Cybersecurity is Stressful: The Impact of Stress on Identifying Phishing Attacks

Christian Bergh
Computing & Software Systems
University of Washington
Bothell, USA
cbergh@uw.edu

Marc J. Dupuis
Computing & Software Systems
University of Washington
Bothell, USA
marcjd@uw.edu

***Abstract*— Phishing remains one of the most pervasive cybersecurity threats, causing billions in annual damages. Despite advances in technical defenses, attackers exploit the human element, often leveraging urgency, authority, and fear to provoke hurried responses. While stress is frequently cited as a factor in phishing susceptibility, empirical evidence has been limited. This study experimentally examined the relationship between stress and phishing detection. Using an online adaptation of the Trier Social Stress Test (TSST), 150 participants recruited via Amazon Mechanical Turk (MTurk) completed a phishing identification task under time pressure, evaluative observation cues, and monetary incentives. Stress levels were assessed pre- and post-task using the State-Trait Anxiety Inventory (STAI). Although overall stress levels remained stable across the full sample, a subset of participants exhibited measurable increases. Within this subgroup, higher stress was significantly correlated with reduced phishing detection accuracy. These results suggest that acute stress impairs cognitive resources required for careful evaluation, increasing vulnerability to phishing. The study makes three contributions: (1) providing empirical evidence that acute stress can compromise phishing detection, (2) demonstrating the feasibility and challenges of applying controlled stress-induction methods in online experiments, and (3) raising critical implications for organizations that rely on employees as a frontline defense. The findings highlight the importance of stress-aware security training, workplace reforms to reduce cognitive burden, and future research integrating physiological and behavioral measures to better understand the role of stress in cybersecurity.** (*Abstract*)

Keywords—stress, cybersecurity, information security, emotional health, phishing attack (*key words*)

I. Introduction

Cybersecurity has become one of the defining challenges of the 21st century. From government espionage to corporate data breaches and identity theft, the scope and scale of cyber threats are vast. At the center of many of these attacks lies phishing, a social engineering strategy that manipulates human psychology rather than exploiting technical vulnerabilities [1], [2]. Phishing emails, text messages, and phone calls succeed not because of advanced code but because they convince individuals to click a link, download an attachment, or disclose personal credentials. As a result, humans are consistently described as the "weakest link" in the security chain [3].

The ubiquity of phishing is well-documented. The Anti-Phishing Working Group (APWG) has reported consistent year-over-year increases in phishing attempts, with campaigns targeting not only large corporations but also small businesses and individuals. During the COVID-19 pandemic, phishing exploited global uncertainty, leveraging themes such as vaccine registration, contact tracing, and government relief funds [4]. These campaigns reveal how attackers adapt to situational stressors and exploit human vulnerabilities in times of crisis. Phishing has also evolved in sophistication, with increasingly professional design, contextual tailoring, and integration of real-world events, making detection more challenging than ever before.

Stress itself has been the subject of extensive psychological and neurobiological research. It is well-established that stress impairs working memory, reduces executive function, and biases individuals toward habitual rather than deliberative decision-making [5], [6]. Acute stress activates the hypothalamic-pituitary-adrenal (HPA) axis, flooding the body with cortisol, which in turn disrupts prefrontal cortex activity. In cybersecurity contexts, these impairments could translate into misclassifying emails, overlooking subtle phishing cues, or failing to pause before clicking. Stress also shapes attentional resources: individuals under stress may fixate on certain aspects of a message, such as bolded deadlines or warnings, while ignoring subtle discrepancies in sender addresses or URLs.

What makes stress particularly insidious in phishing is its dual presence. First, many workplaces are already stress-inducing environments, with employees facing deadlines, multitasking demands, and high expectations for productivity. Second, attackers deliberately craft phishing messages to introduce additional stress, using urgency ("Your account will be locked in 24 hours"), authority ("CEO request"), and fear ("Suspicious login detected") as cues. These tactics exploit the very cognitive vulnerabilities exacerbated by stress. Importantly, stress may not always operate in a linear manner: moderate stress can sometimes improve performance on simple or routine tasks but tends to impair complex decision-making [7]. Phishing detection, which requires careful scrutiny, likely falls into the latter category.

Despite the clear theoretical overlap between stress and phishing susceptibility, empirical research directly linking the

two is limited. Studies on stress in cybersecurity often rely on self-report surveys or focus on compliance behaviors rather than phishing detection [4], [8]. For instance, research on "security fatigue" suggests that overwhelmed individuals are more likely to disregard warnings, even when they recognize the risks [9]. And while individuals may learn from their mistakes by experiencing regret, minimizing the conditions that cause thos mistakes in the first place is essential [10]. Other work has noted the role of occupational stress in shaping broader organizational security culture but stops short of experimentally testing its impact on phishing detection. Thus, while scholars have long speculated about the relationship between stress and phishing, rigorous evidence remains scarce.

This paper addresses that gap by conducting a controlled experiment that measures stress levels before and after a phishing detection task. We hypothesize that increased stress will negatively correlate with phishing detection accuracy. By leveraging an online adaptation of the Trier Social Stress Test (TSST), we sought to replicate the cognitive and emotional strain of real-world stressors while observing its impact on participants' ability to detect phishing attempts.

Our contributions are both empirical and practical. Empirically, we demonstrate that stress can, under certain conditions, impair phishing detection. Practically, we highlight the importance of designing training, policies, and workplace systems that account for stress as a critical variable. These contributions align with a growing recognition in the field of human factors in cybersecurity that cognitive, emotional, and social contexts significantly shape security outcomes [11].

The remainder of this paper is structured as follows. Section 2 reviews related work on phishing, stress, and human factors. Section 3 outlines the methodology, including participants, instruments, and procedures. Section 4 presents the results of stress measurements and phishing accuracy analyses. Section 5 discusses implications and limitations. Section 6 concludes with recommendations for research and practice.

II. Related Work

Research on phishing has consistently emphasized that psychological manipulation, rather than technical sophistication, drives the success of attacks. Although filters, authentication systems, and technical countermeasures have improved over the years, attackers continue to bypass these defenses by targeting human vulnerabilities. Understanding the interplay between psychological tactics and cognitive conditions such as stress is essential for developing holistic security strategies.

A. Phishing Tactics and Psychological Manipulation

Early studies on phishing highlighted how easily users can be deceived by superficially convincing email designs. Dhamija, Tygar, and Hearst [12] demonstrated that even experienced users often misjudge the legitimacy of websites and emails, particularly when visual elements such as logos, headers, and security indicators are manipulated. Jagatic et al. [1] advanced this understanding by showing how contextual cues from social networks dramatically increase success rates in so-called spear-phishing attacks. The personalization of attacks exploits pre-existing trust relationships, thereby reducing skepticism and scrutiny.

These findings have been reinforced by broader research on persuasion and influence. Cialdini [3] identified six principles of persuasion—authority, reciprocity, scarcity, commitment, liking, and social proof—that explain much of human decision-making in uncertain contexts. Phishing emails frequently draw on these principles. Authority is leveraged when attackers impersonate senior executives or trusted institutions. Scarcity and urgency are invoked through statements such as "immediate action required" or "your account will expire within 24 hours." Social proof is embedded when attackers reference widely known brands or suggest that "others in your department have already updated their credentials." Collectively, these cues shape behavior in ways that are difficult for even knowledgeable users to resist when cognitive resources are strained.

Workman [2] extended this line of research by exploring how humor, narrative framing, or other emotional hooks can lower suspicion. His findings suggest that phishing effectiveness is not limited to fear or urgency; rather, any tactic that shifts attention away from deliberate scrutiny can increase vulnerability. These insights underscore the importance of considering emotional and cognitive context, not just message design, in understanding phishing susceptibility.

B. Training and Awareness Efforts

Given the persistence of phishing, a significant body of research has examined educational and training interventions. Kumaraguru et al. [13] evaluated anti-phishing education programs, finding that while short-term improvements in detection were possible, retention and transfer of learning were often limited. This mirrors broader findings in the information security awareness literature, where one-time trainings produce temporary gains but struggle to sustain behavioral change over time.

Gamified approaches have been proposed to increase engagement. Sheng et al. [14] developed "Anti-Phishing Phil," an interactive game designed to teach users how to identify phishing URLs. Results showed promise in improving detection skills, though critics have noted that such interventions may oversimplify the problem by focusing narrowly on technical cues. More recent approaches have suggested embedding phishing simulations into organizational culture, where periodic simulated phishing emails both test and train employees. While effective in raising awareness, such programs sometimes create adversarial relationships between employees and security teams, raising ethical and cultural concerns.

Importantly, none of these approaches fully account for the role of stress. Training often assumes a rational, attentive user who is motivated to apply learned strategies consistently. In reality, employees operate under time pressure, divided attention, and varying emotional states. Training that does not account for these contextual realities may therefore overestimate its effectiveness.

C. Stress and Cognitive Performance

Parallel to phishing research, the psychology and neuroscience literature provide extensive evidence on the impact of stress on cognition. McEwen [6] describes how both acute and chronic stressors affect the brain, particularly the hippocampus and prefrontal cortex, which are critical for memory and executive function. Under stress, individuals experience reduced working memory capacity and impaired attention control. Arnsten [5] adds that acute stress disrupts the neural circuits responsible for higher-order thinking, biasing individuals toward rapid, habitual responses rather than careful deliberation.

The Yerkes–Dodson law [7] provides a useful heuristic for understanding these dynamics. It suggests that while moderate arousal can enhance performance on simple tasks, performance on complex tasks deteriorates as stress increases. Phishing detection arguably falls into the latter category: distinguishing subtle differences between legitimate and malicious messages requires sustained attention, working memory, and critical evaluation—precisely the capacities most impaired under stress.

Schwabe and Wolf [15] provided experimental evidence for this shift by showing that stressed individuals are more likely to rely on habitual responses, even when goal-directed strategies would yield better outcomes. Applied to phishing, this implies that stressed employees may default to automatic behaviors such as clicking links or following apparent instructions rather than scrutinizing details.

D. Stress in Cybersecurity Contexts

Despite the theoretical connections, direct empirical research linking stress to cybersecurity behavior remains limited. Vance, Siponen, and Pahnila [4] examined how stress and habit interact to shape security compliance behaviors. They found that when users are stressed, they are more likely to rely on established habits, which may or may not align with secure practices. This finding suggests that stress can amplify existing behavioral tendencies, including insecure ones.

Henshel et al. [8] studied healthcare workers to explore the relationship between occupational stress and security behaviors. Their findings were inconclusive, reflecting the difficulty of isolating stress as a variable in complex organizational settings. However, they noted that workers experiencing high levels of stress often reported ignoring or bypassing security procedures in order to meet immediate work demands.

Johnston, Warkentin, and Siponen [16] advanced this line of inquiry by connecting stress to "security fatigue," a phenomenon in which users, overwhelmed by repeated security demands, become apathetic toward compliance. Their work highlights the cumulative impact of stress over time, which may erode vigilance and increase vulnerability to attacks.

E. Gaps in the Literature

The literature reviewed above suggests several important themes. First, phishing success is strongly linked to psychological manipulation, and stress is a recurring element in the tactics attackers deploy. Second, stress has well-documented effects on cognition, particularly in reducing the very faculties needed for phishing detection. Third, while studies have begun to explore stress in relation to compliance and fatigue, few have experimentally tested the direct relationship between acute stress and phishing detection.

This gap is significant for both theory and practice. From a theoretical perspective, demonstrating the causal role of stress would advance our understanding of how psychological states interact with cybersecurity behaviors. From a practical perspective, it would inform training and policy design, emphasizing the need to prepare users not just cognitively but also emotionally for security challenges.

The present study builds on these insights by experimentally inducing stress in an online environment and measuring its effect on phishing detection accuracy. By adapting the Trier Social Stress Test (TSST) to a cybersecurity context, we attempt to provide controlled, empirical evidence for the hypothesized relationship. This represents a novel contribution to the human factors literature on cybersecurity and an important step toward designing interventions that address the full complexity of human vulnerability.

III. Methods

A. Participants and Recruitment

This study recruited 150 participants using Amazon Mechanical Turk's (MTurk) Masters Worker pool. MTurk was chosen for its scalability, participant diversity, and demonstrated reliability in behavioral research [17]. Workers in the Masters pool possess a record of high-quality performance across tasks, reducing concerns about inattentive or automated responses. Eligibility criteria required participants to be at least 18 years of age, fluent in English, and to have an approval rating of 95% or higher across previous MTurk tasks. These criteria ensured that participants could reliably comprehend study instructions and that the dataset would meet quality expectations.

Although MTurk offers significant advantages in terms of cost-effectiveness and efficiency, it is not without limitations. Prior research has documented concerns about ecological validity and sample representativeness [18]. MTurk participants often work under unique motivational and environmental conditions, including multitasking, background stressors, and variable attention. While these factors may limit generalizability, they also provide an ecologically relevant context, as many workplace employees likewise navigate competing demands when encountering phishing attempts. To mitigate risks of low-quality data, we implemented attention checks, monitored completion times, and excluded incomplete or failed submissions, which have been shown to help mitigate many of these issues [19], [20].

B. Experimental Design

The experimental design was informed by the Trier Social Stress Test [21], a well-validated laboratory procedure for inducing acute stress. The original TSST involves participants performing speech and mental arithmetic tasks in front of a panel of evaluators, a situation designed to elicit social-evaluative threat and uncontrollability, two hallmarks of stress

induction. Given the online format of this study, a direct replication was infeasible. Instead, we adapted the TSST principles into a phishing detection context through the following mechanisms:

1. **Time Pressure:** Each email stimulus was displayed with a strict 20-second time limit, after which the system automatically advanced. This created a sense of urgency and reduced deliberation time, mimicking real-world workplace conditions where employees face constant interruptions and competing demands.
2. **Performance Incentives:** Participants were told that high accuracy would result in bonus compensation, framing the task as evaluative and consequential. Incentive structures were used to heighten motivation and perceived stakes, both of which have been shown to increase stress levels in cognitive tasks.
3. **Observation Cues:** Instructions informed participants that their responses were monitored for quality control and research integrity. Although no live observers were present, the suggestion of evaluation was intended to replicate the TSST's social-evaluative threat.

Together, these stressors were designed to approximate the psychological ingredients of the TSST while aligning them with the phishing detection task.

C. Materials and Stimuli

The phishing detection task consisted of 25 email screenshots, presented one at a time. Of these, 13 were benign and 12 were phishing. Emails were drawn from real-world datasets, including publicly available phishing repositories and legitimate corporate communication archives. To ensure realism, messages were updated to reflect current contexts, including service provider notifications, workplace directives, and financial account alerts. Care was taken to balance difficulty across stimuli: some phishing emails were intentionally obvious (e.g., spelling errors, implausible sender domains), while others were more sophisticated (e.g., spoofed login requests with subtle URL discrepancies). An example of one such email may be found in Figure 1.

Each email required a binary judgment: participants were asked, "Is this email phishing or legitimate?" Responses were recorded along with reaction times. Accuracy was calculated as the percentage of correct classifications across all items.

D. Stress Measurement

Stress levels were measured using the State-Trait Anxiety Inventory [22]. The STAI distinguishes between state anxiety (temporary, situational stress) and trait anxiety (enduring personality characteristics). In this study, only the state anxiety subscale was used, as it is sensitive to acute changes in stress levels.

Participants completed the STAI immediately before and immediately after the phishing detection task. Each item asked respondents to rate their current feelings (e.g., "I feel tense," "I feel calm") on a 4-point Likert scale. Scores range from 20 to 80, with higher scores indicating greater anxiety. Pre-task and post-task scores were used to calculate delta values, reflecting the change in stress levels attributable to the experiment.

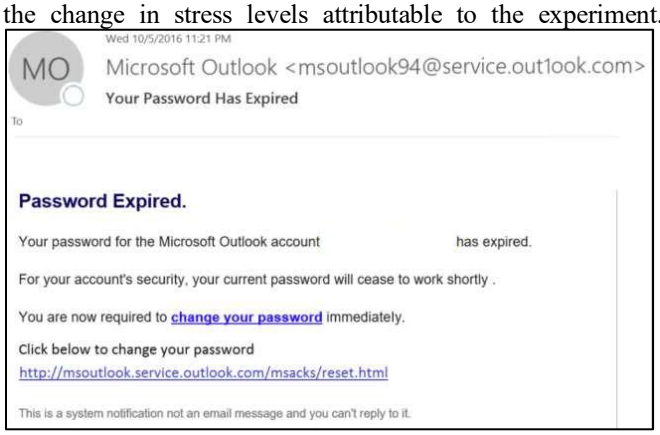

Fig. 1. Example Phishing/Not Phishing Email for Participants

The STAI has been widely validated across laboratory and field settings and is particularly useful in online experiments where physiological measures (e.g., cortisol, galvanic skin response) are impractical. Although self-report measures have limitations, the STAI's sensitivity to short-term fluctuations made it well-suited for this study's objectives.

E. Procedure

The study proceeded in the following sequence:

1. **Consent and Instructions:** Participants first reviewed an informed consent form, which outlined the purpose, procedures, risks, and compensation. They were informed that the task involved identifying phishing emails under time pressure and that performance would influence bonus payments.
2. **Baseline Stress Assessment:** Participants completed the pre-task STAI to measure baseline stress.
3. **Phishing Task:** Participants classified the 25 email stimuli under the 20-second time constraint. They were reminded of accuracy bonuses and monitoring to maintain evaluative stressors.
4. **Post-task Stress Assessment:** Participants completed the post-task STAI to measure stress levels after the phishing detection activity.
5. **Debriefing:** Participants were debriefed about the purpose of the study, reassured that monitoring was not literal, and informed about stress management resources.

F. Data Analysis

Data were analyzed using SPSS 30.0. Descriptive statistics summarized demographic variables, STAI scores, and phishing detection accuracy. Paired-sample t-tests compared pre-task and post-task STAI scores to assess overall stress changes. Pearson correlation coefficients were calculated to test the relationship between stress deltas and accuracy.

Subgroup analyses were conducted by splitting participants into two categories: those whose stress increased (positive delta) and those whose stress decreased or remained constant

(zero/negative delta). This allowed examination of whether stress effects were conditional on actual increases.

Demographic variables (age, gender, education level) were tested as potential moderators using ANOVA and regression analyses. Reaction times were also examined to assess whether stress influenced decision speed.

G. Ethical Considerations

The study protocol was reviewed and approved by the Institutional Review Board (IRB). Ethical considerations included ensuring informed consent, minimizing undue stress, and providing debriefing with stress management resources. Compensation included a base payment consistent with fair hourly wages for MTurk tasks, with bonuses tied to accuracy. Data were anonymized to protect participant confidentiality.

Although the stress induction was mild relative to in-person TSST protocols, safeguards were in place to allow participants to withdraw at any time without penalty. No participants reported adverse reactions during or after the study.

H. Limitations of Methodology

Several methodological limitations are worth noting. First, the online adaptation of the TSST may not fully replicate the stress-inducing power of in-person protocols. The absence of live observers likely reduced the intensity of social-evaluative threat. Second, reliance on self-reported stress via the STAI may underestimate physiological changes. Future work could benefit from integrating wearable devices or saliva-based cortisol measures to capture more objective indicators of stress.

Third, the MTurk sample, while diverse, may not fully represent organizational employees who face phishing in workplace settings. MTurk participants are often younger, more technologically savvy, and habituated to stressors unique to online piecework. Nonetheless, the high baseline stress levels observed in this study suggest that MTurk remains a useful proxy for testing stress-security interactions.

Finally, the binary task design, while ecologically relevant, may oversimplify real-world phishing detection, where users can ignore, report, or seek advice about suspicious messages. Future research should consider more nuanced task designs that capture these alternative responses. Nonetheless, it is worth noting that live phishing experiments also pose their own significant challenges, both logically and ethically [23].

IV. Results

A. Descriptive Statistics

Participants entered the study with elevated baseline stress levels. Pre-task State-Trait Anxiety Inventory (STAI) scores ranged from 22 to 65, with a mean of 44.71 (SD = 10.9). Importantly, 83 percent of participants scored above the clinical cutoff of 40, indicating that most were already experiencing stress levels higher than typical non-clinical samples. Post-task scores ranged from 22 to 63, with a mean of 44.39 (SD = 11.2). The average change in stress (delta) was –0.31 points, suggesting minimal net movement across the full sample. Approximately 53 percent of participants reported stable or increased stress, while 47 percent reported decreases.

From an accuracy perspective, phishing detection performance averaged 75 percent correct (SD = 13.5), spanning a wide distribution from 36 percent to 100 percent. The distribution was approximately normal, with no evidence of extreme ceiling or floor effects. Accuracy varied across items: several phishing attempts were correctly identified by more than 95 percent of participants (e.g., poorly worded "Nigerian prince"–style scams), while others were correctly identified by fewer than half of participants (e.g., sophisticated credential harvesting messages that closely mimicked legitimate service provider notifications).

Reaction times also displayed meaningful variation. On average, participants responded within 11.2 seconds (SD = 3.8). However, early items exhibited longer latencies, suggesting initial caution before participants adapted to the time-limited format. Later responses clustered closer to the 20-second cutoff, with 13 percent of responses across the entire task being auto-submitted when participants failed to respond in time. This figure decreased as the task progressed, implying acclimation to time pressure.

B. Stress and Accuracy Correlations

To test the central hypothesis, Pearson correlation coefficients were computed between stress delta values and phishing detection accuracy. Across the full sample, no significant correlation was found ($r = –.09$, $p = .215$). This suggests that, at the aggregate level, changes in stress did not predict detection performance.

However, aggregate null results may obscure subgroup effects. Indeed, a closer look at participants with stress increases versus those with stress decreases revealed different patterns.

C. Subgroup Analyses

Among participants whose stress increased during the task ($n = 79$), stress delta values negatively correlated with accuracy ($r = –.28$, $p = .011$). This indicates that higher increases in stress were associated with lower phishing detection accuracy, supporting the hypothesis for this subgroup. Conversely, participants whose stress decreased or remained stable ($n = 71$) showed no significant relationship ($r = .04$, $p = .61$).

A regression model including stress delta, age, gender, and education as predictors of accuracy confirmed these results. Stress delta was a significant predictor only in the stress-increase subgroup. Demographic variables were non-significant, though exploratory trends suggested that younger participants exhibited slightly less accuracy decline under stress than older participants.

D. Item-Level Performance

Performance varied significantly across individual email items. Obvious phishing attempts, such as poorly worded requests for wire transfers, were almost universally identified. In contrast, sophisticated phishing messages with professional formatting and legitimate-seeming domains proved challenging, with accuracy rates dipping below 50 percent. Interestingly, items that blended authority and urgency cues (e.g., "CEO request for password reset" or "Suspicious login attempt, act immediately") were among the lowest performing, suggesting

that stress-inducing content may interact with stress levels in particularly detrimental ways.

A post hoc analysis explored whether participants under higher stress deltas were disproportionately misclassifying these authority/urgency emails. Results suggested that stressed participants were more likely to misidentify such items as legitimate, though this effect did not reach statistical significance after correcting for multiple comparisons.

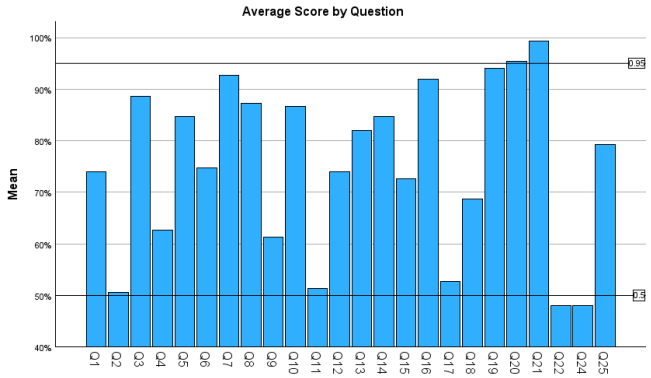

Fig. 2. Average Score by Question

E. Reaction Times and Stress

To assess whether stress influenced decision speed, mean reaction times were compared across participants with increased versus decreased stress. Those with increased stress responded significantly faster on average ($M = 10.3$s) than those with decreased/stable stress ($M = 12.2$s; $t(148) = 2.12$, $p < .05$). Faster responding, however, correlated with lower accuracy ($r = –.24$, $p = .018$), suggesting that stress-induced haste contributed to performance deficits.

This finding aligns with theoretical models predicting that stress biases individuals toward rapid, habitual responses rather than deliberate analysis (Arnsten, 2009; Schwabe & Wolf, 2009). Importantly, the speed–accuracy tradeoff was most pronounced on the more subtle phishing messages, where careful scrutiny was required.

F. Summary of Findings

To summarize, the results provide mixed but meaningful support for the hypothesis. While stress deltas did not predict accuracy across the entire sample, subgroup analysis revealed that participants who experienced stress increases performed worse on phishing detection. The relationship was moderate in size and statistically significant. Reaction time analyses suggest that stress increases may push participants toward quicker, less accurate decisions. Item-level analyses further suggest that authority and urgency cues may be particularly effective against stressed individuals.

Collectively, these results indicate that stress is a conditional but potent factor in phishing susceptibility. Stress alone may not universally reduce performance, but when stressors successfully elevate anxiety, users' accuracy in detecting phishing attempts declines.

V. Discussion

A. Integrating Neuroscience, Psychology, and Security

The results of this study provide empirical evidence supporting the claim that stress can impair phishing detection. Participants whose stress increased during the task performed significantly worse, and their faster reaction times were associated with lower accuracy — a classic speed–accuracy trade-off driven by stress. These findings map closely onto established neurobiological models: acute stress activates the hypothalamic-pituitary-adrenal (HPA) axis, elevating cortisol levels and impairing prefrontal cortex function, which is critical for working memory, attention regulation, and deliberative reasoning [5], [6].

The observed performance deficit aligns with the Yerkes–Dodson law: heightened arousal impairs complex tasks, which phishing detection clearly represents. The fact that authority and urgency cues further degraded performance in stressed individuals supports models of psychological manipulation grounded in Cialdini's principles of persuasion [3]. Attackers intentionally exploit the cognitive shortcuts that stressed individuals default to — a phenomenon also documented in phishing studies [1], [2].

B. Stress, Digital Deception, and Human Factors

This study situates itself within a broader body of work emphasizing the role of human psychology in cybersecurity. Prior research has established that phishing effectiveness is tied less to technical sophistication than to human decision-making under pressure [1], [12]. Similarly, studies of stress demonstrate that acute anxiety biases individuals toward habitual, automatic responses rather than reflective analysis [15]. By showing that stress increases correlate with reduced phishing detection accuracy, the present study links these literatures and provides causal evidence that stress undermines human resilience to digital deception.

These findings also connect to broader work on security fatigue, where repeated exposure to warnings and demands under stressful conditions leads to disengagement (Johnston et al., 2019). Taken together, the evidence suggests that stress is not simply a background factor but a critical variable influencing human responses to security threats.

C. Organizational and Training Implications

From a practical standpoint, the results challenge common assumptions in security training. Traditional awareness programs emphasize knowledge and detection strategies but implicitly assume that users can apply them consistently. In reality, training may fail precisely when users are under stress. Organizations should therefore consider stress-resilient training paradigms, including:

- **Simulated stress in training:** Incorporating time pressure, authority, or urgency cues into phishing simulations may better prepare employees for real-world attacks.
- **Encouraging pauses:** Embedding reflective moments, such as "double-check before you click" prompts, may

help users resist automatic responses triggered by stress.

- **Stress recognition:** Training could include not just technical cues for phishing but also self-awareness cues (e.g., "if you feel rushed or pressured, stop and verify").

Policy implications follow from these insights. Workplaces that push employees to operate under tight deadlines and constant multitasking may inadvertently increase susceptibility to phishing. A balanced approach should recognize that productivity demands and security vigilance are interdependent. Addressing workplace stress is thus both a well-being and a cybersecurity concern.

D. System and Interface Design Considerations

System designers should anticipate degraded performance under stress. Reliance on human vigilance alone is risky, especially in environments where stress is high. Email clients and communication platforms could provide built-in support, such as:

- **Automated anomaly detection:** Highlighting unusual sender domains or unexpected attachments.
- **Contextual alerts:** Drawing attention to urgency or authority cues within messages.
- **Delays for risky actions:** Requiring a short pause before executing sensitive actions (e.g., entering credentials from an email link).

More broadly, adaptive security interfaces could detect behavioral signs of hurried or inattentive action (e.g., extremely rapid clicks) and escalate safeguards accordingly. By recognizing the interaction between stress and cognition, system design can shift some of the burden away from end users.

E. Limitations

Several limitations of this study should be acknowledged. First, the online adaptation of the Trier Social Stress Test (TSST) likely produced weaker stress induction than in-person protocols. Without live evaluators, the social-evaluative threat was attenuated. This may explain why aggregate stress deltas were small and why effects only appeared in a subgroup. Second, the study relied on self-reported stress via the State-Trait Anxiety Inventory (STAI). While validated, self-reports may not fully capture physiological changes. Future studies should incorporate objective measures, such as cortisol levels or galvanic skin response.

Third, the sample consisted of MTurk participants. While diverse, they may not represent organizational employees who encounter phishing in workplace contexts. MTurk workers are accustomed to multitasking in high-turnover, low-reward environments, which may shape both baseline stress and responses. Nonetheless, the elevated baseline stress scores observed here suggest that the MTurk environment provides a meaningful test case for stress–security interactions.

Finally, the binary decision task simplified real-world phishing responses. In practice, users may ignore suspicious emails, seek help from colleagues, or escalate messages to IT departments. Future research should adopt richer task environments to capture these additional behaviors.

F. Future Directions

Several avenues for future work emerge from these findings:

1. **Multimodal stress measurement:** Combining self-report scales with biosensors would provide more robust measures of stress.
2. **Cross-context replications:** Testing in industries where stress is pervasive (e.g., healthcare, emergency services) could assess ecological validity.
3. **Adaptive security tools:** Evaluating the effectiveness of phishing detection aids that adapt to user stress levels or rushed behavior.
4. **Longitudinal research:** Exploring how repeated stress exposure shapes security vigilance or fatigue over time.

G. Summary

This study demonstrates that stress, when elevated, undermines phishing detection performance. The effect is not universal — it depends on whether stress actually increases — but when present, the impact is significant. The findings contribute to both theory and practice: theoretically, they extend models of stress and cognition into the cybersecurity domain; practically, they suggest that training, policy, and system design must account for stress as a critical factor. Cybersecurity strategies that overlook the influence of stress risk overestimating the reliability of human detection in the face of digital deception.

VI. Conclusion

This study set out to examine the relationship between acute stress and phishing detection, a topic that has been frequently theorized but rarely tested in a controlled experimental setting. By adapting elements of the Trier Social Stress Test (TSST) into an online phishing detection task, we sought to measure how stress levels before and after task performance related to detection accuracy. While overall stress levels across the sample remained largely unchanged, participants who experienced increases in stress performed significantly worse at identifying phishing emails. Reaction time analyses further revealed that stressed individuals responded more quickly but less accurately, consistent with established theories of stress-induced shifts toward automatic responding.

These findings provide several important contributions. First, they extend neurobiological and psychological models of stress into the cybersecurity domain, demonstrating that stress can directly undermine users' ability to detect phishing. Second, they highlight the conditional nature of this effect: stress does not universally impair detection, but when stress levels rise, the impact is both statistically and practically significant. Third, they underscore the importance of considering stress not as a background factor but as a core variable in security behavior research, with implications for training, organizational policy, and system design.

The practical implications are clear. Training programs must be stress-aware, preparing users to recognize not only technical cues of phishing but also situational cues of stress that may compromise their vigilance. Organizational policies should address the reality that productivity demands and security expectations are intertwined; efforts to reduce workplace stress may simultaneously strengthen security resilience. System designers, meanwhile, should explore adaptive interfaces that account for human cognitive limitations under stress, shifting some of the detection burden away from individuals.

Of course, limitations must be acknowledged. Stress induction in online environments is inherently weaker than in laboratory settings, and reliance on self-reported measures such as the State-Trait Anxiety Inventory cannot capture the full physiological picture. The MTurk sample, while diverse, does not perfectly represent organizational employees. Finally, the binary classification task oversimplifies real-world phishing responses.

Future research should address these limitations by incorporating physiological stress measures, testing in high-stress occupational contexts, designing richer task environments, and exploring adaptive system supports. Longitudinal studies could also examine how repeated stress exposure interacts with security fatigue and long-term susceptibility to deception.

In conclusion, this study demonstrates that stress matters in cybersecurity. When stress levels rise, users are more likely to misclassify phishing emails, particularly those leveraging authority and urgency. Recognizing and addressing the interplay between stress and security behaviors is essential for building more resilient organizations and systems. Technical defenses alone cannot fully address the human element; only by integrating insights from psychology, neuroscience, and organizational science can we hope to reduce vulnerability to digital deception.

REFERENCES

[1] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Commun. ACM*, vol. 50, no. 10, pp. 94–100, 2007.

[2] M. Workman, "Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security," *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, no. 4, pp. 662–674, 2008.

[3] R. B. Cialdini, *Influence: Science and Practice*. Allyn and Bacon, 2001.

[4] A. Vance, M. Siponen, and S. Pahnila, "Motivating IS security compliance: Insights from Habit and Protection Motivation Theory," *Inf. Manage.*, vol. 49, no. 3–4, pp. 190–198, May 2012, doi: 10.1016/j.im.2012.04.002.

[5] A. F. Arnsten, "Stress signalling pathways that impair prefrontal cortex structure and function," *Nat. Rev. Neurosci.*, vol. 10, no. 6, pp. 410–422, 2009.

[6] B. S. McEwen, "Physiology and neurobiology of stress and adaptation: Central role of the brain," *Physiol. Rev.*, vol. 87, no. 3, pp. 873–904, 2007.

[7] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *J. Comp. Neurol. Psychol.*, vol. 18, no. 5, pp. 459–482, Nov. 1908, doi: 10.1002/cne.920180503.

[8] D. Henshel, C. Sample, M. G. Cains, and B. Hoffman, "Integrating Cultural Factors into Human Factors Framework and Ontology for Cyber Attackers," 2016, pp. 123–137. doi: 10.1007/978-3-319-41932-9_11.

[9] A. C. Johnston, M. Warkentin, and M. Siponen, "An Enhanced Fear Appeal Rhetorical Framework: Leveraging Threats to the Human Asset Through Sanctioning Rhetoric," *MIS Q.*, vol. 39, no. 1, pp. 113–134, 2015.

[10] K. Renaud, M. Dupuis, and R. Searle, "Cybersecurity Regrets: I've had a few .... Je Ne Regrette," 2022.

[11] S. Aurigemma and T. Mattson, "Generally Speaking, Context Matters: Making the Case for a Change from Universal to Particular ISP Research," *J. Assoc. Inf. Syst.*, pp. 1700–1742, 2019, doi: 10.17705/1jais.00583.

[12] R. Dhamija, J. D. Tygar, and M. Hearst, "Why Phishing Works," in *CHI 2006 Proceedings, Security*, Montréal, Québec, Canada: ACM, 2006.

[13] P. Kumaraguru, S. Sheng, A. Acquisti, L. Cranor, and J. Hong, "Teaching Johnny not to fall for phish," *ACM Trans Internet Techn*, vol. 10, May 2010, doi: 10.1145/1754393.1754396.

[14] S. Sheng *et al.*, "Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish," in *Proceedings of the 3rd symposium on Usable privacy and security*, Pittsburgh Pennsylvania USA: ACM, July 2007, pp. 88–99. doi: 10.1145/1280680.1280692.

[15] L. Schwabe and O. T. Wolf, "Stress prompts habit behavior in humans," *J. Neurosci.*, vol. 29, no. 22, pp. 7191–7198, 2009.

[16] A. C. Johnston, M. Warkentin, A. R. Dennis, and M. Siponen, "Speak their Language: Designing Effective Messages to Improve Employees' Information Security Decision Making," *Decis. Sci.*, vol. 50, no. 2, pp. 245–284, Apr. 2019, doi: 10.1111/deci.12328.

[17] G. Paolacci and J. Chandler, "Inside the Turk: Understanding Mechanical Turk as a participant pool," *Curr. Dir. Psychol. Sci.*, vol. 23, no. 3, pp. 184–188, 2014.

[18] J. Chandler and D. Shapiro, "Conducting clinical research using crowdsourced convenience samples," *Annu. Rev. Clin. Psychol.*, vol. 12, pp. 53–81, 2016.

[19] M. Dupuis, B. Endicott-Popovsky, and R. Crossler, "An Analysis of the Use of Amazon's Mechanical Turk for Survey Research in the Cloud," in *International Conference on Cloud Security Management*, Seattle, Washington, Oct. 2013.

[20] M. Dupuis, K. Renaud, and R. Searle, "Crowdsourcing Quality Concerns: An Examination of Amazon's Mechanical Turk," in *The 23rd Annual Conference on Information Technology Education*, Chicago IL USA: ACM, Sept. 2022, pp. 127–129. doi: 10.1145/3537674.3555783.

[21] C. Kirschbaum, K. M. Pirke, and D. H. Hellhammer, "The Trier Social Stress Test – A tool for investigating psychobiological stress responses in a laboratory setting," *Neuropsychobiology*, vol. 28, no. 1–2, pp. 76–81, 1993.

[22] C. D. Spielberger, *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, 1983.

[23] M. J. Dupuis and S. Smith, "Clickthrough testing for real-world phishing simulations," in *Proceedings of the 21st Annual Conference on Information Technology Education*, 2020, pp. 347–347.