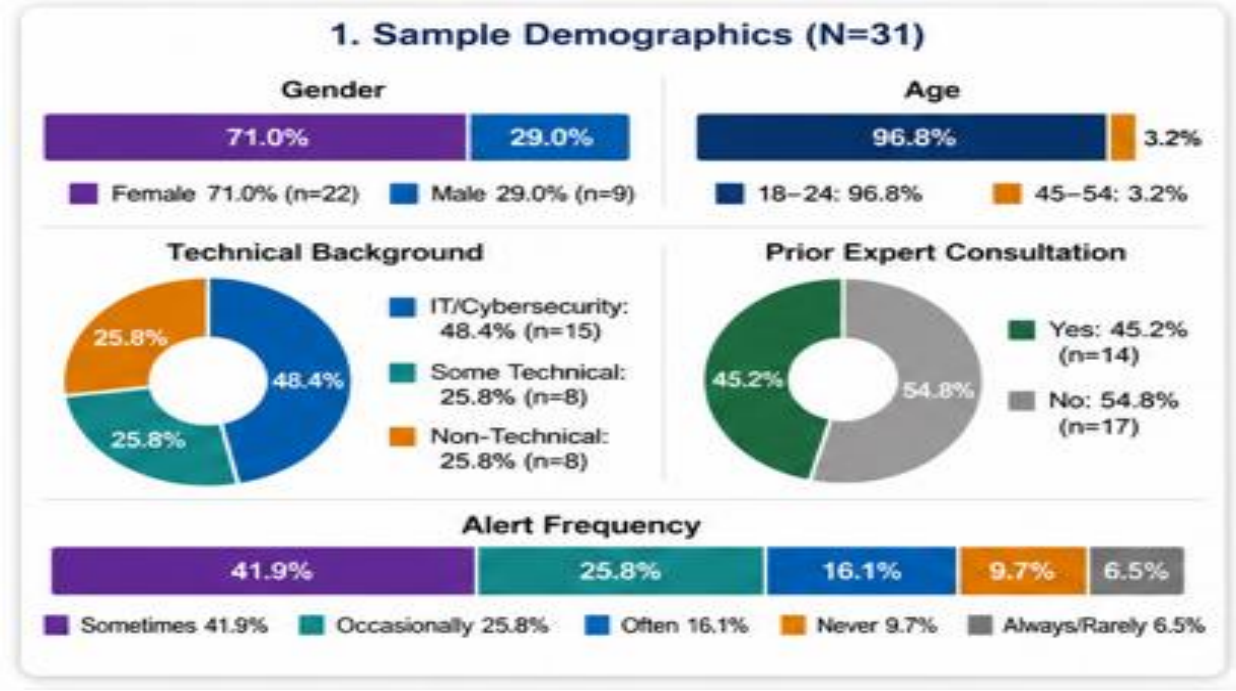


*Figure 3. Charts for sample demographics*

*Table 2: Sample Demographic Profile (N = 31)*

| **Variable** | **Category** | **n** | **% (N=31)** |
|---|---|---|---|
| **Gender** | Female | 22 | 71.0% |
| | Male | 9 | 29.0% |
| **Age Group** | 18–24 | 30 | 96.8% |
| | 45–54 | 1 | 3.2% |
| **Education Level** | Bachelor's degree | 30 | 96.8% |

| | | | |
|---|---|---|---|
| | Master's degree | 1 | 3.2% |
| **Technical Background** | IT/Cybersecurity | 15 | 48.4% |
| | Some technical knowledge | 8 | 25.8% |
| | Non-technical | 8 | 25.8% |
| **Prior Expert Consultation** | Yes | 14 | 45.2% |
| | No | 17 | 54.8% |
| **Alert Frequency** | Sometimes | 13 | 41.9% |
| | Occasionally | 8 | 25.8% |
| | Often | 5 | 16.1% |
| | Never | 3 | 9.7% |
| | Always / Rarely | 2 | 6.5% |

*Note. 'Always' (n=1) and 'Rarely' (n=1) combined as 'Always / Rarely' for frequency brevity.*

Table 3: Security Systems Used by Respondents

| **Security System** | **n** | **% of Respondents** |
|---|---|---|
| Antivirus software | 22 | 71.0% |
| Browser security warnings | 20 | 64.5% |
| Email phishing alerts | 14 | 45.2% |
| Workplace security tools | 8 | 25.8% |
| None | 8 | 25.8% |

*Note. Respondents could select all applicable systems; percentages do not sum to 100%.*

As shown in Tables 2 and 3, the sample was predominantly female (71.0%), aged 18–24 (96.8%), and educated at the bachelor's degree level (96.8%). Technical backgrounds were distributed across IT/Cybersecurity (48.4%), some technical knowledge (25.8%), and non-technical (25.8%). Approximately 45.2% had previously consulted a human security expert. Antivirus software (71.0%) and browser security warnings (64.5%) were the most commonly encountered automated systems, consistent with the everyday digital environment of university students.

**Data Collection Instruments**

Data were collected using a structured, self-administered questionnaire hosted on Google Forms and comprising 23 items organized into five sections. The instrument was designed to align with established trust-in-automation scale conventions adapted to the cybersecurity advisory context (Schemmer et al., 2023; Holland et al., 2024). Section A (5 demographic items) gathered age, gender, education level, technical background, alert frequency, security systems used, and prior expert consultation experience. Sections B and C (5 Likert items each) measured Automated Alert Trust and Human Expert Trust, respectively, using a 5-point scale (1 = Strongly Disagree, 5 = Strongly Agree). Section D (5 items) elicited comparative preferences including source disagreement preferences and scenario-specific advisory choices. Section E (3 items) operationalized behavioral intention through a Likert likelihood-to-follow-alert item and two categorical preference questions. The complete instrument structure is presented in Table 4 and reproduced in full in the Appendix.

Table 4: Survey Instrument Structure

| **Section** | **Content Description** |
|---|---|
| Section A – Demographics (5 items) | Age, gender, education level, technical background, alert frequency, security systems used, prior expert consultation |
| Section B – Automated Alert Trust (5 Likert items) | Perceived timeliness, reliability, confidence in acting alone, risk-reduction capability, accuracy of threat identification |
| Section C – Human Expert Trust (5 Likert items) | Perceived accuracy, relative reliability vs. automation, confidence following expert advice, contextual understanding, judgment accuracy |
| Section D – Comparative Preferences (5 items) | Preferred source when sources disagree, perceived system consistency, ability to handle complexity, preference for quick vs. critical decisions |
| Section E – Behavioral Intentions (3 items) | Likelihood to follow software alert (Likert 1–5), preferred action when sources conflict (4 categorical options), high-risk reliance preference (3 categorical options) |

*Note. All Likert items used a 5-point scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.*

**Data Collection Procedure**

The Google Forms survey link was circulated electronically through university communication channels and peer networks during April 2026, spanning a three-day active data collection window. All potential participants were presented with a full informed consent statement at the outset of the survey, describing the study's purpose, the voluntary and anonymous nature of participation, and the right to withdraw without consequence. Only participants who affirmatively consented were permitted to proceed. No personally identifiable information was collected; responses were recorded by timestamp only. The survey required approximately five to ten minutes to complete, and no incentives were offered for participation. Upon closure of the data collection window, all responses were exported from Google Forms to a spreadsheet, cleaned by removing the one non-consent response, and prepared for analysis.


*Figure 4. Conceptual map for research methodology*

**Data Analysis Technique**

The primary analytical approach employed descriptive statistics supplemented by frequency analysis for categorical items. For each of the 15 Likert-scale items in Sections B, C, and D, the mean (M) and standard deviation (SD) were computed to characterize central tendency and dispersion. Composite trust scores were calculated by averaging the five items within the Automated Alert Trust subscale and the five items within the Human Expert Trust subscale, enabling direct numerical comparison of relative trust in each advisory

channel. For categorical behavioral intention items in Section E, frequency counts and percentages were computed for each response category. Subgroup comparisons were conducted descriptively by partitioning the sample by technical background (IT/Cybersecurity, some technical, non-technical) and by prior expert consultation experience (yes, no). All analyzes were performed in Microsoft Excel. Given the sample size (N = 31) and exploratory objectives, no inferential statistical tests (e.g., Pearson correlation, ANOVA, regression) were applied in the primary analysis, consistent with the descriptive survey design and methodological precedents established by comparable studies (MDPI, 2025).

**Delimitations of the Study**

This study is delimited by five intentional boundary decisions: (1) the target population is limited to university students and recent graduates, excluding professional SOC analysts, organizational leaders, and older adult populations addressed in other bodies of literature; (2) the cross-sectional design does not capture the longitudinal evolution of trust in response to system performance over time; (3) the self-report methodology captures stated behavioral intentions rather than directly observed security behavior; (4) convenience sampling from a single university network limits geographic, institutional, and demographic diversity; and (5) English-language administration limits participation to those with sufficient proficiency. These delimitations collectively define the study as exploratory and hypothesis-generating, intended to provide an empirical foundation for future confirmatory, longitudinal, and cross-cultural research.

## DATA ANALYSIS

This section presents results of the descriptive data analysis in four subsections: item-level descriptive statistics, subgroup comparisons by user characteristics, behavioral intention distributions, and visual representations of key findings. All analyzes are based on the valid analytical sample of N = 31 respondents.

**Descriptive Statistics**

Table 5 presents means and standard deviations for all 19 Likert-scale survey items, including computed composite scores for the Automated Alert Trust and Human Expert Trust subscales.

Table 5: Descriptive Statistics for All Survey Items

| **Survey Item / Subscale** | **M** | **SD** | **Min** | **Max** |
|---|---|---|---|---|
| **AUTOMATED ALERT TRUST SUBSCALE** | | | | |
| Automated systems provide timely warnings | 3.52 | 0.97 | 1 | 5 |
| I believe automated alerts are reliable | 3.61 | 0.88 | 1 | 5 |
| I feel confident acting on automated alerts alone | 3.42 | 0.89 | 1 | 5 |
| Automated systems reduce my risk of threats | 3.42 | 0.99 | 1 | 5 |
| I trust automated alerts to identify threats accurately | 3.52 | 0.85 | 1 | 5 |
| **Automated Trust Composite** | **3.50** | **0.73** | **1.60** | **5.00** |

| | | | | |
|---|---|---|---|---|
| **HUMAN EXPERT TRUST SUBSCALE** | | | | |
| I trust human experts to identify threats accurately | 3.29 | 1.10 | 1 | 5 |
| Human experts provide more reliable advice than automated systems | 3.39 | 0.99 | 1 | 5 |
| I feel confident following advice from a human expert | 3.26 | 1.14 | 1 | 5 |
| Human experts understand context better than automated systems | 3.48 | 1.02 | 1 | 5 |
| I believe human judgment is more accurate than automated alerts | 3.55 | 1.14 | 1 | 5 |
| **Human Expert Trust Composite** | **3.39** | **0.81** | **1.00** | **5.00** |
| **COMPARATIVE PREFERENCE ITEMS** | | | | |
| Automated systems are more consistent than human experts | 3.55 | 0.94 | 1 | 5 |
| Human experts handle complex security situations better | 3.71 | 1.02 | 1 | 5 |
| I prefer automated systems for quick decisions | 3.81 | 0.93 | 1 | 5 |
| I prefer human experts for critical decisions | 3.90 | 1.03 | 1 | 5 |
| Likelihood to follow a software security alert | 3.61 | 0.90 | 1 | 5 |

*Note. Composite scores = mean of five subscale items. SD = standard deviation. Blue rows = automated trust; green = human expert trust; amber = comparative preference items.*

The results indicate that participants expressed moderate trust in both automated security alerts (Composite M = 3.50, SD = 0.73) and human security experts (Composite M = 3.39, SD = 0.81). Both composites fall within the 'neutral to agree' range of the 5-point scale. Automated trust was marginally higher at the composite level, but the standard deviation for human expert trust is slightly larger (0.81 vs. 0.73), reflecting greater inter-individual variability — consistent with the expectation that prior expert consultation experience differentiates respondents on the human subscale. Among automated trust items, 'I believe automated alerts are reliable' recorded the highest mean (M = 3.61), reflecting broad recognition of automated systems' consistency advantage. Among human expert trust items, 'I believe human judgment is more accurate than automated alerts' recorded the highest mean (M = 3.55), reaffirming the enduring attribution of contextual superiority to human judgment. Critically, the highest mean across all 19 items was 'I prefer human experts for critical decisions' (M = 3.90, SD = 1.03) .

**Subgroup Comparison**

Table 6 presents composite trust scores disaggregated by technical background and prior expert consultation experience.

Table 6: Subgroup Comparison of Trust Composites

| Subgroup | Auto M | Auto SD | Human M | Human SD | Pattern Observed |
|---|---|---|---|---|---|
| IT/Cybersecurity background (n=15) | 3.60 | 0.52 | 3.36 | 0.80 | Higher automated trust; assessments |
| Some technical knowledge (n=8) | 3.45 | 1.00 | 3.62 | 0.76 | Slight human trust advantage |
| Non-technical (n=8) | 3.31 | 0.75 | 3.20 | 0.83 | Lowest trust in both; highest uncertainty |
| Prior expert consultation – Yes (n=14) | 3.36 | 0.75 | 3.57 | 0.76 | Expert experience raises human trust |
| Prior expert consultation – No (n=17) | 3.61 | 0.70 | 3.25 | 0.82 | No-consultation users lean toward automation |

*Note. All composite scores are means of five Likert items. Subgroup Ns: IT/Cybersecurity = 15, Some technical = 8, non-technical = 8, Prior expert Yes = 14, No = 17.*

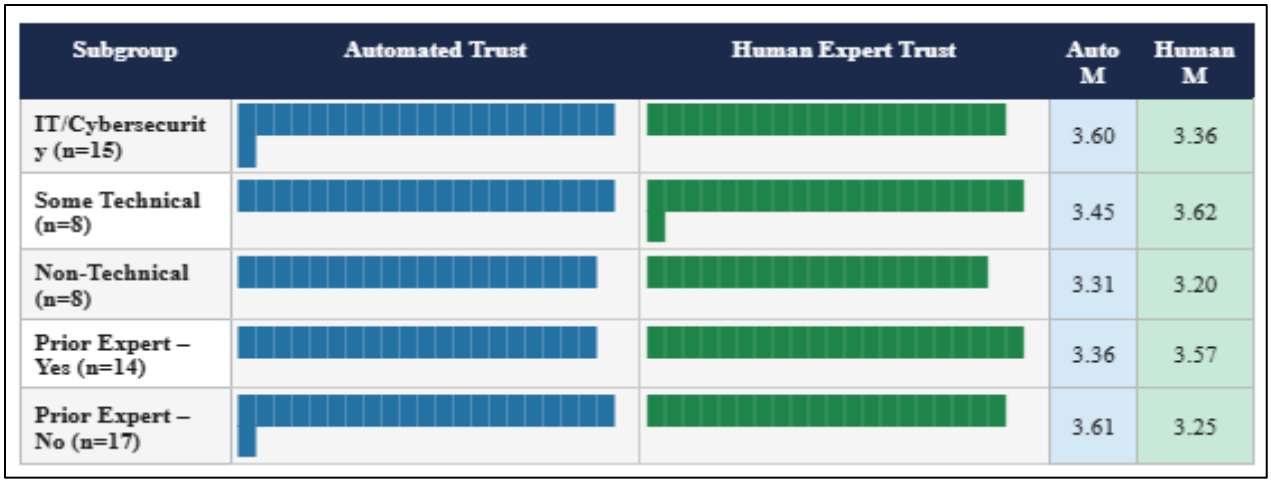


Figure 5. Comparative Results

IT/Cybersecurity background respondents showed the highest automated trust (M = 3.60, SD = 0.52) and the lowest standard deviation — indicating more calibrated, confident assessments consistent with TAM predictions for high-perceived-usefulness users. Their human expert trust was slightly lower (M = 3.36), suggesting nuanced appreciation of automation's reliability advantages. Respondents with some technical knowledge showed a reversal: slightly lower automated trust (M = 3.45) but higher human trust (M = 3.62), suggesting that partial knowledge sensitises users to automation's limitations while still valuing contextual expert guidance. Non-technical respondents exhibited the lowest scores on both subscales (Auto M = 3.31, Human M = 3.20) and the highest proportional uncertainty in conflict resolution preferences, consistent

with the algorithm aversion pattern documented by Horowitz and Kahn (2024) for low-knowledge users. Respondents with prior expert consultation experience showed significantly higher human trust (M = 3.57 vs. 3.25 for non-consulted users) and lower automated trust (M = 3.36 vs. 3.61), directly demonstrating the trust-updating effect of direct advisory experience predicted by Trust Calibration Theory (Holland et al., 2024).

**Behavioral Intention Distributions**

Table 7: Behavioral Intention Frequency Distributions (N = 31)

| **Behavioral Intention Item** | **n** | **Percentage** | **Interpretation** |
|---|---|---|---|
| **Conflict Action (When Sources Disagree)** | | | |
| Seek additional information | 24 | 77.4% | Dominant response; deliberative triangulation pattern |
| Follow the human expert | 4 | 12.9% | Second most common; human preference in conflict |
| Follow the automated alert | 2 | 6.5% | Minority; indicates high automation confidence |
| Ignore both sources | 1 | 3.2% | Rare; extreme distrust of both advisory channels |
| **High-Risk Reliance Preference** | | | |
| Both sources equally | 16 | 51.6% | Majority prefer hybrid advisory model in high stakes |
| Human experts only | 10 | 32.3% | Strong human preference as decision stakes increase |
| Automated systems only | 5 | 16.1% | Minority prefer automation even in high-risk scenarios |
| **When Sources Disagree – Trust Preference** | | | |
| Both equally / not sure (combined) | 22 | 71.0% | Majority express balanced or unresolved trust |
| Human expert | 6 | 19.4% | Human-leaning preference when forced to choose |
| Automated system | 3 | 9.7% | Automation-leaning preference when forced to choose |

*Note. 'Both equally' (n=13, 41.9%) and 'Not sure' (n=9, 29.0%) combined in the 'When sources disagree' category for narrative purposes.*

The behavioral intention data reveal a strongly deliberative pattern. When automated alerts and human expert advice conflict; 77.4% of respondents prefer to seek additional information dramatically outnumbering those who exclusively defer to either the human expert (12.9%) or the automated alert (6.5%). This distribution demonstrates that neither pure automation reliance nor unconditional human deference characterizes the behavioral reality of this population. In high-risk scenarios, 51.6% prefer to rely on both sources equally, 32.3% prefer human experts, and only 16.1% prefer automated systems — a strong directional shift toward human judgment as stakes increase, directly corroborating the risk-stakes moderation predicted by algorithm aversion theory (Dietvorst et al., 2015; Jussupow et al., 2024).


*Figure 6. Horizontal bar chart of mean scores for all 12 core trust items (bars shown as | characters) plus composites. Blue = Automated Alert Trust items; Green = Human Expert Trust items. X-axis range: 1–5. Values labeled on right.*

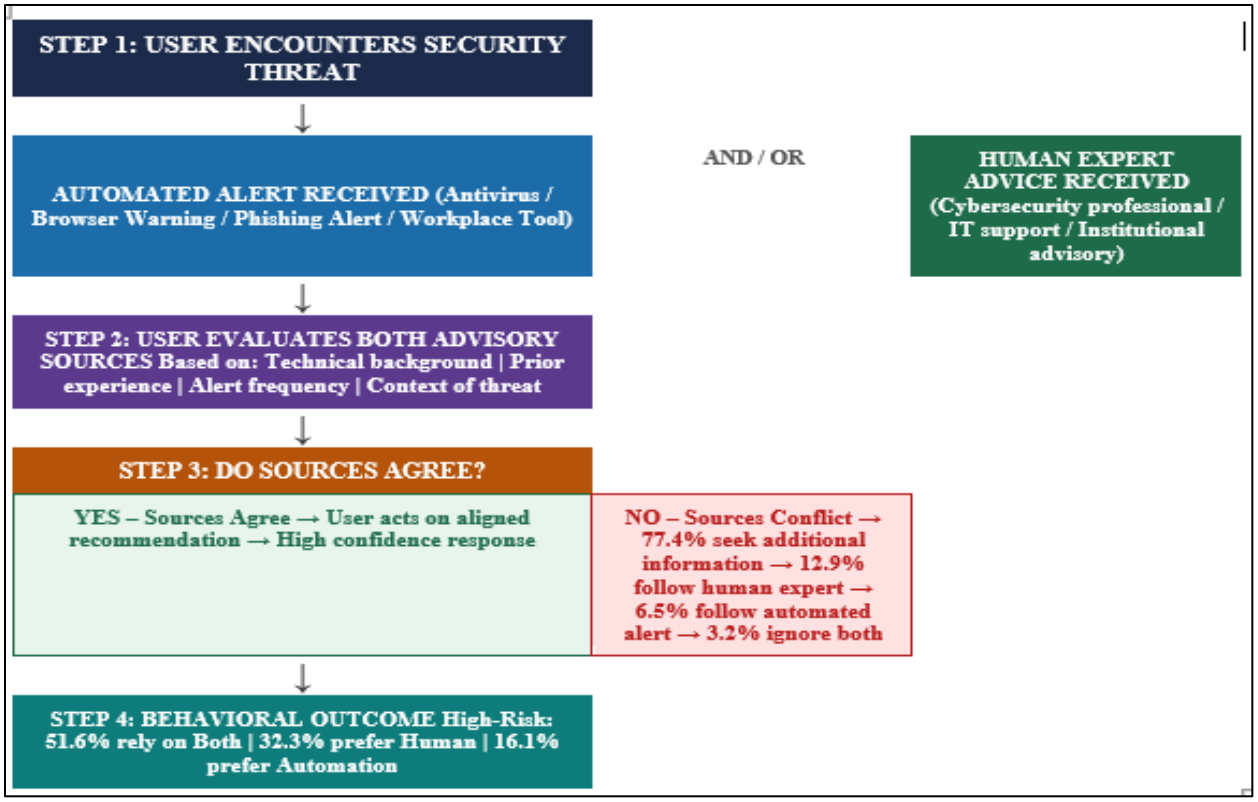


*Figure 8. Flowchart illustrating the user trust decision-making process from threat encounter through source evaluation, conflict resolution, risk assessment, and final behavioral outcome. Empirical frequencies embedded at decision nodes.*

## IMPLICATIONS OF THE STUDY

1. This study extends the Technology Acceptance Model to a comparative advisory trust context within cybersecurity, providing empirical support for treating automated alerts and human expert advice as complementary rather than competing information channels. The finding that both trust composites are moderate and closely proximate (Auto M = 3.50; Human M = 3.39) challenges simple substitution models, while the subgroup data advance Trust Calibration Theory by demonstrating that direct experience with a human security expert significantly updates trust in that channel (M = 3.57 vs. 3.25 for non-consulted users), an asymmetric updating effect not previously documented in cybersecurity research.

2. For cybersecurity system designers, the finding that 77.4% of users prefer to seek additional information when sources conflict implies that binary comply-or-dismiss security alert interfaces are fundamentally misaligned with users' natural decision-making preferences. Interfaces should instead implement graduated disclosure mechanisms — providing contextual threat information, confidence levels, and historical false-positive rates alongside alerts — to support deliberative triangulation. For university IT departments, non-technical students (25.8% of the sample), who exhibited the lowest trust in both channels and the highest decision uncertainty, require targeted training in source evaluation heuristics that enable them to engage effectively with both automated alerts and human expert advice.

3. At the institutional level, findings support the development of dual-channel security advisory policies that formally recognize automated alerts and human expert consultation as complementary and mutually reinforcing components of effective security response, rather than privileging one channel exclusively. At the national level, the evidence supports integrating human-AI collaborative decision-making training into cybersecurity certification programs (such as CompTIA Security+ and CISSP), equipping professionals not only with technical competencies but with trust calibration skills for evaluating automated outputs relative to human judgment. Additionally, the alert fatigue crisis documented at false-positive rates up to 99% in some SOC environments (ACM, 2025) calls for regulatory frameworks governing automated security product certification to incorporate maximum allowable false-positive rates as a trust-preservation standard.

## CONCLUSION

This study set out to address a meaningful gap in the cybersecurity trust literature by directly comparing end-user trust in automated security alerts versus human security experts among 31 university students and recent graduates with varying levels of technical background. Using a structured, validated 5-point Likert-scale questionnaire administered via Google Forms in April 2026, the study generated quantitative data on comparative trust dispositions, behavioral intentions under source conflict, and advisory preferences in high-risk security scenarios. The central empirical finding is that users occupy a deliberative middle ground between pure automation reliance and unconditional human deference. Automated alert trust (Composite M = 3.50, SD = 0.73) and human expert trust (Composite M = 3.39, SD = 0.81) were both moderate and closely proximate, indicating that respondents neither exclusively valorize nor systematically distrust either advisory channel. When sources conflict, 77.4% prefer to seek additional information — a behavioral pattern characterized in this study as deliberative triangulation and representing a theoretically distinct epistemic position that warrants incorporation into future trust models. In high-risk scenarios, 51.6% prefer both sources equally and 32.3% prefer human experts, confirming the risk-stakes moderation of human preference predicted by algorithm aversion theory. Subgroup analysis revealed that IT/Cybersecurity background users show higher, more calibrated automated trust; prior expert consultation experience significantly raises human trust; and non-technical users exhibit the lowest scores on both subscales and the highest decision uncertainty. These findings collectively support a user-centered, multi-source approach to cybersecurity advisory design and policy. Future research should extend this work through larger probability samples enabling inferential analysis, longitudinal designs capturing trust evolution over time, and experimental studies testing whether deliberative triangulation interventions improve security decision quality relative to single-channel advisory designs.

## RECOMMENDATIONS

Based on the findings of this study, three categories of recommendations are proposed for system designers, educational institutions, and future researchers. Cybersecurity user interface designers should redesign security alert systems to support deliberative triangulation rather than binary comply-or-dismiss responses. Practically, this means implementing graduated disclosure mechanisms that provide contextual information — threat category, system confidence level, historical false-positive rates for the alert type, and links to human expert commentary — alongside automated alerts, enabling users to make informed judgments rather than reflexive compliance decisions. University IT security departments should implement tiered, technically differentiated security awareness programs: for non-technical students, training should focus on practical source evaluation heuristics; for IT/Cybersecurity-background students, on trust calibration and the conditions under which automation bias is most likely; and for all students, on structured exposure to human security expert consultation as a trust-building experience with demonstrated benefits for subsequent decision quality (M = 3.57 for consulted vs. 3.25 for non-consulted users).

Future research should prioritize four directions: (1) replication with larger probability samples from multiple institutions enabling robust inferential statistical analysis (Pearson correlation, ANOVA, and regression) of the IV-DV-OV relationships identified in this study's conceptual framework; (2) longitudinal designs that track how trust in automated alerts and human experts evolves in response to system performance experience, security incidents, and education interventions over academic years or employment cycles; (3) experimental studies testing whether deliberative triangulation interventions such as graduated alert disclosure interfaces or structured expert consultation protocols improve measurable security decision quality relative to single-channel advisory designs; and (4) cross-cultural replication studies examining whether the trust patterns documented here are culturally universal or reflect culturally specific dispositions toward algorithmic authority and institutional trust.

**REFERENCES**

ACM Computing Surveys. (2025). Alert fatigue in security operations centres: Research challenges and opportunities. https://dl.acm.org/doi/10.1145/3723158

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Dietvorst, B. J., Logg, J. M., & Logg, J. M. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126. https://doi.org/10.1037/xge0000033

Ebert, N., Ackermann, K. A., & Bearth, A. (2022). When information security depends on font size: How the saliency of warnings affects protection behavior. Journal of Risk Research. https://doi.org/10.1080/13669877.2022.2142952

HAT-Lab Research Group. (2025). Are you sure? An empirical study of human perception vulnerability in LLM-driven agentic systems. ArXiv. https://arxiv.org/pdf/2602.21127

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. Human Factors, 57(3), 407–434. https://doi.org/10.1177/0018720814547570

Holland, C., Perry, G., & Neyedli, H. F. (2024). Calibrating trust, reliance and dependence in variable-reliability automation. Human Factors. https://doi.org/10.1177/10711813241277531

Horowitz, M. C., & Kahn, L. (2024). Bending the automation bias curve: A study of human and AI-based decision making in national security contexts. International Studies Quarterly, 68(2). https://doi.org/10.1093/isq/sqae020

IJCTECE. (2025). Human-AI collaboration in security operations: Measuring alert trust, automation bias, and analyst upskilling in AI-augmented SOC environments. International Journal of Computer Technology and Electronics Communication. https://ijctece.com/index.php/IJCTEC/article/view/233

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2024). An integrative perspective on algorithm aversion and appreciation in decision-making. MIS Quarterly, 48(4). https://doi.org/10.25300/misq/2024/17702

KPMG. (2024). KPMG 2024 cybersecurity survey. https://kpmg.com/us/en/media/news/2024-cybersecurity-survey.html

Keepnet Labs. (2026). 2025 phishing statistics (Updated January 2026). https://keepnetlabs.com/blog/top-phishing-statistics-and-trends-you-must-know

Korka, D., Salehzadeh Niksirat, K., & Cherubini, M. (2023). Revisiting the design agenda for privacy notices and security warnings. CHI 2023 Workshop on Privacy Interventions and Education. https://arxiv.org/pdf/2304.08780

Kueper, A., Lodde, G., Knopp, M., Mueller-Birn, C., & Smeddinck, J. (2025). Exploring automation bias in human-AI collaboration: A review and implications for explainable AI. AI & Society. https://doi.org/10.1007/s00146-025-02422-7

Lim, B., et al. (2025). Too much to trust? Measuring the security and cognitive impacts of explainability in AI-driven SOCs. ArXiv. https://arxiv.org/pdf/2503.02065

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

MDPI. (2025). AI-based phishing detection and student cybersecurity awareness in the digital age. Big Data and Cognitive Computing, 9(8), 210. https://doi.org/10.3390/bdcc9080210

Proofpoint. (2024). 2024 state of the phish report. https://www.proofpoint.com/us/blog/security-awareness-training/2024-state-of-phish-report

Schemmer, M., Kuehl, N., Benz, C., Bartos, A., & Satzger, G. (2023). Measuring and understanding trust calibrations for automated systems. Proceedings of the 28th ACM CHI Conference on Human Factors in Computing Systems. https://dl.acm.org/doi/full/10.1145/3544548.3581197

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? International Journal of Human-Computer Studies, 51(5), 991–1006. https://doi.org/10.1006/ijhc.1999.0252

Springer Nature. (2025). User acceptance of AI-powered training: Extending the technology acceptance model (TAM). Future Business Journal. https://doi.org/10.1186/s43093-025-00665-w

ThreatQuotient. (2023). The state of cybersecurity automation adoption 2023. https://techhq.com/2023/11/cybersecurity-automation-survey-results-released-what-does-it-say/

Xie, L., & Zhou, W. A. (2025). Trust under threat: How AI vs. human mistakes in disaster alerts shape public perception and response. SSRN Working Paper. https://doi.org/10.2139/ssrn.5227870

**Appendix**

**Survey Questionnaire**

**Comparative Study of Trust in Automated Security Alerts vs. Human Security Experts**

*Survey Instrument — Google Forms | April 2026 | Air University Multan Campus*

**Informed Consent Statement:** I have read the information above and agree to participate in this research study. I understand that my participation is voluntary and anonymous, and that I may withdraw at any time without consequence.

| # | Item | Response Options | Type |
|---|---|---|---|
| | ***SECTION A: DEMOGRAPHICS*** | | |
| 1 | Have you read and agree to participate? | Yes / No | |
| 2 | What is your age? | 18–24 / 25–34 / 35–44 / 45–54 / 55+ | |
| 3 | What is your gender? | Male / Female / Non-binary / Prefer not to say | |
| 4 | What is your highest education level? | High school / Bachelor's / Master's / PhD / Other | |
| 5 | What is your technical background? | IT/Cybersecurity / Some technical knowledge / Non-technical | |
| ***SECTION B: AUTOMATED ALERT TRUST (1=Strongly Disagree, 5=Strongly Agree)*** | | | |
| 6 | Automated systems provide timely warnings. | 1 2 3 4 5 | Likert |
| 7 | I believe automated alerts are reliable. | 1 2 3 4 5 | Likert |
| 8 | I feel confident acting based on automated alerts alone. | 1 2 3 4 5 | Likert |
| 9 | Automated systems reduce my risk of security threats. | 1 2 3 4 5 | Likert |
| 10 | I trust automated security alerts to identify threats accurately. | 1 2 3 4 5 | Likert |
| ***SECTION C: HUMAN EXPERT TRUST (1=Strongly Disagree, 5=Strongly Agree)*** | | | |

| | | | |
|---|---|---|---|
| 11 | I trust human security experts to identify threats accurately. | 1 2 3 4 5 | Likert |
| 12 | Human experts provide more reliable advice than automated systems. | 1 2 3 4 5 | Likert |
| 13 | I feel confident following advice from a human security expert. | 1 2 3 4 5 | Likert |
| 14 | Human experts understand context better than automated systems. | 1 2 3 4 5 | Likert |
| 15 | I believe human judgment is more accurate than automated alerts. | 1 2 3 4 5 | Likert |
| ***SECTION D: COMPARATIVE PREFERENCES*** | | | |
| 16 | When automated systems and human experts disagree, I trust: | Automated system / Human expert / Both equally / Not sure | Cat. |
| 17 | Automated systems are more consistent than human experts. | 1 2 3 4 5 | Likert |
| 18 | Human experts are better at handling complex security situations. | 1 2 3 4 5 | Likert |
| 19 | I prefer automated systems for quick security decisions. | 1 2 3 4 5 | Likert |
| 20 | I prefer human experts for critical security decisions. | 1 2 3 4 5 | Likert |
| ***SECTION E: BEHAVIORAL INTENTIONS*** | | | |
| 21 | If you receive a security alert from software, how likely are you to follow it? | 1 2 3 4 5 | Likert |

| | | | |
|---|---|---|---|
| 22 | If a human expert advises differently than an automated alert, what would you do? | Follow alert / Follow expert / Seek additional info / Ignore both | Cat. |
| 23 | In high-risk situations, I would rely more on: | Automated systems / Human experts / Both equally | Cat. |