A Platform for Physiological and Behavioral Security

Felix Dietz
felix.dietz@ifi.lmu.de
LMU Munich / University of the Bundeswehr Munich
Munich, Germany

Peter Heubl
peter.heubl@ruhr-uni-bochum.de
Ruhr University Bochum
Bochum, Germany

Luke Haliburton
luke.haliburton@ifi.lmu.de
LMU Munich
Munich, Germany

David Bothe
bothe@internet-sicherheit.de
Institut für Internet-Sicherheit - if(is), Westfälische Hochschule Gelsenkirchen
Gelsenkirchen, Germany

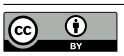
Jan Hörnemann
jan@aware7.de
aware7 GmbH
Gelsenkirchen, Germany

M. Angela Sasse
martina.sasse@rub.de
Ruhr University Bochum
Bochum, Germany

Florian Alt
florian.alt@ifi.lmu.de
LMU Munich / University of the Bundeswehr Munich
Munich, Germany

Abstract

Human-centered security research traditionally leverages self-reports and high-level behavioral data. However, the increasing ubiquity of sensors integrated into personal, wearable devices (e.g., smartphones, smartwatches) and in users' environments (e.g., cameras) enables researchers to unobtrusively collect rich physiological and behavioral signals. These real-time data streams can reveal user states—such as attention or workload—that can be employed to design adaptive security mechanisms. In this paper, we present a platform that supports designing, building, and evaluating next-generation user interfaces that leverage physiological and behavioral data for enhanced security. First, we introduce the physio-behavioral security paradigm, highlighting how sensor-based insights into user states can inform individualized security interventions and accurately identify moments of vulnerability. We then outline the requirements, system architecture, and implementation details of the platform, illustrating how multiple data streams (e.g., gaze, heart rate, keystrokes, mouse movements) are integrated and securely processed. Finally, we report on an exploratory deployment in a mid-sized organization, showcasing how the tool captures real-time security behaviors and enables context-aware interventions. The deployment yields insights into factors influencing acceptance across different stakeholders (management, IT department, employees). Our results suggest that adaptive approaches, informed by physiological and behavioral signals, can improve security outcomes and user acceptance.



This work is licensed under a Creative Commons Attribution 4.0 International License.
NSPW '25, Aerzen, Germany
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1875-5/2025/08
https://doi.org/10.1145/3774761.3774915

CCS Concepts

• **Security and privacy → Usability in security and privacy**.

Keywords

Human-Centered Security, Physiological Sensing, Behavior

ACM Reference Format:
Felix Dietz, Peter Heubl, Luke Haliburton, David Bothe, Jan Hörnemann, M. Angela Sasse, and Florian Alt. 2025. A Platform for Physiological and Behavioral Security. In *New Security Paradigms Workshop (NSPW '25), August 24–27, 2025, Aerzen, Germany.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3774761.3774915

1 Introduction

Over the past two decades, human-centered security research has made substantial progress in understanding how individuals perceive and respond to threats - from phishing attempts to password misuse - by emphasizing user experience and behavior. Seminal studies, such as 'Why Johnny Can't Encrypt' [62], and 'Users Are Not the Enemy' [7], demonstrated that security incidents often arise from usability obstacles or user misconceptions rather than purely technical flaws.

Despite these advances, much work still relies on self-reported data or high-level activity logs, generally captured through interviews and surveys, e.g., about phishing susceptibility or password practices [29, 54]. These methods yield valuable subjective insights yet capture only a fraction of users' situational states.

Simultaneously, the proliferation of sensors in personal devices (e.g., smartphones, smartwatches) and work environments (e.g., meeting-room cameras) has created avenues for the unobtrusive collection of physiological signals - such as heart rate, gaze, and skin conductance. These signals offer real-time indicators of user states like attention, workload, and stress, critically influencing security-relevant decisions. For instance, a user under time pressure or high cognitive load may unknowingly click a malicious link

or reuse passwords across multiple accounts. Identifying such states on the fly opens new opportunities for adaptive security mechanisms that respond to users' moment-to-moment needs. Building on our earlier introduction of the physio-behavioral security paradigm [9], this paper focuses on realizing that vision through a practical platform that enables its implementation and study in real-world environments.

At the same time, this emerging paradigm of physio-behavioral security - leveraging physiological and behavioral data in real-world security contexts - remains relatively unexplored. Existing efforts often address isolated applications, such as detecting cases where humans reuse passwords [5] or identifying exposure to phishing [3] or fake news [4], rather than broader frameworks for detecting and mitigating security risks. Moreover, stakeholders often raise concerns about continuous physiological monitoring, underscoring the need to balance potential security gains with user privacy, transparency, and consent.

In this paper, we present a platform for designing, building, and evaluating next-generation security user interfaces driven by real-time physiological and behavioral input.

(1) We present a novel platform architecture for physio-behavioral security that securely integrates multi-modal sensing (e.g., gaze, heart rate, keystrokes) in real-world contexts, outlining its requirements, architecture, and implementation.
(2) We demonstrate the platform's potential and efficacy through a field deployment in a mid-sized organization, revealing how real-time monitoring and targeted interventions can improve security behaviors and user acceptance.
(3) Finally, we provide insights into privacy, scalability, and stakeholder concerns to inform the broader adoption of physiology-aware security solutions and guide future deployment of adaptive, context-aware mechanisms.

Our results indicate that adaptive approaches informed by a user's situational state can improve immediate security outcomes and foster greater acceptance, as interventions can be delivered when users are most receptive. We conclude by discussing practical considerations - such as data privacy, consent, and scalability - and offer guidance on bringing physiology- and behavior-aware security solutions into broader use.

2 Related Work

We introduce essential concepts and discuss how sensor-based approaches in ubiquitous computing can enrich human-centered security by revealing user states such as stress, workload, or inattention. Afterwards we review sensing platforms proposed by the ubicomp community.

2.1 Behavior and Physiology in Human-Centered Security

2.1.1 Research Methodology in Usable Security. Studies in Usable Privacy and Security frequently employ interviews, surveys, or lab-based user tests [29], with minimal use of sensor-based data. Such approaches capture attitudes and self-reported behaviors but can miss in-the-moment user states (e.g., stress, workload) that lead to insecure decisions [48]. Self-reports also rely on participants' memory, risking biased recall. Where click-based metrics (e.g., phishing link clicks) are collected, they often lack granularity in pinpointing why a user clicked.

Recent works illustrate the potential of more nuanced sensor data. For instance, Almehmadi [8] leveraged micro-behavioral cues to discriminate accidental clicks from intentional ones, highlighting how fine-grained analysis of movement patterns can address slip-based security errors. Eye-tracking research suggests that prolonged fixations on suspicious email elements can reflect heightened cognitive or emotional response [45, 47], while Rozentals [55] found that stress and email overload impede secure decision-making.

2.1.2 Behavioral and Physiological Research in Usable Security. Prior research leveraged different types of sensors in various application areas.

Behavioral Biometrics. A substantial body of work has examined *behavioral biometrics*, focusing on how fine-grained user actions can authenticate or continuously verify identity [17, 19, 25]. Typing rhythms (keystroke dynamics), touch-targeting patterns, and motion gestures have proven predictive of an individual's identity [20, 26], leading to potential second-line security checks (e.g., detecting impostors). Beyond identity, advanced usage data (such as *where* and *when* a user lingers during sign-in) offers insights into confusion or intention [18]. However, such work typically stops short of real-time adaptation to detect or mitigate attacks.

Eye Gaze and Security. Gaze tracking has been used for authentication [27, 40], but also to gauge security behavior. For example, Arianezhad et al. [11] found that security experts exhibit different gaze durations on browser security indicators, while Miyamoto et al. [47] introduced EyeBit, which observes whether users visually inspect email elements that signal phishing. Recent analyses have shown that gaze behaviors can unveil password reuse intentions [5] and highlight overlooked cues in fraudulent emails [45, 47].

Physiological Indicators and Stress. Other physiological signals—heart rate (HR), skin conductivity (GSR/EDA), and electroencephalography (EEG)—were explored. Neupane et al. [50] combined EEG and eye tracking to examine user responses to phishing emails and malware warnings, while Hashem et al. [31] studied ECG readings to flag insider threats. These signals can track arousal or mental workload, which heavily influences risk-taking tendencies [28]. Building on psychological findings that stress or cognitive overload compromise vigilance [48, 64], researchers envision context-aware defenses responding to real-time physiological triggers.

2.1.3 The Interplay of Physiology and Behavior. Users often exhibit combined physiological and behavioral cues under stress or high workload. A spike in HR or a sharper EDA response may coincide with erratic mouse movements [65], slower reaction times, or more false clicks [8]. Attackers frequently exploit these states by inducing fear or urgency—for example, threatening account closure [30].

Likewise, situational factors affect user states: email overload or time pressure can make a user more prone to unsafe actions [55]. Real-time monitoring of these cues can inform interventions, such as intensifying phishing alerts when signals of confusion or stress arise [47]. Schwarz et al. [57] argue that such states (attention, engagement, workload, stress) mediate human decision-making; thus, measuring them in situ could identify the precise moments

users need safeguarding. In short, *behavior plus physiology* offers a high-resolution lens on user cognition, bridging the gap between static security training and dynamic, contextually tailored defenses.

2.2 Sensing Platforms

Recent advances in wearable technology have enabled innovative applications in health monitoring and human-computer interaction. For example, Zhang et al. [66] demonstrated how wearable sensors can predict post-operative complications in patients undergoing pancreatic surgery. Similarly, Ley-Flores et al. [41] examined the design of sensory feedback for wearables, showing how body sensations can support physical activity.

Bari et al. [14] proposed an automated approach for detecting stressful conversations using wearable physiological and inertial sensors. Complementing this work, Mishra et al. [46] evaluated the reproducibility of physiological stress detection models across multiple studies, sensor types, and populations.

Research has also focused on using physiological and facial cues to assess cognitive performance. Sharma et al. [58] demonstrated that features extracted from physiological responses and facial expressions can reliably predict cognitive performance across tasks such as gaming, coding, and adaptive assessment.

The intersection of human-computer interaction and physiological monitoring has also been explored in the context of security. Earlier work by Hercegfi et al. [35] presented a methodology for assessing user interaction with computer interfaces by monitoring rapid physiological responses, such as heart period variability, alongside keystroke and mouse data. Similarly, Möller et al. [49] modeled user behavior in security-related contexts using a mixed probabilistic and rule-driven state machine.

Expanding beyond stress detection, affective computing in immersive environments has also gained attention. Tabbaa et al. [59] introduced the VREED dataset, a multimodal affective dataset combining eye tracking and physiological measures (ECG and GSR) collected during exposure to immersive 360° video-based virtual environments.

2.3 Summary

Human-centered security has historically focused on self-reports and static observations, overlooking the interplay of user physiology and behavior in the moment. Recent studies demonstrate how physiologically driven signals (e.g., stress, cognitive load) and nuanced behavioral data (e.g., keystroke timing, gaze fixation) can yield deeper insights into why users make insecure choices. Meanwhile, sensing platforms, including wearable and environmental sensors, facilitate unobtrusive, continuous data collection. Building on these advances, the next step is to develop architectures and privacy-preserving frameworks that fuse these data streams for real-time, context-aware security interventions.

3 An Introduction to Physio-Behavioral Security

With the rapid proliferation of ubiquitous sensing in the form of smartphones, smartwatches, wearable devices, and ambient sensors (e.g., office cameras or voice assistants), a rich set of behavioral and physiological signals has become available to researchers. These signals - ranging from keystroke dynamics and gaze patterns to heart rate and skin conductance - can reveal user states such as attention, stress, or workload in real time [39, 57]. This advance represents a potentially transformative development for human-centered security: rather than relying solely on user self-reports or posthoc behavioral logs, it becomes feasible to adapt security interventions as they unfold, considering an individual's current cognitive or emotional state.

This section provides an introduction to human-centered behavioral and physiological security [9] - hereafter "physio-behavioral security". We define core terminology and then explain how physiological and behavioral signals can inform novel security mechanisms. Next, we discuss how established human-centered attacks might be mitigated through better awareness of user states and how this relates to broader security habits. We then present a structured research space illustrating the interplay between user states and security threats. Throughout, we highlight open challenges, emphasizing areas of synergy between usable security and emerging physiological computing techniques.

3.1 Defining Physio-Behavioral Security

Alt et al. define *physio-behavioral security* as an approach to designing, analyzing, and evaluating security mechanisms (e.g., authentication, phishing detection, warning systems) by leveraging sensor data about the user's current physiological state (e.g., heart rate, skin conductance, gaze) and behavior (e.g., typing patterns, mouse movements, gestures) [9]. The overarching goals are twofold:

Improving Understanding Researchers, designers, and evaluators can collect richer insights into *how* users experience and respond to security-critical moments. Traditionally, many assumptions about user states (such as attention or workload) have been inferred indirectly. By examining actual physiological or behavioral signals, it becomes possible to map states more precisely to moments of risk or vulnerability.

Enhancing Security Mechanisms Security systems can be made context-aware by incorporating real-time knowledge of a user's state. For instance, a system might detect increased stress and provide an extra confirmation step for a high-stakes transaction. It could also recognize low attentional focus in an email application and proactively highlight critical security indicators to help the user discern genuine messages from phishing attempts.

Although behavior has long been a part of user-centered security research - click metrics or user logs, for instance - the kind of granular, sensor-derived behavioral data we consider here (e.g., keystroke flight times, micro-pauses in mouse movement) goes beyond standard event capture. Likewise, collecting physiological data (e.g., heart rate variability, pupillary responses) has traditionally been confined to controlled lab experiments (e.g., EEG-based workload studies [39, 42]) or specialized user research. Today, consumer-grade devices (e.g., smartwatches) and ubiquitous environmental sensors (e.g., camera-based eye tracking) create a broader potential for real-world usage. This potential has previously been identified by Katsini et al. [36]. It is important to note that such signals capture

activation and workload but are not sufficient to reliably infer nuanced emotional valence, personal intent, or complex mental states. Physiological surges, for instance, may reflect stress or physical exertion, making fine-grained distinctions inherently uncertain. A key research problem is to determine the utility of these signals in practice.

A related but narrower concept is behavioral biometrics [53], where the primary focus is using behavior (e.g., typing or gait) to authenticate or identify individuals. Although physiological-behavioral security includes insights from behavioral biometrics, it aims to extend well beyond authentication, considering how knowledge of dynamic user states, not just static identity, can inform protective actions, warnings, or adapt the entire security interface.

3.2 Why User States Matter for Security

Security vulnerabilities often emerge from lapses in user attention or from misconceptions [34, 51, 63]. For instance, a user under high cognitive load might ignore subtle indicators of a phishing attempt, such as a suspicious URL or spelling mistakes in an email. Similarly, stress or fatigue can lead users to select weak or reused passwords rather than carefully creating a unique one for every new service. Attackers also exploit these states: many social engineering strategies create a sense of urgency or anxiety to prompt rash user actions [30, 64].

Real-time detection of such states - attentional drift, elevated stress, or fatigue - thus holds particular promise for security applications. State-detection could enable an interface to intervene at the exact moment when a user is most susceptible to error, say by offering a gentle reminder or an automated fallback mechanism (e.g., generating a secure password). Such interventions could not only reduce security incidents but also spare users from the constant vigilance typical security "best practices" demand. Precisely targeting interventions at high-risk moments may also alleviate warning fatigue, a well-known problem where repeated or poorly timed alerts cause users to dismiss warnings altogether [10].

3.3 From Observation to Intervention: Physio-Behavioral Data in Security Research

Physio-behavioral data has not only deepened our understanding of security-relevant behavior but also has the potential to drive adaptive, real-time interfaces that address user vulnerabilities in context.

Enhancing Understanding of Security Practices. A focus in security research is examining how individuals behave in tasks such as creating passwords, responding to phishing, or configuring devices. Traditional methods rely primarily on self-reports and recall-based accounts, which may overlook subconscious reactions. By contrast, physiological signals (e.g., elevated heart rate or pupil dilation) and fine-grained behavioral data (e.g., keystroke latencies) can uncover hidden indicators of confusion, stress, or cognitive effort. These insights help researchers formulate theories about why certain risky behaviors persist despite awareness campaigns, and how transient emotional states can influence security decisions.

Building Adaptive Security Systems. Adaptive security interfaces can dynamically respond to user states, akin to an IT help desk that senses user confusion. For instance:

Password Guidance If sensor data indicates high workload during password creation, the system might offer passphrase suggestions or password management assistance.

Phishing Defenses Real-time gaze or mouse tracking could detect user inattention in email clients; in such moments, the interface might highlight critical sender details or briefly block suspicious links.

Insider Threat Detection Indicators of unusual stress or anxiety patterns - captured by workstation cameras or wearables - could signify malicious behavior or signal a need for immediate support.

Leveraging these continuous, real-time cues, physio-behavioral approaches could enable security systems to adapt to *ongoing user states*, bridging the gap between human cognition and protection.

3.4 Human-centered Threats, Security Habits, and the Research Space

Human-centered attacks target users, aiming to obtain sensitive information - most commonly credentials that can unlock financial or enterprise resources. Such compromises allow attackers to extract payment details, exfiltrate sensitive data, or escalate privileges across a broader network. Below, we highlight four major categories of these threats: guessing attacks, observation attacks, social engineering, and reconstruction attacks. This list is not meant to be comprehensive but to illustrate the potential of the approach.

Guessing attacks involve systematically probing potential secrets (e.g., passwords, PINs, lock patterns) until the correct one is discovered. Offline brute force attacks cycle through vast combinations of usernames and passwords, while dictionary attacks rely on pre-compiled lists of common passphrases *(e.g., "password123")* to exploit weak credential choices. Credential stuffing leverages stolen usernames and passwords gleaned from data breaches, with attackers rapidly testing these pairs across services.

Observation attacks capitalize on physically or digitally monitoring the user. Shoulder surfing occurs when an attacker covertly watches someone typing a password, either in person or via camera feeds, thus capturing credentials without the victim's knowledge. Keyloggers similarly record keystrokes or screenshots through malicious software, archiving passwords, PINs, and other sensitive data. In a more passive approach, network sniffing tools (e.g., WireShark) capture data traversing insecure connections, often revealing user logins or session tokens that can be reused by attackers.

Social engineering exploits trust and human error to persuade victims into divulging private information. Phishing scams, including spear phishing, send convincing emails containing malicious links or attachments, often evading automatic spam filters. Attackers tailor these messages to specific individuals or companies, making detection challenging. Vishing attacks operate similarly through voice calls or voice messages, where impostors impersonate authoritative entities (e.g., a bank) to manipulate victims into handing over

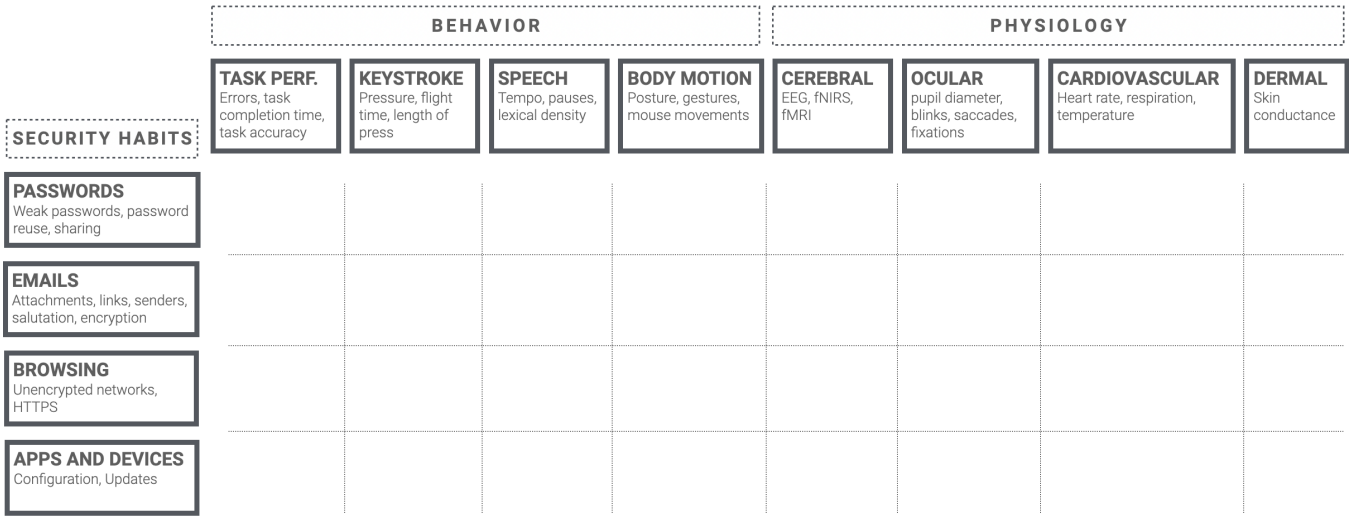


Figure 1: Research Space for User States in Security Tasks: Security habits lead to observable changes in user behavior and physiology. These changes can be used to better target user interfaces to users' states.

account details or transferring funds. Other variants range from "smishing" (SMS-based phishing) to pop-up phishing on web browsers, all drawing on deception and urgency to trick users.

Reconstruction attacks arise when adversaries gain physical possession of a device and seek to infer credentials from residual traces. Smudge attacks focus on the skin-oil residue users leave on touch surfaces, which can reveal lock patterns or PINs when viewed under angled lighting [13]. Similarly, thermal attacks capture heat signatures on keyboards or touchscreens using thermal imaging, allowing attackers to infer recently entered sequences [1, 15].

Overall, each attack highlights the role of human vulnerabilities - such as weak password habits, unguarded behavior, and the ease of deception. A platform that assesses real-time physiological and behavioral cues can mitigate these threats by detecting instances of elevated stress or diminished attention, intervening precisely when individuals are most likely to make security-critical errors. This integration of context-aware detection and tailored guidance complements existing technical safeguards, potentially reducing both the frequency and impact of human-centered attacks.

In the following, we present a research space structured along two dimensions (see Figure 1): human security habits and physio-behavioral data as a source for mitigation technologies.

3.4.1 Dimension 1: Human Security Habits. The above-mentioned threats become possible through habits that establish themselves around the use of security mechanisms or the general behavior in security situations.

Users' day-to-day interactions - choosing passwords, handling emails, browsing the web, and installing apps or devices - often determine their susceptibility to the previously described human-centered attacks. Although not exhaustive, the examples below illustrate how knowledge of user states (e.g., workload, attention, or stress) can inform more effective interventions.

Passwords. Weak or carelessly chosen passwords increase susceptibility to guessing attacks. Many users have misconceptions about password strength [60], and stronger credentials are less likely when users are rushed (e.g., forced to set a password at checkout time). Systems that are aware of a user's workload or attention level could identify more suitable moments to encourage secure choices. Password reuse further exacerbates risk by enabling credential stuffing, yet many users struggle to grasp the resulting dangers, despite tools like *Have I Been Pwned*. Knowing when users type a frequently used password could trigger targeted warnings or proactive explanations. In addition, users sometimes share passwords through insecure channels, creating further vulnerabilities. Monitoring situational awareness could prompt safer sharing methods (e.g., secure password managers). Finally, when entering passwords, real-time feedback about potential shoulder surfing or reconstruction attacks [56] may help protect users without causing unnecessary friction.

Emails. Phishing remains a major threat, often exploiting fatigue or heightened emotions. Attackers send malicious links or attachments when users are distracted, for instance near workday's end. Real-time sensing of attention might nudge individuals to scrutinize suspicious links more carefully or confirm unusual sender addresses. Indeed, research suggests users mostly focus on the header and body of an email [52], indicating that interventions should prioritize these areas. Tools like EyeBit [47] demonstrate how gaze-based awareness can encourage checking a URL's legitimacy before interacting. Similarly, if a user's emotional state indicates stress or urgency, email clients could delay or visually highlight suspicious attachments, helping mitigate impulsive clicks.

Browsing. A third set of habits emerges when surfing the web, especially on untrusted networks (e.g., airport Wi-Fi). Many such networks remain unencrypted, leaving user data vulnerable unless the user manually enables a VPN. Systems that detect user confusion or lack of familiarity could prompt the immediate use of encryption. Likewise, ensuring that HTTPS is enabled for secure transactions is sometimes overlooked if users fail to notice security

indicators in the browser. Monitoring attentional cues might trigger timely reminders or shift these indicators into more prominent places.

Applications and Devices. New devices (e.g., routers, smart home appliances) and installed apps require proper configuration - often with default passwords, complex security settings, or potential privacy pitfalls. Manufacturers may ship secure defaults, but a confused or impatient user might disable critical protections. Systems that gauge user workload or frustration during setup could provide simplified instructions or postpone prompts until the user is more receptive. Similarly, scheduling software updates at moments when the user is less burdened may improve compliance. By tailoring interventions to the user's current state, security-critical tasks become less of an annoyance and more of a timely safeguard.

3.4.2 Dimension 2: Detecting User States via Physiology and Behavior. Behavioral and physiological signals can reveal moment-to-moment user states, such as cognitive load, emotion, and stress [39]. The ability to detect these states in real time opens the door to more adaptive, context-aware security interventions.

Behavioral Signals. Users' observable interactions - typing, speech, and body motion - often mirror their underlying cognitive and emotional conditions. For instance, metrics such as task completion time and error rates can reflect increased mental load, while keystroke dynamics can signal uncertainty or elevated stress [16, 44]. Speech features, including tempo and pauses, similarly provide clues about cognitive load and affect [22, 37]. Likewise, body posture and gestures - captured via on-body sensors, cameras, or mouse-movement data - have been linked to user states ranging from attention to stress [12, 61]. These behavioral observations require minimal specialized hardware (microphones or webcams embedded in phones and laptops) and can often be collected unobtrusively.

Physiological Signals. In tandem with behavior, physiological responses offer deeper insight into user states. Neuroimaging techniques like EEG and fNIRS can reveal workload or engagement [33, 39] without requiring a laboratory environment, thanks to advances in lightweight, wearable sensors [32]. Ocular measures - ranging from pupil dilation to fixation and saccade patterns - can predict cognitive load [38] or highlight security-relevant actions (e.g., verifying a URL) [5]. Cardiovascular, respiratory, and thermal signals collected via chest straps, smartwatches, or thermal cameras further correlate with stress or emotional arousal [2, 24]. Meanwhile, electrodermal activity (EDA) readings, also available on some commercial smartwatches, can capture physiological arousal related to stress or heightened vigilance [23, 39].

Implications for Security. When leveraged appropriately, these behavior and physiological indicators can signal whether a user is distracted, overloaded, or experiencing heightened arousal at a critical juncture - such as reading a suspicious email or setting a new password. By integrating and analyzing these signals, security systems can adapt their real-time interventions, prompting users to double-check critical indicators or momentarily offering additional guidance. As low-cost sensors and embedded hardware mature, the potential to deploy such adaptive solutions in everyday digital environments grows, making it increasingly feasible to safeguard users precisely when they are most at risk of falling for an attack.

3.5 Examples Illustrating Physio-Behavioral Security in Practice

Prior studies illustrate how combining physiological and behavioral data can enhance security designs.

3.5.1 Understanding Secure Password Habits. Early work in this domain examined how users' cognitive states - tracked via eye gaze and pupil dilation - can shape their password choices [5, 6]. In one study, participants were asked to create strong and weak passwords. Pupil dilation proved consistently higher for stronger passwords, reflecting increased cognitive load. This insight suggests that future systems could detect user strain in real time - e.g., via an AR headset - and proactively nudge people when they are about to choose a weak password without knowing the actual password.

Further research investigated password reuse, revealing that certain gaze and keystroke features (e.g., fixation duration on the keyboard) strongly predict when users are reusing passwords. This recognition happens even before typing begins, hinting at "just-in-time" interventions. While existing password managers only flag reuse after a password is saved, sensing user states during creation might preempt insecure choices, potentially lowering the risks of credential stuffing.

3.5.2 Detecting Phishing Email Exposure. A second example focused on users' eye gaze and mouse movements while handling emails [3]. Participants, role-playing an office worker, categorized emails into appropriate folders. Features like hover speed and fixation counts often align with users' awareness of potentially malicious content. Although results were less robust than for password tasks - partly due to email complexity and personal risk-taking behavior - they still underscore how real-time sensing of attentional cues may enable adaptive alerts. For instance, if a system infers the user is distracted or fatigued, it could highlight suspicious links or slow down the click-through process to reduce the chance of hasty decisions.

These examples illustrate how integrating physiological (e.g., pupil dilation, heart rate) and behavioral (e.g., keystroke or mouse dynamics) data offers nuanced ways to detect and mitigate security lapses. Through state-aware interfaces, interventions could be personalized and timed to moments of highest vulnerability - like prompting users before finalizing a reused password or calling attention to suspicious email indicators.

3.5.3 Supporting Reflection and Planning. Beyond direct security interventions, physio-behavioral sensing could also support self-regulation. For instance, detecting heightened arousal while composing sensitive emails might prompt users to pause and reconsider before sending a message they could later regret. Rather than enforcing a specific action, such feedback could provide individuals with moments for reflection and planning, highlighting how physio-behavioral insights can extend beyond immediate threat mitigation toward fostering more mindful security and communication practices.

3.6 Challenges and Considerations in Physio-Behavioral Security

Physio-behavioral security approaches face notable challenges, including ensuring accurate detection of user states (given individual differences and complex tasks), managing potential privacy concerns surrounding the monitoring of physiological signals, and calibrating the frequency of interventions to avoid user frustration.

3.6.1 Privacy and Ethical Implications. While new sensing capabilities open possibilities for more personalized security, continuous monitoring of physiological or behavioral data raises pressing concerns about user privacy. Personal data such as heart rate or stress could reveal sensitive health or emotional information. In a workplace setting, employees might fear surveillance or employer misuse (e.g., evaluating performance based on stress patterns). Likewise, storing these signals can create new attack surfaces if stolen or misused by third parties.

Hence, consent models, data minimization, and secure data handling protocols are essential. Future systems must balance the necessity of real-time data analysis with respect for user autonomy, giving individuals control over the scope and duration of physiological monitoring. Methods such as local on-device processing (to avoid raw data transmission to external servers) and privacy-preserving machine learning techniques (e.g., differential privacy) may become essential design strategies.

3.6.2 Uncertainty and Accuracy Trade-offs. Physiological and behavioral signals often exhibit noise and variability across individuals, contexts, and devices. A single sensor reading, like a heart-rate spike, might result from climbing a flight of stairs rather than encountering a malicious link. Approaches must robustly integrate multiple features to reduce false positives (unnecessary or annoying interventions) and false negatives (missing real threats). Designers must also communicate the uncertain nature of these inferences, perhaps showing confidence levels (e.g., "High probability you're reusing a password here - please verify?"). This challenge is particularly acute for real-time warning systems, where even modest false-positive rates can quickly erode user trust. In practice, this means that highly accurate detection is required before deploying direct physiology-based warnings. Until such accuracy is reliably achieved, pre-emptive or context-sensitive adaptations (e.g., delaying risky tasks or offering supportive defaults) may represent a more realistic application path.

3.6.3 User Experience and Acceptance. A central tenet of human-centered design is ensuring that added complexity does not degrade usability. If physiologically adaptive security systems trigger too many false alarms or if they feel intrusive (e.g., frequent requests to "calm down" or "pay attention"), users may abandon or circumvent them. Longitudinal studies are necessary to evaluate how acceptance evolves over time and to identify the thresholds at which adaptive systems become more helpful than burdensome.

3.6.4 Technical Integration and Scalability. Finally, while some forms of physiological sensing (e.g., heart rate from a smartwatch) are becoming commonplace, others (e.g., EEG or advanced eye tracking) remain less accessible or require specialized hardware. Scaling up from a controlled lab or small pilot study to broad workplace or consumer adoption entails:

- Cross-device consistency: Dealing with heterogeneous sensors across smartphones, smartwatches, and in-room cameras, each with different sampling rates and noise characteristics.
- Secure data pipelines: Ensuring data is encrypted and processed either locally or in a trusted cloud environment with minimal risk of leaks.
- Maintenance: Updating sensor drivers, handling data storage, and retraining machine learning models as hardware or usage contexts change.

3.7 From Concept to Practice: Looking Ahead

Realizing the vision of physio-behavioral security requires moving beyond static, universal defenses toward adaptive systems that detect and mitigate threats in real time. By combining insights from usable security and physiological computing, researchers can better understand why security failures occur, whether due to stress, fatigue, or cognitive overload, and design interventions that respond to users' momentary states.

In this paper, we take a concrete step towards this goal by introducing a platform that operationalizes the physio-behavioral security paradigm. Our system integrates multi-modal sensing with secure data handling, extensible interfaces, and privacy safeguards to support the rapid prototyping and evaluation of adaptive mechanisms in real-world settings. Future research can build on this foundation to refine real-time detection methods, explore additional sensing modalities, and design context-aware interventions that integrate seamlessly into users' workflows. By offering an adaptable architecture and clear deployment pathways, our platform bridges the gap between conceptual frameworks and practical, deployable solutions for physiology-aware security.

4 A Platform for Physio-Behavioral Security – Reference Implementation

4.1 Requirements Analysis

This section outlines the key requirements for a platform that supports designing and developing user interfaces leveraging physiological and behavioral sensing in security contexts. Drawing on prior work, early prototypes, and feedback from security experts and end users, we identify five core needs: *(i) flexible sensor integration*, *(ii) real-time data processing*, *(iii) security and privacy*, *(iv) data management and scalability*, and *(v) developer and researcher support*. Together, these requirements establish the foundation for an infrastructure deployable across various environments, from controlled lab studies to large-scale organizational rollouts.

4.1.1 Flexible Sensor Integration. A physio-behavioral security platform must support a wide range of sensing modalities - such as eye tracking, keystroke dynamics, mouse movement, electrodermal activity, and heart rate - to capture diverse aspects of user state. This variety is essential given the variability in human behavior and physiology; no single signal suffices for assessing attention, stress, or workload. Real-world use also demands device-agnostic integration, as users rely on different hardware (e.g., wearables, phones,

webcams). A modular architecture should allow easy addition or removal of sensor drivers, minimizing engineering overhead. This flexibility is key for accommodating new sensing technologies and tailoring configurations to specific research or deployment needs.

4.1.2 Real-Time Data Processing. To support adaptive security, the platform must enable real-time or near-real-time analysis of the user state, as physiological signals are most useful when promptly interpreted, e.g., detecting stress surges to guide immediate interventions. Low-latency processing requires efficient handling of continuous data via feature extraction pipelines that convert raw input (such as gaze or photoplethysmography) into actionable metrics (e.g., heart rate variability, pupil dilation). Adaptive sampling helps manage resource use and privacy, increasing sensor frequency during critical tasks (e.g., password entry) and lowering it otherwise. The system must also tolerate noisy or missing data, ensuring reliable inferences despite artifacts from movement or sensor drift. In practice, this means that individual sensor failures (e.g., a smartwatch losing connection) do not invalidate the analysis, as the platform aggregates across multiple modalities. Aggregate features provide robustness by smoothing over noisy channels while still allowing researchers to inspect fine-grained data reliability when needed.

4.1.3 Security and Privacy. Physio-behavioral security involves sensitive data, such as stress levels and attention patterns, or possibly even medical conditions, requiring strong protection. The platform must encrypt data in transit and when stored, support role-based access controls, and offer transparent consent mechanisms for each sensing modality. For example, users might allow mouse tracking but decline heart rate monitoring. Supporting fine-grained consent fosters trust and meets legal requirements. The system should also minimize raw data retention, favoring aggregated or anonymized metrics to reduce privacy risks and storage needs. Overall, privacy-by-design and security-by-design are essential foundations.

4.1.4 Data Management and Scalability. Even moderate-scale deployments can generate large volumes of high-frequency sensor data, requiring robust infrastructure for storage and management. Time-series databases and scalable storage enable efficient querying across sensors and time intervals. Clear retention policies are essential to limit raw data storage and uphold privacy such as aggregated gaze metrics may suffice over raw data. To support both small studies and large-scale deployments, the platform should leverage distributed or cloud architectures, with load balancing and failover mechanisms to ensure stable performance under varying loads.

4.1.5 Developer and Researcher Support. Supporting the full research and development lifecycle requires well-designed APIs and toolkits for sensor ingestion, data processing, model training, and adaptive interventions. Documentation, sample code, and reference implementations ease adoption and promote academic replication. Dashboards for monitoring data quality and real-time inferences aid rapid prototyping. Beyond raw data capture, the platform should integrate key machine learning workflows and support testing of intervention strategies. Standardized pipelines and version control enhance reproducibility, while accessible tools ensure extensibility and broader adoption of physio-behavioral security systems.

Table 1: High-Level Requirements for a Physio-Behavioral Security Platform

Category	Key Requirement	Description
Flexible Sensor Integration	Multi-Modal Data	Collect multiple signals (e.g., heart rate, gaze, keystrokes) to capture diverse user-state indicators.
	Device-Agnostic	Support heterogeneous hardware (wearables, desktops, mobiles) and operating systems.
	Modular Framework	Enable easy addition and removal of sensor drivers with minimal integration overhead.
Real-Time Data Processing	Immediate Analysis	Provide low-latency inference so interventions can be delivered in a timely manner.
	Adaptive Sampling	Dynamically adjust sensor frequency based on user context or task criticality.
	Noise Robustness	Use filtering and fusion techniques to manage artifacts and missing data.
Security & Privacy	Secure Pipelines	Encrypt data in transit and at rest, applying fine-grained access controls.
	On-Device Processing	Favor local computation where possible to limit sharing of raw signals.
	Granular Consent	Permit opt-in/opt-out for specific modalities and ensure transparent usage policies.
Data Management & Scalability	Scalable Storage	Handle high-volume time-series data in multi-user deployments.
	Lifecycle Policies	Automate archiving or deletion of older raw data to reduce privacy risks.
	Load Balancing	Maintain reliability and responsiveness under variable sensor loads.
Developer & Researcher Support	ML Integration	Provide configurable modules for feature extraction, model training, and evaluation.
	Intervention Logic	Offer a rules engine to define triggers for real-time nudges or warnings.
	Documentation	Include APIs, sample code, and dashboards to streamline deployment and experimentation.

Summary. These five requirements - flexible sensor integration, real-time data processing, security and privacy, data management and scalability, and developer/researcher support - form the core framework for building a robust and ethically responsible platform. The subsequent subsections will detail how the system is architected to meet these requirements and how it can be deployed in a variety of real-world scenarios to advance both research and practical security interventions.

4.2 Description of Platform Components

The experimental setup for this study consists of a self-contained network designed to analyze the physiological data of users when interacting with emails, particularly phishing emails. The system architecture (see Figure 2) ensures data integrity and security while allowing continuous recording over multiple working days.

4.2.1 Threat Model. In designing the platform, we consider threats at multiple levels. On the hardware/software side, compromised recording PCs, insecure sensor firmware, or man-in-the-middle attacks on data transmission could expose raw signals. To mitigate these risks, our architecture uses encrypted channels, local

preprocessing, and minimal raw data retention. At the data level, adversaries could attempt to reconstruct sensitive information such as passwords or personal health traits from physiological streams. At the personal level, physiological data itself could reveal individual vulnerabilities, such as heightened susceptibility to phishing when under stress. While our system aims to use such signals for supportive interventions, in a malicious context they could be weaponized for targeted social engineering. Recognizing these risks informs both our technical safeguards and our recommendations on deployment boundaries.

4.2.2 Data Collection: Recording PC. Each workstation is equipped with physiological recording devices connected to a dedicated PC, which controls the recording and provides access to the data. These PCs communicate with a server that collects and stores the data. A virtual private network facilitates data transfer between these PCs and the central server, ensuring isolation from the users' productivity network. We developed a Java application that aggregates all data streams and periodically sends them compressed to the server. We wrote Python scripts that connect the eye tracker and RGB camera using the *tobii-research packages*[1] and *deepface*[2] packages, respectively, and stream the extracted data to the Java application using tcp sockets. The recording PCs use NTP to ensure accurate timestamps between devices and log every recorded data point with a current timestamp.

To monitor the study remotely, a remote desktop connection allows researchers to check the status of data collection. Periodic secure data transfers ensure that collected physiological data is transferred to university servers for further analysis.

Each PC is equipped with an ESP32 microcontroller connected via USB, which interfaces with a 3-inch touch display and two buttons. This setup allows for experience sampling, either by collecting user input at scheduled intervals throughout the day or in response to security-relevant events detected by the sensors or email infrastructure. Survey questions can be sent via the server, and users' responses are transmitted back to the server for further analysis.

4.2.3 Data Storage. The recording PCs compressed each aggregated dataset and transferred it via an encrypted HTTPS connection to the university-hosted server. Internally, we saved the data in CSV files in subfolders created based on the participant IDs. While time-series databases could improve performance of data processing, this was not necessary at our scale and enabled researchers to directly access and work with the saved data.

4.2.4 Sensing.

Eye Tracker for Gaze Monitoring. Physiological components are designed to be interchangeable and adaptable, allowing flexibility based on the specific study requirements, necessary measurements, data protection considerations, and user preferences.

For eye-tracking, we use a Tobii Pro Spark, which is mounted on the user's desk using a custom-designed, 3D-printed mounting case. This setup allows precise alignment with the user's workstation while ensuring easy removal after the study. To address privacy concerns, participants were provided with a tangible privacy cover, enabling them to block the eye tracker at any time physically.

The eye tracker records gaze data at 60 Hz, logging data when the user is present. Since the recording PC is not directly connected to the user's productivity PC or display, conventional eye-tracker calibration methods cannot be used. Instead, we apply a standard calibration procedure and refine the data using an averaged calibration approach, which estimates the display borders based on the aggregated dataset.

RGB Camera for Emotion Detection. We integrated an RGB camera module mounted directly onto the eye tracker (Figure 3), ensuring identical alignment. The camera captures images at 24 Hz, allowing synchronized recording of facial data. To analyze user expressions, we use DeepFace [21] to extract facial features and detect dominant emotions. For enhanced control, the provided privacy cover can conceal the RGB camera, enabling users to manage when data is captured. All processing is handled on the recording PCs themselves, while only the extracted emotional data is sent to the server. Recorded images are immediately deleted after processing.

Wearables for Detecting Heart Rate and Galvanic Skin Response. Our setup supports the integration of Bluetooth wristbands, such as the Garmin Instinct 2, and more precise yet less convenient chest straps, like the Polar H10. The latter could also indirectly measure the user's galvanic skin response in the chest area. However, its accuracy in this use case has yet to be fully evaluated.

To ensure reliable data tracking, we recorded the serial numbers of each heart rate monitoring device and stored them on an ESP32 microcontroller. When a registered wristband or chest strap is within range, the devices automatically establish a connection and transfer data. The ESP32 microcontroller then transfers the data to the PC, which aggregates and transfers it securely to the study server.

Environment Sensors for Context Assessment. We integrated several sensors with the ESP32, including temperature, humidity, multiple brightness sensors to capture light levels from different angles, and a CO2 sensor to monitor air quality. This environmental data helps assess whether the workspace conditions are optimal for the user's performance. If conditions are suboptimal, the system can dynamically adjust its UI behavior, such as delaying security-critical tasks or providing additional assistance until a more favorable environment is established. In addition, this data can be used to actively inform users about their workspace conditions and encourage them to make improvements, such as opening a window for better air circulation.

4.2.5 Security Behavior Elicitation: Email Distribution and Tracking. We used a modified version of the open-source tool GoPhish [43], which enables us to customize additional aspects of each simulated phishing email, including headers and metadata, to reduce the likelihood of detection by spam or phishing filters. Each email contains a tracking pixel and an embedded remote resource linked to our server. When a participant opens the email, their email client attempts to load this resource, allowing us to log the event. However, this method is only effective if the user's email client is configured to load remote resources. Similarly, tracking clicks on embedded links is achieved by directing users to a resource hosted

[1] https://pypi.org/project/tobii-research/
[2] https://pypi.org/project/deepface/

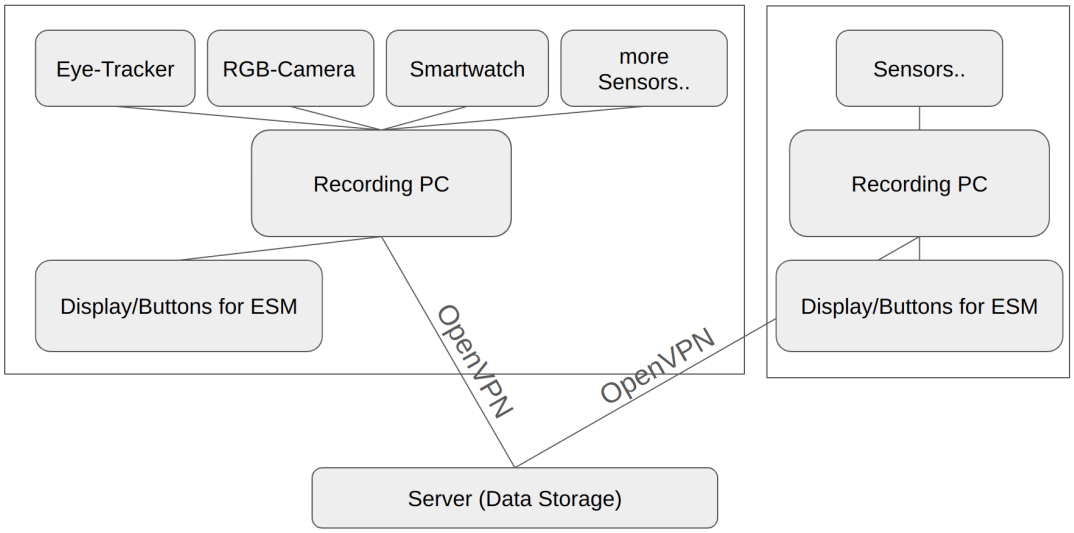


Figure 2: Architectural overview of the different soft- and hardware components.

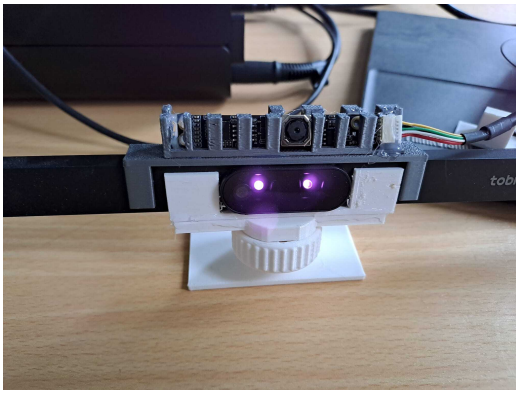


Figure 3: 3D-printed mount of for the eye-tracker and RGB camera. Both sensors have the same alignment and can be rotated on an embedded ball joint. The mount for the RGB camera includes gaps for heat dissipation. All individual 3D printed parts are designed to be printed without support and can be finally assembled using hot glue. We additionally used double-sided tape for steady positioning on the participants desk.

on our email infrastructure, which can replicate legitimate login pages. While GoPhish provides functionality to log all user input, including usernames and passwords, we deliberately disabled this feature for ethical and data protection reasons.

5 Deployment and User Study

To evaluate users' real-world behaviors when exposed to phishing attempts, we conducted a two-week deployment of our platform at a Germany-based cybersecurity consulting firm. Our partner organization facilitated a dedicated email-sending infrastructure to bypass spam filters for simulated phishing messages, enabling us to observe participants' unguarded responses. We recruited 8 employees spanning technical and non-technical roles but initially framed the study as investigating workplace stress rather than phishing detection, thereby minimizing potential bias. Over the course of the study, we discreetly sent three phishing emails to capture participants' physiological and behavioral reactions in situ. In addition, we interviewed participants both before and after the deployment to assess acceptance and perceptions of the sensors. After data collection, we fully debriefed them about the study's true purpose to ensure transparency and ethical compliance. Participants participated in our study during their regular paid working hours thus did not receive additional remuneration. We conducted the study in a country which does not require IRB approval, although we followed our universities' ethic guidelines.

5.1 Setup and Procedure

We evaluate how eye tracking and camera-based emotion recognition could be integrated into everyday workflows without disrupting routine tasks. Each participant's workstation was equipped with a Tobii Pro Spark eye tracker, aligned to cover one primary monitor (participants typically used two or more screens). The Tobii Pro software provided positioning guidelines, and we asked participants to sit in their usual work posture for optimal calibration.

5.2 Multi-Stakeholder Perspectives

Employees. Day-to-day users expressed curiosity and mild apprehension about the platform's ability to track gaze data. Although we framed the study as a stress-related intervention, several participants raised questions about data privacy and the potential for "surveillance." However, a large majority were willing to participate once they were assured that raw footage would not be stored and that personal identifiers would be removed wherever possible.

Follow-up interviews revealed that most employees found the sensors non-disruptive and had largely forgotten about them during daily work.

IT Department. The IT team oversaw local network configurations and sensor software. Their concerns involved ensuring that the platform's network traffic would not burden other services and that collected data were stored securely. They noted that integrating eye trackers into existing hardware setups was straightforward - although they flagged potential scalability issues if future expansions required more sensors per employee (e.g., multiple eye trackers). From a security standpoint, they appreciated the system's potential for "just-in-time" phishing alerts but also highlighted the importance of robust privacy safeguards to avoid internal backlash.

Management. Company leadership viewed the pilot as an opportunity to test novel security solutions and to position their firm at the forefront of user-centered cybersecurity. Still, their willingness hinged on transparent communication about data handling and strong consent protocols. Management liked that we conducted pre- and post-study briefings, citing this practice as vital for building trust. They also valued our gradual platform introduction, starting with minimal sensing (eye tracking, camera) before adding wearables or other modalities.

5.3 Reflections and Ethical Considerations

Overall, the pilot deployment uncovered key insights regarding acceptance and feasibility. Participants did not feel overtly monitored once they understood that all gaze and facial data would be processed locally on the devices and no raw images would be transmitted or stored. A few participants did note that the presence of cameras could affect behavior in sensitive tasks (e.g., handling personal emails). Nonetheless, this initial pilot suggests that targeted, privacy-conscious physiological monitoring can be introduced without significant pushback.

By engaging employees, IT staff, and management, we obtained a holistic view of challenges - ranging from device calibration to potential privacy and policy conflicts. This experience supports the notion that a multi-stakeholder focus is essential for deploying physio-behavioral security platforms at scale.

6 Lessons Learned

Designing and deploying a platform that leverages physiological and behavioral data for adaptive security presented practical and conceptual challenges. This section synthesizes our experiences from prototype development to real-world deployment, spanning technical and human-centric concerns. We highlight how sensor integration and calibration issues impacted system design, discuss the user experience implications of monitoring sensitive signals, and examine the balance of stakeholder interests within the deploying organization. Each subsection offers concrete recommendations to guide future efforts, whether in research labs or enterprise environments.

6.1 Development

Creating a unified platform for eye trackers, webcams, smartwatches, and other sensors posed unique engineering and design challenges. Differences in sampling rates, noisy signals, and diverse organizational infrastructures required careful attention to synchronization, data quality, scalability, and forward compatibility. This section outlines four main obstacles, each followed by a recommendation for future implementations.

6.1.1 Sensor Synchronization. Physiological and behavioral devices generate data at varying latencies and frequencies. Slight misalignments (e.g., a few milliseconds off) can cause gaze events to mismatch with corresponding changes in heart rate or facial expression, skewing subsequent analyses. We addressed this by implementing time-alignment routines that standardized timestamps across devices, retriggering them whenever a new sensor was added or network conditions changed.

Recommendation: Adopt a robust synchronization protocol that regularly cross-checks and updates sensor clocks. This approach ensures that moment-to-moment user states remain accurately correlated.

6.1.2 Noise and Artifact Handling. Physiological signals are inherently prone to artifacts from environmental variability (e.g., lighting fluctuations, seating changes) and transient sensor disconnections (e.g., a smartwatch losing Bluetooth). We integrated filtering pipelines that automatically flagged and replaced anomalous readings, reducing the burden on downstream processes dependent on stable input.

Recommendation: Incorporate automated artifact detection and error-handling modules that either correct or gracefully discard corrupted data. Where possible, use statistical methods resilient to outliers.

6.1.3 Scalability and Interoperability. Organizations differ widely in hardware, network configurations, and security policies. We opted for standardized APIs and modular drivers, allowing us to attach or remove sensors with minimal disruptions. Extension points for existing IT systems (e.g., email gateways) also facilitated seamless deployment in heterogeneous environments.

Recommendation: Structure the platform around modular, well-documented interfaces that support straightforward integration into diverse technical ecosystems. Early collaboration with IT teams can streamline configuration and gain buy-in from security administrators.

6.1.4 Preparing for Future Expansion. As sensor technology advances and user needs evolve, systems must flexibly incorporate new modalities. We used a multi-tiered design that processes data locally while aggregating high-level features in a central repository. This architecture reduces network load and accommodates additional sensing devices without extensive re-engineering.

Recommendation: Anticipate growth by distributing computational tasks across local nodes and retaining only essential aggregated metrics in a secure store. This design mitigates latency and scales with both new sensors and increasing user bases.

6.2 Deployment

Introducing our physiology-based platform into a real-world work environment revealed a range of practical, organizational, and

human-centric considerations. On a technical level, sensor calibration emerged as a major challenge. Although precise calibration can yield higher data fidelity, it also increases the burden on participants, who often expect minimal disruption to their regular tasks. We also encountered hardware constraints, such as desk layouts with multiple monitors and inconsistent lighting conditions, which necessitated adaptive calibration strategies and frequent configuration checks. Organizational factors, such as local IT policies and employees' privacy expectations, further influenced our approach - underscoring the need to integrate seamlessly with existing infrastructures and to address user concerns transparently.

Below, we present four recommendations arising from these observations. Each recommendation summarizes the rationale behind our design decisions and highlights how future deployments can generalize lessons learned.

6.2.1 Minimize Calibration Overhead. Sensor calibration can affect user acceptance. While a multi-step process might enhance data accuracy (e.g., better gaze tracking), participants often view it as an added task, interrupting their workflow or daily routines. In our deployment, employees showed a clear preference for brief, automated calibration and for using metrics inherently tolerant of minor misalignments (e.g., fixation length or pupil dilation) over metrics such as absolute gaze position, requiring meticulous calibration.

Recommendation: Deploy sensors and metrics that demand minimal user intervention. Where possible, conduct automated calibration or use passive indicators (like pupil dilation) that do not require repeated user actions.

Reducing calibration steps shortens onboarding time, encourages participation from less tech-savvy users, and lowers the risk of calibration drift during prolonged usage. This approach ultimately expands the range of contexts where physiology-based sensing can be seamlessly introduced.

6.2.2 Design for Environmental Variability. Office environments are rarely uniform. Lighting conditions can change over the course of the day, seats and monitors can be rearranged, and employees may frequently switch tasks or devices. These variations can degrade the accuracy of vision-based sensing (e.g., eye tracking, facial expression analysis) and introduce artifacts in physiological data (e.g., from changes in posture or desk height).

Recommendation: Develop adaptive or periodic re-calibration protocols and consider integrating additional environmental sensors (e.g., ambient light detectors) to adjust image-processing parameters automatically.

Proactive adaptation ensures the system maintains robust performance under changing conditions. It also offloads effort from end users, who otherwise would need to manually re-calibrate or troubleshoot upon changes.

6.2.3 Respect User Constraints. Employees' concerns about privacy and workload shaped their openness to physiological monitoring. Many worried about being constantly recorded or believed physiological data might be used to evaluate their job performance. Our pilot suggested these worries subside when users understand how data is handled, can pause or disable sensors, and see a tangible benefit from the system.

Recommendation: Combine transparent consent mechanisms with clear communication of benefits. Provide users with the ability to opt-out or temporarily disable sensors and clarify what data is recorded, how it is stored, and who has access.

Giving employees control and reassurance fosters trust, lowers the risk of backlash, and can increase engagement. Genuine user buy-in is vital for sustained real-world use - particularly in contexts involving sensitive data like stress or emotional states.

6.2.4 Challenges of Integrating with Existing Applications. Although our platform enables us to capture physiological and behavioral data, integrating adaptive security interventions directly into core productivity tools, such as modifying the email interface to highlight suspicious links, proved significantly more difficult. Organizational policies often restrict third-party modifications to email clients, and IT departments must balance research goals against reliability, security, and employee productivity. These constraints limited our ability to experiment with dynamic UI elements or gaze-aware interaction cues in situ. Future work may explore sandbox simulations, staged roll-outs, or working with software vendors to enable such interventions.

6.2.5 Reliability challenges. Because the deployment was our initial real-world pilot, we limited the sensing modalities to eye tracking and an RGB camera for facial analysis. Since participants used multiple monitors and the Tobii Pro Spark eye tracker featured a narrow field of view, eye tracking frequently dropped out when participants moved their heads or shifted their attention across screens. In addition, several participants changed their workplace during the study and reassembled the sensors without proper recalibration, which further decreased data accuracy. We encouraged participants to behave naturally to ensure ecological validity, but these factors led to a substantial amount of incomplete or inconsistent gaze and facial data. As a result, we were not able to reliably correlate physiological signals with phishing email interactions. In future work, we plan to extend the setup to multiple eye trackers (one per monitor) or devices with a wider field of view to capture a more complete picture of user attention patterns. During the study, employees carried out their normal tasks while wearing no additional devices. After the study we debriefed participants and conducted short semi-structured interviews to address their experiences and potential concerns.

Intermediate Summary. Taken together, these recommendations reflect the multifaceted realities of deploying a physio-behavioral security system. Although our observations stem from a single organizational setting, the core principles of minimizing calibration, designing for changing environments, aligning with existing infrastructure, and respecting user constraints are broadly applicable to other domains - from education to healthcare to government agencies. By proactively considering these dimensions in new deployments, developers can avoid common pitfalls, streamline adoption, and ultimately deliver more reliable, user-centered security solutions. One additional challenge was the difficulty of modifying existing interfaces such as email clients, which are often tightly controlled by corporate IT policies - highlighting the practical gap between conceptual interventions and real-world implementation constraints.

6.3 User View

Users' perceptions and willingness to adopt physiological monitoring emerged as crucial in determining the system's overall acceptance. Privacy concerns and fears of surveillance surfaced early on, with some participants worrying that data might be used to evaluate performance or scrutinize personal habits. Over time, transparent communication about how data would be collected, analyzed, and stored effectively alleviated most apprehensions.

> "In the beginning I noticed the sensors a lot, but after a few days I fell accustomed to it since I it is just another camera around me anyways (besides the normal webcam on my PC). One behavior I did change, however, is that I did not eat at my desk anymore since I felt watched doing that."

One participant was particularly concerned about client data:

> "It was important for us that no audio data was recorded, since we have a lot of confidential client data. For example, when we directly talk to clients or colleagues. I would not worry too much about my personal data, since I mostly talk about personal stuff in the coffee kitchen."

Users particularly valued the ability to pause or disable monitoring devices, such as covering the camera lens or deactivating eye tracking, which bolstered their sense of control.

> "I found the onboarding process good and necessary. I also found the research context with the involvement of a university reassuring."

Opinions on the potential benefits of the system varied. Although some participants appreciated adaptive security alerts (for example, additional prompts when they appeared inattentive), others considered them intrusive, especially if false positives interrupted daily workflows. One participant mentioned his previous experience with smart tech:

> "I am not a fan of "High-End devices", for example my washing machine that wants to tell me using AI how I should wash my laundry. I would never use that, I think I could decide better myself what is correct."

This balance between proactive defense and minimal disruption appeared to be a key factor in building long-term trust and engagement.

Participants also mentioned their altered work environment:

> "We have limited space on our desk so this additional thing took even more space away."

Below, we outline three core recommendations to maximize user acceptance and comfort with physio-behavioral security platforms.

6.3.1 Provide Transparent Data Handling. Many users fear misuse or accidental exposure of highly personal physiological data. Clear disclosures regarding data storage, anonymization or aggregation, and permitted uses went a long way toward easing these worries.

Recommendation: Publicize policies detailing how, when, and why the system collects data, and give users easy-to-read summaries of retention periods and access rights.

Being upfront about data handling practices builds credibility and reassures participants that their privacy remains a priority. This transparency can also reduce speculative fears about covert surveillance or hidden data-sharing practices.

6.3.2 Offer Direct User Control. A recurring theme in user feedback was the desire for personal agency. Even if the data collection seemed benign, users favored having the option to disable monitoring temporarily - such as toggling off eye tracking.

Recommendation: Integrate opt-out or pause features, and provide physical means to block sensors (e.g., camera covers)

Granting control over data capture not only respects personal comfort levels but can also boost user buy-in, as participants perceive the system as voluntary rather than imposed.

6.3.3 Design Minimally Intrusive Interventions. Adaptive security alerts should align with user workflow and cognitive load. Even well-intentioned warnings can feel disruptive if triggered too frequently or at inopportune moments - particularly if false positives are common.

Recommendation: Personalize or throttle interventions to occur only when the risk is high or users appear genuinely inattentive. Provide customization settings that let individuals fine-tune intervention frequency and intensity.

Overly invasive or misaligned interventions can erode trust and lead to alert fatigue, undermining the system's goal of improving security.

Intermediate Summary. By combining transparent data handling, user-driven controls, and thoughtfully timed interventions, practitioners can mitigate common concerns around privacy and autonomy. Although individual preferences may differ across workplaces or cultural contexts, these strategies form a scalable foundation for user-centric deployment. When users feel respected and genuinely benefit from adaptive security measures, they are more likely to engage meaningfully, thereby amplifying the system's overall effectiveness.

6.4 Stakeholder View

Our deployment revealed differing priorities across stakeholders: IT, management, and employees. While all saw value in physiological monitoring for security, concerns varied: IT teams focused on technical integration, processing overhead, and maintenance; management emphasized strategic benefits, ROI, and privacy safeguards; and employees were split between appreciating real-time alerts and fearing surveillance or misinterpretation of data. A key challenge was separating stress from legitimate threats versus everyday job pressures. The following recommendations aim to balance these concerns and guide stakeholder-aligned implementation of physio-behavioral security.

6.4.1 Institute Iterative Feedback Loops. Stakeholders may have evolving or even conflicting needs over time. Regular check-ins can reveal overlooked issues (e.g., rising maintenance costs or employee discomfort with a specific sensor) and support agile adjustments.

Recommendation: Schedule periodic reviews involving IT staff, management, and user representatives to assess data governance, usability concerns, and evolving security requirements.

This inclusive process helps maintain alignment, increases buy-in, and surfaces implementation challenges early, reducing the risk of large-scale user resistance or technical dead ends.

6.4.2 Clarify and Contextualize Physiological Signals. Real-world deployments require clearly communicating what physiological data can and cannot reveal. Many false alarms or misinterpretations arise from normal fluctuations in stress unrelated to security threats.

Recommendation: Couple physiological metrics with contextual cues - such as user-initiated feedback or system logs - to avoid conflating legitimate security risks with routine job stressors.

Contextualizing signals fosters trust among employees, reassuring them that data-driven security decisions are informed by multiple sources rather than simplistic or purely automated interpretations.

6.4.3 Balance Security Gains with Operational Feasibility. While management may support innovative solutions, IT teams must ensure sustainable integration, and employees must not feel overwhelmed. Management may hesitate to adopt additional hardware due to higher costs and commitment, preferring software-only options that are easier to deploy and remove. However, this can reduce detection accuracy when relying solely on existing inputs such as mouse movements, typing behavior, or gaze tracking through standard webcams.

Recommendation: Invest in robust but lightweight architectures that minimize system overhead, ensure easy maintenance, and integrate seamlessly with existing tools. Evaluate return on investment through pilots that quantify both security improvements and operational costs.

A balanced solution that delivers tangible security benefits without overwhelming technical or human resources is more likely to gain lasting acceptance across the organization.

6.5 Utopian and Dystopian Scenarios

Looking beyond our deployment, physio-behavioral security can be imagined along two very different trajectories. In a utopian vision, the platform enhances security while remaining largely invisible to users. Interventions would occur only at the right moments – for example, highlighting a suspicious email when a user is distracted, or deferring a password change until workload is low. In such a scenario, users feel supported rather than burdened, and organizations benefit from reduced incidents without compromising trust. Security becomes a cooperative partner, adapting seamlessly to the rhythms of daily work.

At the same time, a dystopian vision is easy to imagine. The very same signals that enable adaptive interventions could also be repurposed for surveillance and control. Physiological measures of stress or attention could be used to evaluate productivity, enforce compliance, or even discipline employees. In contexts with strong power asymmetries, such as schools, prisons, or authoritarian regimes, this misuse could extend far beyond security, turning physiological monitoring into a tool of coercion. Even in corporate settings, function creep may gradually transform a supportive security tool into a mechanism of performance evaluation.

These dual scenarios highlight the importance of drawing clear boundaries. We argue that the platform should **not** be implemented in environments where meaningful consent cannot be given, or where data is likely to be exploited for non-security purposes. Physiological signals should never become inputs for managerial oversight, hiring or firing decisions, or political surveillance. Instead, the system's use should remain confined to contexts where transparency, privacy safeguards, and user autonomy can be assured.

Intermediate Summary. Addressing the needs of IT, management, and users calls for a holistic, collaborative technology adoption model. By establishing iterative feedback loops, contextualizing physiological data, and actively weighing security benefits against deployment complexity, organizations can tailor solutions respecting stakeholder concerns. This multi-perspective approach not only eases initial rollouts but also enhances long-term sustainability, ensuring physio-behavioral security becomes a trusted component of organizational defenses. At the same time, considering both utopian and dystopian scenarios reminds us that the very same capabilities that enable supportive, well-timed interventions could also be misused for surveillance or control, underscoring the importance of strong governance and clear boundaries for deployment.

7 Adaptive User Interfaces

While our platform focuses on real-time physiological and behavioral signal acquisition and processing, a critical next step is to get user interfaces to react meaningfully in response to inferred user state. Leaning on our data streams and on principles established in the physio-behavioral security paradigm, we propose two forms of adaptive interventions: **(1) general state-based adaptations**, responding to long-term user states such as fatigue or stress, and **(2) in-situ reactive changes**, responding to present or sudden changes in user physiology or behavior while engaged in security-related tasks.

7.1 General User State-Based Adaptions

To move beyond passive sensing, our platform enables interfaces that adapt to user states like fatigue, stress, or inattention, thus avoiding security-critical tasks to when users are most receptive or at risk.

Phishing Email Timing: When signs of fatigue or low attention are detected, the system can hold back potentially suspicious emails until the user is more alert, helping prevent inattentive interactions.

Deferred Security Prompts: Tasks such as password changes, software updates, or configuration dialogs can be postponed if the user's physiological signals indicate cognitive overload, frustration, or poor environmental conditions (e.g., high CO_2, low light).

7.2 In-Situ Reactive UI Changes During Security Tasks

Reactive interfaces respond to sudden changes in the state of the user during security tasks, offering timely nudges to reduce risks.

Phishing Awareness Support: If eye tracking reveals a lack of fixation on sender details or links in an email, the interface can reactively highlight these elements, slow link activation, or request confirmation before interaction.

Password Reuse Nudging: Based on gaze and typing dynamics, the system can infer the likelihood of password reuse even before text input is complete [6]. A gentle in-line prompt can then suggest stronger alternatives or trigger the password generator.

Stress-Triggered Flow Interruption: Sudden spikes in heart rate variability or electrodermal activity during sensitive tasks (e.g., email handling or device setup) can trigger UI pauses, simplified views, or an option to defer the task.

Emotion-Aware Guidance: Real-time facial analysis (via the RGB camera) can detect confusion or frustration and proactively offer in-context explanations, visual cues, or calming interface adjustments.

7.3 Implementation Considerations

These examples show how adaptive interfaces can be integrated into our platform, but practical use requires balancing support and intrusiveness. Although our architecture enables technical integration, our current deployment did not include such interventions due to inconsistent sensor reliability in real-world conditions. Future work should refine UX patterns, escalation logic, and personalization. Interventions must remain transparent, overrulable, and minimally disruptive to maintain user autonomy and privacy.

8 Future Work

Merging Data Streams and Real-Time Interventions. This research provides a solid foundation for collecting and processing physiological data in real-world conditions. Moving forward, a key priority is combining diverse data streams into a unified, real-time monitoring and intervention system. Leveraging edge computing would minimize latency and strengthen privacy by retaining sensitive sensor data on the user's device rather than transmitting everything to the cloud.

Expanding the Range of Physiological Signals. Additional modalities such as EEG or voice-based stress analysis could refine user state estimation and expand the coverage of cognitive and emotional cues. Integrating consumer wearables, for instance, fitness trackers, might also yield more context-aware interventions that account for daily rhythms or physical activity. These enhancements could better personalize security responses without imposing excessive user burden.

Overcoming Multi-Monitor Constraints. One limitation observed during our deployment stemmed from single-eye-tracker setups in multi-monitor workspaces, which left gaps in gaze data whenever participants shifted focus to a secondary screen. Future iterations should explore multi-tracker configurations and develop new methods for calibrating and synchronizing gaze streams across devices. This approach will require further testing to evaluate feasibility, computational overhead, and user acceptance in productivity-driven settings.

Personalizing Security Interventions. Scalability depends on tailoring security mechanisms to individual users, based on past behavior and physiological responses. Adapting the frequency and intensity of interventions according to user context and preferences can prevent intrusiveness while preserving effectiveness. This direction includes implementing granular privacy controls that let users selectively opt in or out of specific sensors, fostering trust and adoption in heterogeneous organizations.

Enhancing Usability and Engagement. Effective alerting mechanisms should prioritize critical warnings while suppressing trivial ones, reducing the risk of alert fatigue or user desensitization. Gamified or interactive security training - integrated with real-time physiological feedback - could further reinforce positive security habits and boost user engagement, making the learning experience more intuitive and less disruptive.

Detecting Insider Threats. Another research avenue involves examining whether shifts in stress or cognitive load patterns can predict malicious behavior from within an organization. By identifying precursors to critical events such as data leaks or unauthorized access, physio-behavioral indicators might inform proactive risk models and help organizations intervene before damage occurs.

Scaling Up Through Field Studies. Finally, large-scale field evaluations in corporate and other professional contexts are essential for validating physiological security over extended periods. Such studies would reveal how variations in user behavior, device configurations, and network environments affect system performance and acceptance. Building on these insights, future research can drive the advancement of more adaptive, user-centric, and privacy-preserving security solutions that embed seamlessly into everyday workflows.

9 Conclusion

This work presents a platform that integrates physiological and behavioral signals to enhance security mechanisms through adaptive, context-aware interventions. By leveraging multi-modal data sources such as gaze tracking, heart rate variability, and keystroke dynamics, our approach enables real-time assessment of user states, allowing security measures to be dynamically tailored to individual cognitive and emotional conditions. Our exploratory deployment in a real-world setting demonstrates the potential of this approach to improve security effectiveness while maintaining user acceptance.

However, our findings also highlight key challenges, including the need for robust data processing to handle physiological signal variability, user concerns regarding privacy and transparency, and the complexities of integrating such systems into existing IT infrastructures. Our findings further underscore that while activation-related states such as attention and stress can be captured with reasonable reliability, inferences about valence, intent, or broader mental states should be treated with caution. Future research should focus on refining adaptive security mechanisms, ensuring privacy-preserving data processing through edge computing, and conducting longitudinal studies to evaluate long-term user acceptance and effectiveness.

Ultimately, this research highlights the potential of physiology-aware security solutions to bridge the gap between human factors and cybersecurity. By adapting security interventions to real-time user states, such approaches can enhance protection without imposing unnecessary friction. However, their success depends on balancing effectiveness with privacy, ensuring transparent user control, and integrating seamlessly into existing security frameworks. At the same time, our work underscores that physiology-based warning systems demand very high accuracy to avoid false positives and user fatigue. As a result, pre-emptive, context-aware interventions may be the more feasible application in the near term.

We hope our work will pave the way for more research in this area. We plan to release our platform for other researchers to use.

Acknowledgments

The presented work received funding from the German Research Foundation (DFG) under project no. 425869382 and Technology Research Center of the Bundeswehr (Voice of Wisdom). dtec.bw is funded by the European Union - NextGenerationEU. The authors used AI-Tools such as Grammarly and ChatGPT to improve grammar, punctuation, and vocabulary. We are also grateful to the anonymous reviewers from the NSPW program committee for their constructive input, and to Anil Somayaji for his guidance throughout the shepherding process and for helping us sharpen and more clearly communicate the key ideas of this work. Lastly, we thank the workshop attendees for their invaluable feedback.

References

[1] Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay Cool! Understanding Thermal Attacks on Mobile-based User Authentication. In *Proceedings of the 35th Annual ACM Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA. doi:10.1145/3025453.3025461

[2] Yomna Abdelrahman, Eduardo Velloso, Tilman Dingler, Albrecht Schmidt, and Frank Vetere. 2017. Cognitive Heat: Exploring the Usage of Thermal Imaging to Unobtrusively Estimate Cognitive Load. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 33 (sep 2017). doi:10.1145/3130898

[3] Yasmeen Abdrabou, Felix Dietz, Ahmed Shams, Pascal Knierim, Yomna Abdelrahman, Ken Pfeuffer, Mariam Hassib, and Florian Alt. 2023. Revealing the Hidden Effects of Phishing Emails: An Analysis of Eye and Mouse Movements in Email Sorting Tasks. arXiv.org. doi:10.48550/arXiv.2305.17044Focustolearnmore arXiv:2305.17044 [cs.HC] abdrabou2023arxiv.

[4] Yasmeen Abdrabou, Elisaveta Karypidou, Florian Alt, and Mariam Hassib. 2023. Investigating User Behaviour Towards Fake News on Social Media Using Eye Tracking and Mouse Movements. In *Proceedings of the Usable Security Mini Conference 2023 (USEC'23)*. Internet Society, San Diego, CA, USA. doi:10.14722/usec.2023.232041 abdrabou2023usec.

[5] Yasmeen Abdrabou, Johannes Schütte, Ahmed Shams, Ken Pfeuffer, Daniel Buschek, Mohamed Khamis, and Florian Alt. 2022. "Your Eyes Tell You Have Used This Password Before": Identifying Password Reuse from Gaze and Keystroke Dynamics. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 400, 16 pages. doi:10.1145/3491102.3517531

[6] Yasmeen Abdrabou, Ahmed Shams, Mohamed Omar Mantawy, Anam Ahmad Khan, Mohamed Khamis, Florian Alt, and Yomna Abdelrahman. 2021. GazeMeter: Exploring the Usage of Gaze Behaviour to Enhance Password Assessments. In *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Full Papers)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3448017.3457384

[7] Anne Adams and Martina Angela Sasse. 1999. Users are not the enemy. *Commun. ACM* 42, 12 (Dec. 1999), 40–46. doi:10.1145/322796.322806

[8] Abdulaziz Almehmadi. 2021. Micro-Behavioral Accidental Click Detection System for Preventing Slip-Based Human Error. *Sensors* 21, 24 (2021). doi:10.3390/s21248209

[9] Florian Alt, Mariam Hassib, and Verena Distler. 2023. Human-centered Behavioral and Physiological Security. In *Proceedings of the 2023 Workshop on New Security Paradigms (NSPW '23)*. ACM, New York, NY, USA.

[10] Bonnie Brinton Anderson, Anthony Vance, C. Brock Kirwan, Jeffrey L. Jenkins, and David Eargle. 2016. From Warning to Wallpaper: Why the Brain Habituates to Security Warnings and What Can Be Done About It. *Journal of Management Information Systems* 33, 3 (July 2016), 713–743. doi:10.1080/07421222.2016.1243947

[11] Majid Arianezhad, L. Jean Camp, Timothy Kelley, and Douglas Stebila. 2013. Comparative Eye Tracking of Experts and Novices in Web Single Sign-On. In *Proceedings of the Third ACM Conference on Data and Application Security and Privacy (CODASPY '13)*. ACM, New York, NY, USA, 105–116. doi:10.1145/2435349.2435362

[12] Syed Arshad, Yang Wang, and Fang Chen. 2013. Analysing Mouse Activity for Cognitive Load Detection. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (OzCHI '13)*. ACM, New York, NY, USA, 115–118. doi:10.1145/2541016.2541083

[13] Adam J. Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M. Smith. 2010. Smudge Attacks on Smartphone Touch Screens. In *Proceedings of the 4th USENIX Conference on Offensive Technologies (WOOT'10)*. USENIX Association, USA, 1–7.

[14] Rummana Bari, Md. Mahbubur Rahman, Nazir Saleheen, Megan Battles Parsons, Eugene H. Buder, and Santosh Kumar. 2020. Automated Detection of Stressful Conversations Using Wearable Physiological and Inertial Sensors. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 117 (Dec. 2020), 23 pages. doi:10.1145/3432210

[15] Paul Bekaert, Norah Alotaibi, Florian Mathis, Nina Gerber, Aidan Christopher Rafferty, Mohamed Khamis, and Karola Marky. 2022. Are Thermal Attacks a Realistic Threat? Investigating the Preconditions of Thermal Attacks in Users' Daily Lives. In *Nordic Human-Computer Interaction Conference (NordiCHI '22)*. ACM, New York, NY, USA, Article 76, 9 pages. doi:10.1145/3546155.3546706

[16] David Guy Brizan, Adam Goodkind, Patrick Koch, Kiran Balagani, Vir V. Phoha, and Andrew Rosenberg. 2015. Utilizing linguistically enhanced keystroke dynamics to predict typist cognition and demographics. *International Journal of Human-Computer Studies* 82 (2015), 57–68. doi:10.1016/j.ijhcs.2015.04.005

[17] Ulrich Burgbacher and Klaus Hinrichs. 2014. An Implicit Author Verification System for Text Messages Based on Gesture Typing Biometrics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2951–2954. doi:10.1145/2556288.2557346

[18] Daniel Buschek, Benjamin Bisinger, and Florian Alt. 2018. ResearchIME: A Mobile Keyboard Application for Studying Free Typing Behaviour in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 1–14. doi:10.1145/3173574.3173829

[19] Daniel Buschek, Alexander De Luca, and Florian Alt. 2015. Improving Accuracy, Applicability and Usability of Keystroke Biometrics on Mobile Touchscreen Devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1393–1402. doi:10.1145/2702123.2702252

[20] Daniel Buschek, Alexander De Luca, and Florian Alt. 2016. Evaluating the Influence of Targets and Hand Postures on Touch-based Behavioural Biometrics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1349–1361. doi:10.1145/2858036.2858165

[21] Hardie Cate, Fahim Dalvi, and Zeshan Hussain. 2017. Deepface: Face generation using deep learning. *arXiv preprint arXiv:1701.01876* (2017).

[22] Rui Chen, Tiantian Xie, Yingtao Xie, Tao Lin, and Ningjiu Tang. 2016. Do Speech Features for Detecting Cognitive Load Depend on Specific Languages?. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16)*. ACM, New York, NY, USA, 76–83. doi:10.1145/2993148.2993149

[23] Francesco Chiossi, Robin Welsch, Steeven Villa, Lewis Chuang, and Sven Mayer. 2022. Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience. *Big Data and Cognitive Computing* 6, 2 (2022). doi:10.3390/bdcc6020055

[24] Burcu Cinaz, Bert Arnrich, Roberto La Marca, and Gerhard Tröster. 2013. Monitoring of mental workload levels during an everyday life office-work scenario. *Personal and ubiquitous computing* 17 (2013), 229–239.

[25] Heather Crawford. 2010. Keystroke dynamics: Characteristics and opportunities. In *2010 Eighth International Conference on Privacy, Security and Trust*. 205–212. doi:10.1109/PST.2010.5593258

[26] Alexander De Luca, Alina Hang, Frederik Brudy, Christian Lindner, and Heinrich Hussmann. 2012. Touch Me Once and i Know It's You! Implicit Authentication Based on Touch Screen Patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 987–996. doi:10.1145/2207676.2208544

[27] Alexander De Luca, Roman Weiss, and Heiko Drewes. 2007. Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN-Entry. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI '07)*. ACM, New York, NY, USA, 199–202. doi:10.1145/1324892.1324932

[28] Verena Distler. 2023. The Influence of Context on Response to Spear-Phishing Attacks: An In-Situ Deception Study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, Article 619, 18 pages. doi:10.1145/3544548.3581170

[29] Verena Distler, Matthias Fassl, Hana Habib, Katharina Krombholz, Gabriele Lenzini, Carine Lallemand, Lorrie Faith Cranor, and Vincent Koenig. 2021. A Systematic Literature Review of Empirical Methods and Risk Representation in Usable Privacy and Security Research. *ACM Trans. Comput.-Hum. Interact.* 28, 6, Article 43 (dec 2021), 50 pages. doi:10.1145/3469845

[30] Christopher Hadnagy. 2010. *Social engineering: The art of human hacking*. John Wiley & Sons.

[31] Yassir Hashem, Hassan Takabi, Mohammad GhasemiGol, and Ram Dantu. 2015. Towards insider threat detection using psychophysiological signals. In *Proceedings of the 7th ACM CCS International Workshop on Managing Insider Security Threats*. 71–74.

[32] Mariam Hassib, Mohamed Khamis, Susanne Friedl, Stefan Schneegass, and Florian Alt. 2017. Brainatwork: Logging Cognitive Engagement and Tasks in the Workplace Using Electroencephalography. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM '17)*. ACM, New York,

NY, USA, 305–310. doi:10.1145/3152832.3152865

[33] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. 2017. EngageMeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5114–5119. doi:10.1145/3025453.3025669

[34] Franziska Herbert, Steffen Becker, Leonie Schaewitz, Jonas Hielscher, Marvin Kowalewski, Angela Sasse, Yasemin Acar, and Markus Dürmuth. 2023. A World Full of Privacy and Security (Mis)conceptions? Findings of a Representative Survey in 12 Countries. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–23. doi:10.1145/3544548.3581410

[35] K. Hercegfi, O. Kiss, Krisztina Bali, and L. Izsó. 2006. Interface: assessment of human-computer interaction by monitoring physiological and other data with a time-resolution of only a few seconds. (2006), 2288–2299.

[36] Christina Katsini, Yasmeen Abdrabou, George E. Raptidis, Mohamed Khamis, and Florian Alt. 2020. The Role of Eye Gaze in Security and Privacy Applications:Survey and Future HCI Research Directions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, US) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3313831.3376840 katsini2020chi.

[37] Ruhul Amin Khalil, Edward Jones, Mohammad Inayatullah Babar, Tariqullah Jan, Mohammad Haseeb Zafar, and Thamer Alhussain. 2019. Speech emotion recognition using deep learning techniques: A review. *IEEE Access* 7 (2019), 117327–117345.

[38] Thomas Kosch, Mariam Hassib, Daniel Buschek, and Albrecht Schmidt. 2018. Look into My Eyes: Using Pupil Dilation to Estimate Mental Workload for Task Complexity Adaptation. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, 1–6. doi:10.1145/3170427.3188643

[39] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Woźniak. 2023. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *ACM Comput. Surv.* 55, 13s, Article 283 (jul 2023), 39 pages. doi:10.1145/3582272

[40] Manu Kumar, Tal Garfinkel, Dan Boneh, and Terry Winograd. 2007. Reducing Shoulder-Surfing by Using Gaze-Based Password Entry. In *Proceedings of the 3rd Symposium on Usable Privacy and Security (SOUPS '07)*. ACM, New York, NY, USA, 13–19. doi:10.1145/1280680.1280683

[41] Judith Ley-Flores, Laia Turmo Vidal, Elena Márquez Segura, Aneesha Singh, Frederic Bevilacqua, Francisco Cuadrado, Joaquín Roberto Díaz Durán, Omar Valdiviezo-Hernández, Milagrosa Sánchez-Martin, and Ana Tajadura-Jiménez. 2024. Co-Designing Sensory Feedback for Wearables to Support Physical Activity through Body Sensations. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 40 (March 2024), 31 pages. doi:10.1145/3643499

[42] Wei Lun Lim, Yisi Liu, Salem Chandrasekaran Harihara Subramaniam, Serene Hui Ping Liew, Gopala Krishnan, Olga Sourina, Dimitrios Konovessis, Hock Eng Ang, and Lipo Wang. 2018. EEG-Based Mental Workload and Stress Monitoring of Crew Members in Maritime Virtual Simulator. In *Transactions on Computational Science XXXII: Special Issue on Cybersecurity and Biometrics*, Marina L. Gavrilova, C.J. Kenneth Tan, and Alexei Sourin (Eds.). Springer, Berlin, Heidelberg, 15–28. doi:10.1007/978-3-662-56672-5_2

[43] Andy Luse and Jim Burkman. 2021. Gophish: Implementing a real-world phishing exercise to teach social engineering. *Journal of Cybersecurity Education, Research and Practice* 2020, 2 (2021), 5.

[44] Aicha Maalej and Ilhem Kallel. 2020. Does keystroke dynamics tell us about emotions? A systematic literature review and dataset construction. In *2020 16th International Conference on Intelligent Environments (IE)*. IEEE, 60–67.

[45] John McAlaney and Peter J. Hills. 2020. Understanding Phishing Email Processing and Perceived Trustworthiness Through Eye Tracking. *Frontiers in Psychology* 11 (2020). doi:10.3389/fpsyg.2020.01756

[46] Varun Mishra, Sougata Sen, Grace Chen, Tian Hao, Jeffrey Rogers, Ching-Hua Chen, and David Kotz. 2020. Evaluating the Reproducibility of Physiological Stress Detection Models. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 147 (Dec. 2020), 29 pages. doi:10.1145/3432220

[47] Daisuke Miyamoto, Takuji Iimura, Gregory Blanc, Hajime Tazaki, and Youki Kadobayashi. 2014. EyeBit: Eye-Tracking Approach for Enforcing Phishing Prevention Habits. In *2014 Third International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS)*. 56–65. doi:10.1109/BADGERS.2014.14

[48] Rosana Montañez, Edward Golob, and Shouhuai Xu. 2020. Human Cognition Through the Lens of Social Engineering Cyberattacks. *Frontiers in Psychology* 11 (2020). doi:10.3389/fpsyg.2020.01755

[49] Sebastian Möller, Noam Ben-Asher, Klaus-Peter Engelbrecht, R. Englert, and Joachim Meyer. 2011. Modeling the behavior of users who are confronted with security mechanisms. *Comput. Secur.* 30 (2011), 242–256. doi:10.1016/j.cose.2011.01.001

[50] Ajaya Neupane, Md Lutfor Rahman, Nitesh Saxena, and Leanne Hirshfield. 2015. A multi-modal neuro-physiological study of phishing detection and malware warnings. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. 479–491.

[51] Dimitra Papatsaroucha, Yannis Nikoloudakis, Ioannis Kefaloukos, Evangelos Pallis, and Evangelos K. Markakis. 2021. A Survey on Human and Personality Vulnerability Assessment in Cyber-security: Challenges, Approaches, and Open Issues. doi:10.48550/arXiv.2106.09986

[52] Kevin Pfeffel, Philipp Ulsamer, and Nicholas H. Müller. 2019. Where the User Does Look When Reading Phishing Mails – An Eye-Tracking Study. In *Learning and Collaboration Technologies. Designing Learning Experiences*, Panayiotis Zaphiris and Andri Ioannou (Eds.). Springer International Publishing, Cham, 277–287.

[53] Kenneth Revett. 2008. *Behavioral Biometrics: A Remote Access Approach*. John Wiley & Sons.

[54] Emils Rozentals. 2021. *Email load and stress impact on susceptibility to phishing and scam emails*. https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-85403

[55] Emils Rozentals. 2021. *Email load and stress impact on susceptibility to phishing and scam emails*. Student Thesis, Lulea, Sweden.

[56] Alia Saad, Michael Chukwu, and Stefan Schneegass. 2018. Communicating Shoulder Surfing Attacks to Users. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM '18)*. ACM, New York, NY, USA, 147–152. doi:10.1145/3282894.3282919

[57] Jessica Schwarz, Sven Fuchs, and Frank Flemisch. 2014. Towards a more holistic view on user state assessment in adaptive human-computer interaction. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1228–1234. doi:10.1109/SMC.2014.6974082

[58] Kshitij Sharma, Evangelos Niforatos, Michail Giannakos, and Vassilis Kostakos. 2020. Assessing Cognitive Performance Using Physiological and Facial Features: Generalizing across Contexts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 95 (Sept. 2020), 41 pages. doi:10.1145/3411811

[59] Luma Tabbaa, Ryan Searle, Saber Mirzaee Bafti, Md Moinul Hossain, Jittrapol Intarasisrisawat, Maxine Glancy, and Chee Siang Ang. 2022. VREED: Virtual Reality Emotion Recognition Dataset Using Eye Tracking & Physiological Measures. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 178 (Dec. 2022), 20 pages. doi:10.1145/3495002

[60] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M. Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2015. "I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab. In *Proceedings of the Eleventh USENIX Conference on Usable Privacy and Security (SOUPS '15)*. USENIX Association, USA, 123–140.

[61] Chang Zhi Wei. 2013. Stress emotion recognition based on RSP and EMG signals. In *Advanced Materials Research*, Vol. 709. Trans Tech Publ, 827–831.

[62] Alma Whitten and J. Doug Tygar. 1999. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.. In *USENIX security symposium*, Vol. 348. 169–184. https://www.usenix.org/legacy/events/sec99/full_papers/whitten/whitten.ps

[63] Craig Williams, Helen M. Hodgetts, Candice Morey, Bill Macken, Dylan M. Jones, Qiyuan Zhang, and Phillip L. Morgan. 2020. Human Error in Information Security: Exploring the Role of Interruptions and Multitasking in Action Slips. In *HCI International 2020 - Posters*, Constantine Stephanidis and Margherita Antona (Eds.). Springer International Publishing, Cham, 622–629. doi:10.1007/978-3-030-50732-9_80

[64] Emma J. Williams, Joanne Hinds, and Adam N. Joinson. [n. d.]. Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies* 120 ([n. d.]), 1–13. doi:10.1016/j.ijhcs.2018.06.004

[65] Kun Yu, Ronnie Taib, Marcus A Butavicius, Kathryn Parsons, and Fang Chen. 2019. Mouse behavior as an index of phishing awareness. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part I 17*. Springer, 539–548.

[66] Jingwen Zhang, Dingwen Li, Ruixuan Dai, Heidy Cos, Gregory A. Williams, Lacey Raper, Chet W. Hammill, and Chenyang Lu. 2022. Predicting Post-Operative Complications with Wearables: A Case Study with Patients Undergoing Pancreatic Surgery. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 87 (July 2022), 27 pages. doi:10.1145/3534578