
There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

© 2020 Copyright is held by the author/owner(s). This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in the CHI ’20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ISBN 9781450367080
https://doi.org/10.1145/3313831.3376840

http://eprints.gla.ac.uk/205475/

Deposited on: 14 January 2020
The Role of Eye Gaze in Security and Privacy Applications: 
Survey and Future HCI Research Directions

Christina Katsini 
Human Opsis 
Patras, Greece 
ckatsini@humanopsis.com

Yasmeen Abdrabou 
Bundeswehr University 
Munich, Germany 
yasmeen.essam@unibw.de

George E. Raptis 
Human Opsis 
Patras, Greece 
graptis@humanopsis.com

Mohamed Khamis 
University of Glasgow 
Glasgow, United Kingdom 
mohamed.khamis@glasgow.ac.uk

Florian Alt 
Bundeswehr University 
Munich, Germany 
florian.alt@unibw.de

ABSTRACT
For the past 20 years, researchers have investigated the use of eye tracking in security applications. We present a holistic view on gaze-based security applications. In particular, we canvassed the literature and classify the utility of gaze in security applications into a) authentication, b) privacy protection, and c) gaze monitoring during security critical tasks. This allows us to chart several research directions, most importantly 1) conducting field studies of implicit and explicit gaze-based authentication due to recent advances in eye tracking, 2) research on gaze-based privacy protection and gaze monitoring in security critical tasks which are under-investigated yet very promising areas, and 3) understanding the privacy implications of pervasive eye tracking. We discuss the most promising opportunities and most pressing challenges of eye tracking for security that will shape research in gaze-based security applications for the next decade.

Author Keywords
Eye tracking; Gaze Interaction; Security; Privacy; Survey

CCS Concepts
• Security and privacy → Human and societal aspects of security and privacy; • Human-centered computing → Human computer interaction (HCI);

INTRODUCTION
The security community is an early adopter of eye tracking. Security researchers have explored the use of eye tracking for biometric authentication [14, 67, 107, 108] and password entry [53, 93] since the early 2000s. Twenty years later, eye tracking algorithms and technologies have matured significantly. Recent advances in visual computing, gaze estimation algorithms, cameras, and processing power of computing devices have led to eye tracking being no longer constrained to desktop computers but being also available on head-mounted displays [6, 45, 49, 52, 103], handheld mobile devices [73], and public displays [171]. Today, laptops such as Alienware 17 R4 and Acer Predator 21 X come with integrated eye trackers, and smartphones such as the iPhone X and Huawei Mate 30 Pro are equipped with front-facing depth cameras, capable of accurate gaze estimation. Eye trackers are also now used in driver monitoring systems in BMW [17] and Volvo cars [166].

These are important developments for security applications. The ubiquity of eye tracking means that researchers can finally take their gaze-based security applications to the real world. The benefits of large-scale eye tracking include: 1) higher adoption of gaze-based security applications (e.g., gaze-based privacy protection or gaze-based authentication), which in turn leads to better understanding their effectiveness and performance in daily scenarios, 2) allowing more gaze data to be collected that can be used to improve existing approaches (e.g., to improve the accuracy of biometric authentication), and 3) the ability to unobtrusively understand users through their gaze behaviour during security critical tasks (e.g., understanding gaze behavior when subject to phishing attacks).
Despite the significant potential this creates from both a security and a user experience perspective, a holistic view of how research on gaze-based security applications developed in the past decades is missing. This paper closes this gap by a) surveying and organizing knowledge in this field based on previous work in the area, b) clustering existing work into three main application areas: gaze-based authentication, gaze-based privacy protection, and gaze monitoring during security tasks, and c) highlighting the most promising opportunities and most pressing challenges that require further research.

We present 8 important and promising directions for future research in this area and identify 2 challenges that need to be addressed as they hinder the adoption of gaze-based security applications. For example, we highlight that the recent developments in the area make it possible to conduct field studies for gaze-based security applications. This was until recently infeasible due to limitations in mobile eye tracking hardware.

We discuss other research opportunities, such as blending implicit gaze-based authentication with everyday tasks and privacy protection by estimating users’ and bystanders’ gaze. We also underline challenges that are important to address in upcoming HCI research, such as understanding the privacy implications of pervasive eye tracking and the trade off between the accuracy and speed of gaze-based security applications.

RELATED WORK
There have been attempts to organize the existing research in eye tracking for security applications. Most of those focused on the use of physiological biometrics (e.g., iris [24, 111, 117], retina [24, 111, 117], periorcular [117]) for building and evaluating authentication schemes, without considering the use of gaze-based features. Many works apply gaze-based features to build an authentication scheme, aiming to overcome the limitations that are introduced by the use of physiological biometrics, such as inability to revoke passwords once compromised, or unintentional authentications [146]. Most surveys considering eye gaze do not address the HCI perspective but focus on technical [174] and methodological [47] aspects of such systems (e.g., the use of bio-signals for human identification [41]). None of the existing surveys consider multi-factor gaze-based authentication, gaze-based privacy protection, or the use of gaze to understand users during security tasks. To our knowledge, this is the first survey to holistically cover the three major utilities of gaze in security applications.

EXPLICIT GAZE-BASED AUTHENTICATION
Explicit gaze-based authentication refers to the use of eye movements to explicitly verify identity. In this type of authentication, the user has to first define a password that involves consciously performing certain eye movements (step 1: password creation). The user then authenticates by recalling these eye movements and providing them as input (step 2: password recall). The system detects the eye movements and compares them to the password defined in step 1 to verify users’ identity.

Researchers have explored a wide variety of eye movements that could be used for authentication. This includes fixations [93], gestures [33, 34], and smooth pursuit eye movements [8, 27, 35, 164]. There are two dimensions to consider in the use of gaze for explicit authentication: a) password type: legacy vs gaze-based password symbols, and b) used modalities: unimodal vs multimodal gaze-based authentication.
Legacy vs Gaze-based Password Symbols

The first dimension refers to the type of password. Passwords in gaze-based explicit authentication can have two forms: 1) gaze can be used as a modality for entering legacy passwords (e.g., PINs, text passwords or graphical password), or 2) gaze can be used to enter a gaze-based password where the password’s symbols are made of eye movements (e.g., a password that involves gazing to the left, then gazing to the right).

Legacy Passwords

Each password consists of a series of symbols. Traditional systems have used PINs and passwords (i.e., a series of digits and alphanumeric characters, respectively). Digits and alphanumeric characters are examples of legacy symbols that were argued to have been superseded, but it is difficult to replace them because of their wide use as they are easy to implement and easy to reset. Examples of systems that use legacy password symbols include banks and online websites. Gaze can support entering legacy symbols by providing a certain mapping between gaze behaviors and certain symbols. There are many examples of schemes employing gaze to enter legacy passwords. For example, Kumar et al. [93] proposed one of the first gaze-based authentication schemes where users fixated characters on an on-screen keyboard and then pressed the space button to select them. The same scheme was used on ATMs by Seetharama et al. [142]. Similar work was also done by Kasprowski et al. [63] who used gaze for pointing at PINs and confirmed selection by pressing a key. EyePassShapes uses eye movements to enter alphanumeric passwords [32].

Another body of work focused on using gaze to enter PINs. EyeDent [168] allows users to authenticate on desktops by entering 4-digit PINs using eye gaze. Users do not dwell or press triggers. Instead, the system automatically clusters the gaze points to estimate which targets the user intended to select. PathWord [8] is another system where users enter 4-digit PINs by performing smooth pursuit eye movements that follow the trajectory of the respective digits. GazeTouchPIN [79] allows users to enter 4-digit PINs on mobile devices using touch and gaze input. Liu et al. [102] explored using gaze gestures to enter 4-digit PINs on smartphones. Best and Duchowski [15] proposed using gaze to enter PINs on a rotary dial interface.

Several works used smooth pursuit eye movements to allow users to enter 4-digit PINs on a public display [27, 80, 83]. Other researchers explored the augmentation of graphical password schemes by using gaze input. For example, in the work of Forget et al. [43,44], Bulling et al. [20] and others [10,86,157], users fixated points on a shown image, using their gaze as an alternative to clicking with the mouse. Similarly, George et al. [49] used gaze to input 3D graphical passwords in VR. Another authentication scheme that has been extended using gaze input is PassFaces [129]; several works [37,53] extended PassFaces to allow gaze-based selection based on fixations. In EyeSec [98], the authors propose using gaze for input on multiple existing systems, including PIN pads and Patterns.

The advantage of using gaze to enter legacy password symbols is that they can easily integrate with existing backends. For example, to employ EyePIN [33] at an ATM that accepts 4-digit PINs, all that is needed is a camera to track the user’s eyes. However, the disadvantage is that the schemes might induce additional cognitive load on the user in order for them to understand the mapping between their gaze and the resulting symbol. Furthermore, most of these schemes are significantly slower to use compared to traditional, less secure alternatives. For example, GazeTouchPIN [79] and EyePassword [93] require 10.8 and 9.2 seconds respectively to authenticate, while classical PINs require 1-1.5 seconds only [167].

Gaze-based Passwords

Gaze-based passwords are based on gaze behavior. These schemes transform the password space and, hence, are likely to have a different impact on memorability. Examples include GazeTouchPass [74] and GTmoPass [78], where gaze gestures constitute part of the password. Similarly, in EyePass [34] and another work by De Luca et al. [33], the password consists of a series of gaze gestures. In DyGazePass [126, 127], the user’s input is a series of smooth pursuit movements that are supported by cues in the form of 2D geometric objects. In EGBP [141], users authenticate by gazing in one of four directions and then performing a blink to confirm input.

The advantage of gaze-based passwords is that they expand the password space by incorporating new sets of password symbols. Unlike legacy passwords where PINs, alphanumeric, and other widely used passwords are entered by gaze, gaze-based passwords might have a steeper learning curve as users are not accustomed to them, and require in-depth analysis of memorability as users would be required to learn and memorize unfamiliar password symbols which could be less intuitive.

Selecting and Recalling Gaze-based Passwords

Authentication schemes that leverage touch, mid-air gestures, or tactile input make use of the user’s muscle memory [144]. At the same time, the eyes are perceptual rather than control organs. There is no evidence so far that eye movements tap into muscle memory. This raises questions about the memorability of gaze-based authentication. For example, is it harder to recall legacy passwords when entered using gaze as opposed to other modalities? Does the process of recalling a password rely more on the input behaviour (e.g., finger movements vs eye movements during password input), or does it rely on the memorized password itself (e.g., 1234 vs gaze up, gaze down, gaze left). There has not yet been a direct comparison between memorability of legacy and gaze-based passwords. Also, it is unclear how users select gaze-based passwords and if their selections and perceptions of strong gaze-based passwords match reality. Prior work showed that users’ perceptions of strong text passwords do not always match reality [161], but this has not yet been investigated for gaze. These gaps have not been addressed and are open directions for future research. Understanding this will allow the community to identify whether gaze-based passwords offer additional benefits, or if it is rather more practical to use gaze to input legacy passwords.
Unimodal vs Multimodal Gaze-based Authentication
Gaze has been used as the sole input method when authenticating. We refer to that type as unimodal gaze-based authentication. On the other hand, multimodal gaze-based authentication is when gaze is combined with other modalities.

Unimodal Gaze-based Authentication
The advantage of unimodal gaze-based schemes is that they are handsfree. This makes them particularly useful for interfaces that are physically unreachable (e.g., displays behind glass windows [31]), for users with motor disabilities, and for touchless hygienic interactions. The disadvantage, however, is that unimodal gaze-based interfaces need a mechanism to distinguish input and perception. That is, the system needs to detect whether a user is gazing at a target to select it, or merely to perceive it. This has been traditionally addressed by introducing a dwell duration, i.e., the user has to gaze at the target continuously for a brief predefined period of time in order to select it [60]. This method has been adopted by several unimodal authentication schemes [20, 33, 44]. An alternative is to detect certain gaze behaviors that would indicate input, such as gaze gestures [36] as done in EyePassShapes [32], or smooth pursuit eye movements [27, 83, 126, 127, 164].

Multimodal Gaze-based Authentication
In multimodal schemes, gaze is used a) as a pointing mechanism while another modality is used for selection (we refer to this as gaze-supported multimodal authentication); or b) alongside a second modality to improve resistance to observation attacks by splitting the observer’s attention. The advantage of the former type is that, opposed to unimodal gaze input, the system can clearly distinguish input from perception as the user would use the second modality to confirm the intention to select the target being gazed at. Examples include EyePassword [93], GazeTouchPIN [79] and others [63], where users select each password symbol in two steps. Each step involves one of the modalities. In the latter type, gaze and other modalities are used together for improved protection against observation attacks. For example, users of GazeTouchPass [74], GazeGestureAuth [3] and GTmoPass [78] enter a series of gaze input alongside either touch input [74, 78] or mid-air gesture [3]. This requires shoulder surfers to observe two elements which in turn reduces attack success rates. The general disadvantage of multimodal authentication is that it usually complicates password entry. This could influence the memorability of the password symbols.

How to Evaluate Explicit Gaze-based Authentication
Users will, in general, not adopt a secure mechanism that is complicated to use and will find workarounds that reduce security (e.g., write down passwords). Therefore, it is important to make sure new schemes are both usable and secure.

Usability and Memorability Evaluation
Usability studies often include participants entering passwords using the new scheme, comparing it to a baseline – usually a state-of-the-art scheme. The user’s task is to enter a password provided by the experimenter. The password could be verbally communicated to the user [79] or read by a text-to-speech system [83]. Requiring the participant to read passwords from a piece of paper or a screen [82] is not recommended for gaze-based authentication schemes as it might impact the tracked gaze behavior. Error rates are often measured by detecting input failures. A less used approach, albeit equally important, is to present users with incorrect inputs and ask them to correct them [82]. This provides insights on how recognizable errors are, and how easy, fast, and accurately users can correct them.

Memorability is often evaluated by querying participants after a period of time to understand whether they remember a) how to use the scheme, and b) their secret. Note that memorability is important to consider regardless of whether legacy or gaze-based password symbols are used. Even if users are entering the commonly used 4-digit PINs, the input method impacts the user’s memory [32]. This underlines the importance of understanding memorability of proposed schemes.

Security Evaluation
Security studies of gaze-based input often focused on side channel attacks that can be performed by bystanders or located adversaries. Examples of studied attacks include observation attacks (e.g., shoulder surfing) [39] and video attacks [169]. In security studies that investigate shoulder surfing resistance, the user is often recorded while authenticating from the best observation angle possible. The recorded videos are then presented to a different set of participants who have been trained to attack passwords entered via the respective authentication scheme. For example, to evaluate GazeTouchPass [74], the experimenters recorded three videos: a) a video showing the user’s eyes to simulate an attacker observing the user’s eyes, b) a video showing the user’s screen to simulate shoulder surfing the touch input, and finally c) a video from an angle that allows observing both the gaze and touch input.

Guessing attacks, which help understanding how likely an attacker will find a password by random or smart guesses, are also important to investigate. However, the lack of knowledge on how users select their gaze-based passwords (Research Direction 1) means that there are no established strategies that attackers use to make informed guesses.

Evaluation Metrics
Regarding the evaluation metrics, usability often entails efficiency (i.e., entry time) and efficacy (i.e., error rate due to system malfunction or user errors). Other aspects that can be considered include error tolerance (i.e., how likely it is that users perform errors), learnability (i.e., how easy/fast users can learn how to use the scheme), and user preference. In memorability studies, the metrics are often the recall rate and the recall accuracy. The latter is a measure of similarity of how close the user’s recalled password is to the actual password.

Metrics that are commonly used in security studies are: a) attack success rate and b) attack accuracy. The former refers to whether or not attacks were successful while the latter measures how similar the attacker’s guess is to the actual password. These values are measured under a threat model that simulates an attack scenario. Suitable threat models should be employed when evaluating the security of authentication schemes. For example, previous work evaluated multimodal authentication schemes against two observers [76] and by using video record-
ings from two cameras [74]. Commonly studied threats include shoulder surfing attacks [39], video attacks [32], thermal attacks [1], and smudge attacks [11]. While gaze input is not vulnerable to thermal or smudge attacks, multimodal authentication schemes involving touch or tangible input might leak heat or smudge traces that make the scheme vulnerable to said attacks. Studying the aforementioned threat models is important because new input methods do not impact backend security, but rather impact the possible side channel attacks.

Further metrics that could be used to evaluate security include the number of guesses required until an attack is successful. Finally, the theoretical password space should also be computed for any new authentication scheme, while the practical password space can be computed through a longitudinal in the wild study to understand what kind of passwords users create.

Evaluating Explicit Authentication in the Wild

There are several evaluation paradigms in HCI research, each suitable for answering certain types of research questions. Lab studies are generally suited for quantitative measurements and collecting data in a controlled setting where external factors are reduced or ideally eliminated. However, lab studies suffer from low ecological validity as they do not reflect real life conditions. This is why the HCI community also embraces field studies where systems are evaluated in real scenarios such as in public.

Explicit gaze-based authentication was mainly investigated in the lab. As eye tracking technologies become cheaper [68], and soon to become ubiquitous on handheld mobile devices [73, 75], an emerging research opportunity is to evaluate authentication schemes in the wild. This would allow studying learnability to investigate if users’ performance improve after continued daily usage, as well as social implications when using the schemes in public, such as social embarrassment, or unintentionally looking at bystanders when performing gaze input.

There is a lack of – and hence an opportunity for – field studies of explicit gaze-based authentication. This will allow a) collecting ecologically valid findings, not impacted by eye fatigue caused by multiple entries during studies, b) studying the impact of learning effects, c) studying long term memorability, and d) studying use in different contexts.

Implicit Gaze-based Authentication

Implicit gaze-based authentication refers to the use of eye movements to implicitly verify identity; it does not require the user to remember a secret, but it is based on inherent unconscious gaze behavior and can occur actively throughout a session. It consists of two steps: the enrolment phase during which a digital representation of eye movements is acquired and stored as a template and the recognition phase during which the eye movement is tracked, processed and compared to the template to establish the identity of the individual.

Ideally, the tracked eye movements should possess the characteristics of an ideal biometric: universality, uniqueness, permanence and collectability. Research in the field mainly focused on assessing unique eye movements when performing activities with varying visual stimuli and type (e.g., time-dependent). Collectability mainly depends on the context of use (e.g., device capabilities, eye tracking metrics). The permanence of the eye movements has not been fully explored in this context.

Context of Use and Design Perspectives

Identification vs Verification

In implicit authentication it is important to distinguish verification (verifying user’s identity through a 1:1 comparison) from identification (i.e., discovering the user’s identity through a 1:N search) as it affects the authentication performance [149]. In identification scenarios, the more users the system has, the more realistic and more difficult the problem is, and thus, the performance of the system heavily depends on the sample size, the classification models, the device capabilities, and the required resources. Several surveyed works (52) focus on identification (e.g., [13, 21, 26, 30, 55, 65, 114, 115, 132, 139, 163]) while relatively less (21) focus on verification (e.g., [4, 22, 56, 61, 149, 173]), which requires significantly less processing effort.

Eye Tracking Metrics

In contrast to explicit gaze-based authentication, eye-tracking metrics for implicit authentication schemes are more diverse, such as gaze entropy [16], fixation density map [97], angular saccade velocity [104], and scan-paths [55]. Researchers, after acquiring gaze data, extend their data sets considering metrics that build upon the fundamental acquired data (e.g., fixation duration, angles velocity) and calculations on them (e.g., mean, maximum, minimum values). The more eye tracking metrics are available for building identification and verification models, the higher the likelihood for improved performance. Several toolkits (e.g., EMDAT, EALab) can be used to process eye-tracking data and generate larger, more inclusive data sets.

The eye-tracking metrics are often complemented with physiological eye metrics [4, 14, 30, 90, 150] or technology-based metrics, such as key-stroke sequences [147] and mouse dynamics [65, 136] aiming to improve the performance of implicit gaze-based authentication. However, the integration with metrics from multiple and diverse sources introduces a higher level of interdependence and complexity, which could be a barrier when attempting to adopt such implicit authentication schemes in real-life scenarios and everyday tasks.

Device Capabilities

The extracted eye-tracking metrics depend on the eye tracker’s capabilities, as they are associated with device-dependent specifications, such as operating distance, frequency, and operating window. Considering that eye trackers vary from sophisticated systems to simple embedded cameras, they have varying characteristics that influence the universality, the acceptance, and the performance of gaze-based implicit authentication. Considering that people use multiple devices with diverse characteristics, that issue becomes more intense. Very few research teams have considered the equipment when analysing and discussing the findings of their studies [38, 87, 165]. For example, Eberz
Two types of visual stimuli have been used in implicit gaze authentication: controlled and continuous. For controlled visual stimuli, people interrupt other tasks and focus their attention on this stimulus, thus being aware that they are going through an authentication task. Controlled visual stimuli can be either static or dynamic. Tasks that are based on static stimuli include text-based tasks [13, 14, 48, 56, 124, 133, 147], such as reading a passage excerpt, and static image-based tasks [14, 21, 25, 56, 104, 105, 115, 119, 132, 133, 150], such as the exploration of a photograph. The complexity of the images affects the accuracy of the scheme [150, 170]. Tasks that are based on dynamic stimuli elaborate the goal-oriented visual search approach of individuals as they typically are asked to track dynamic stimuli, such as moving target [6, 12, 14, 48, 56, 61, 64, 67, 88, 91, 106, 118, 136, 139, 148–150, 165, 172, 173, 175] or video recordings [23, 85, 133, 134, 143].

In contrast, when users authenticate through continuous stimuli, they may not be aware that they are being authenticated, as the stimuli are embedded to everyday tasks, such as reading emails and web browsing [38, 163, 172]. While continuous visual stimuli are of major importance for HCI as they enable unobtrusive authentication, they typically present lower accuracy and it is under-researched field, in comparison to controlled visual stimuli. Research with controlled visual stimuli has focused on refining authentication (e.g., improve accuracy, reduce time) to make implicit gaze-based authentication practical. Understanding the interplay between the diverse factors that influence the controlled-based authentication (e.g., task type, eye tracking metrics, minimum time) would help to move towards feasible and efficient continuous-based authentication.

et al. [38] used a downsampling approach to show that different sampling rates affected the quality of eye-tracking metrics. Similar work was reported by others [57, 62].

The trade-off between equipment features, the effort of developing sophisticated algorithms for implementing authentication mechanisms depending on sampling data, and processing requirements for using such a scheme in the wild remain unexplored.

### Context & Design of Implicit Authentication

**RESEARCH DIRECTION 3**

Different contexts of use of implicit gaze authentication have not been explored. Yet, context may pose different requirements to interface design and processing. Questions such as how authentication is triggered, if identification or verification is required, etc. could guide the design decisions for implicit gaze-based authentication. Given that identification requires more resources compared to verification, understanding which contexts require identification (e.g., claim an online profile), which require verification (e.g., unlock mobile phone) and which require both (e.g., access email from unknown device) would enable research to focus on realistic scenarios. Apart from the interface, the authentication factors need to be designed (e.g., data type, task type, time to authenticate) and scenarios where implicit authentication is a better fit need to be explored.

### Continuous vs Controlled Visual Stimuli

**RESEARCH DIRECTION 5**

Tasks that are based on dynamic stimuli elaborate the goal-oriented visual search approach of individuals as they typically are asked to track dynamic stimuli, such as moving target [6, 12, 14, 48, 56, 61, 64, 67, 88, 91, 106, 118, 136, 139, 148–150, 165, 172, 173, 175] or video recordings [23, 85, 133, 134, 143].

In contrast, when users authenticate through continuous stimuli, they may not be aware that they are being authenticated, as the stimuli are embedded to everyday tasks, such as reading emails and web browsing [38, 163, 172]. While continuous visual stimuli are of major importance for HCI as they enable unobtrusive authentication, they typically present lower accuracy and it is under-researched field, in comparison to controlled visual stimuli. Research with controlled visual stimuli has focused on refining authentication (e.g., improve accuracy, reduce time) to make implicit gaze-based authentication practical. Understanding the interplay between the diverse factors that influence the controlled-based authentication (e.g., task type, eye tracking metrics, minimum time) would help to move towards feasible and efficient continuous-based authentication.

### Blending Authentication with Everyday Tasks

**RESEARCH DIRECTION 4**

In implicit gaze authentication, the user authenticates based on unconscious eye movements when performing a task. So far, most research in the field is based on controlled stimuli. The users are aware that they are going through an authentication task and the same task is used for the enrolment and the recognition phase. Before attempting to move to continuous visual stimuli, it is essential to conduct more research where different stimulus is used for each phase. This is particularly desirable in an everyday setting. It is possible that the required accuracy is achieved by performing multiple tasks. Different tasks may provide better results for different stimuli. A combination of the above could provide the required accuracy for the authentication process. Another approach can be the use of gaze-based features that are task-independent. Another future research direction is to investigate which tasks result in the most useful data for implicit authentication.

### Task vs Time as Authentication Factor

**RESEARCH DIRECTION 4**

Time is important for the acceptance of an authentication mechanism [51]. Hence, implicit authentication should be performed fast. The majority of the surveyed papers is concerned with tasks as a whole (e.g., [14, 18, 22, 65, 119, 132]), meaning that the authentication process starts after the user has performed one or more tasks. Time varies from a few seconds (e.g., less than 10 seconds [26, 67]), 30 seconds [65], 40 seconds [132]) to a few minutes (e.g., 5 minutes [14], 25 minutes [85]), with tasks that are based on dynamic visual stimuli being faster. To improve performance, tasks can be repeated several times in the same session [46, 135, 152, 165, 175], resulting in longer duration. Very few works consider time as a factor (e.g., time-based analysis [38]). Five seconds seem to be a significant turning point for achieving good accuracy, with dynamic stimuli outperforming static ones [148, 150].

### Time as Factor in Implicit Authentication

**RESEARCH DIRECTION 5**

Authentication is a secondary task for users and needs to be fast and easy. For example, fingerprint authentication is gaining market share because users authenticate in a few seconds. In implicit gaze-based authentication, time is scarcely used as an analysis variable. More systematic research is needed for minimizing the authentication duration while maintaining a high performance, either by performing time series analysis of the authentication or by reducing the duration of the authentication tasks. To do that, there is need to study how performance is interrelated with different task types and visual stimuli, whether authentication can be achieved in less time for certain eye-tracking metrics, whether there is an interrelation between eye-tracking metrics and types of visual stimuli, and whether a combination of certain tasks would enable faster authentication.
How to Evaluate Implicit Gaze-based Authentication

Implicit gaze-based authentication has been evaluated towards performance, security, usability, and resources consumption. Long-term evaluation studies have also been conducted.

Performance Evaluation

The majority of the works in implicit gaze-based authentication focus on evaluating the proposed schemes towards performance (i.e., efficiency of the classification mechanisms) and explore the efficiency of different features against identification accuracy [26]. Several metrics have been used towards this direction, such as equal error rate – EER (e.g., in [55, 56, 85, 148, 150, 172]), receiver operating characteristic curve – ROC curve (e.g., in [6, 13, 25, 132, 149]), false acceptance rate – FAR (e.g., in [55, 88, 97, 132, 172]), and false rejection rate – FRR (e.g., in [55, 66, 88, 89, 132, 150, 172]).

In the authentication domain, reporting only the accuracy of identification or verification algorithms could conceal critical information about the efficiency of the mechanism and raise serious privacy issues. For example, reporting a 90% accuracy suggests 90% of the attempted users were correctly matched, but does not explain whether the remaining 10% were granted access to the system or not. The used evaluation metrics should assess both the probability of false acceptance and that of false rejection. We underline the importance of adopting the respective ISO/IEC standard for biometric evaluation [59].

When evaluating the performance of identification and verification mechanisms, the sample size is key to ensure the reliability of the obtained results. While in literature several suggestions regarding the sample size of eye tracking studies have been made [40, 116], there are no specific guidelines that have been proposed regarding the implicit gaze-based authentication. The number of participants may vary between less than fifty [14, 150], a few hundreds [21, 28, 131], or even thousands [12]. Moreover, several works are based on publicly available datasets [5, 28, 64, 114, 115, 132, 151, 152] to optimize the identification and verification algorithms. There is to date no gold standard for the sample size when evaluating implicit gaze-based authentication schemes.

Security Evaluation

Security evaluation in implicit gaze-based authentication is most often concerned with impersonation [50, 61, 118, 148, 149, 172] and replay attacks [148, 149]. In impersonation attacks, an adversary tries to fraudulently impersonate a legitimate user. Sluganovic et al. [148] simulated this type of attack and showed that increasing the threshold above which a sample is considered legitimate, decreases the likelihood of falsely accepting but at the same time increases the likelihood of falsely rejecting a legitimate user. Security of such attacks is also related to the stimuli complexity [172] and the type of the attackers: internal attackers (i.e., attackers known to the system) and external attackers (i.e., attackers unknown to the system) [149]. Access to the legitimate user’s calibration data also affects the success rate of impersonation attacks [118]. In replay attacks, a valid data transmission is maliciously or fraudulently repeated or delayed. To prevent this type of attacks the visual stimulus shown to the user should never be reused. For example, every time the user starts a new authentication session using a moving dot stimulus, the dot should move to different positions and in different order [148, 149].

Usability and Resources Consumption Evaluation

Very few works focus on evaluating the implicit gaze-based authentication schemes towards usability, such as time efficiency [150] and user experience [150]. Likewise, resources consumption, such as CPU and memory footprint [150] and energy consumption [172], has received little research interest.

Long Term Evaluation

While studying the long-term use is critical for authentication, very few works [6, 92, 118, 150, 173] report such studies. A degradation of accuracy is observed with time regardless of the type of visual stimuli used [92, 150]. To maintain high authentication accuracy, it is necessary to update regularly the owner template. Despite the fact that there is evidence that eye movements change with time [87], the aging factor in not well researched and understood.

Evaluating Gaze-based Implicit Authentication Schemes

Evaluation in implicit gaze-based authentication focuses on evaluating the performance of the proposed scheme. There is a research opportunity in expanding the work on security and usability evaluation. This will allow to a) better understand the possible threats of such schemes, b) understand which factors (e.g., stimulus) introduce security vulnerabilities, c) identify the factors that relate to usability and d) study the user acceptance dimension. Implicit gaze-based authentication schemes has only been studied in the lab. Most often the user is instructed to use a chin rest [12, 14, 89, 118, 135] aiming to achieve good calibration and ensure the performance is not a result of inaccurate gaze data. It is possible that a calibration-free gaze cannot provide the required data accuracy for this type of authentication. There is a research opportunity in evaluating such schemes in the wild. This will allow to a) understand how accuracy is influenced in real-life settings, b) study impact of aging effects, c) study impact of learning effects, and d) study different usage contexts.

GAZE-SUPPORTED MULTI-FACTOR AUTHENTICATION

Multi-factor authentication schemes are those that involve two or more authentication factors. The three most common authentication factors are knowledge, possession, and biometric [120]. Eye gaze can be used for supporting either the knowledge factor, by requiring the user to explicitly move their eyes to demonstrate “knowledge” of the authentication pattern, or it can be used for the biometric factor, by processing the user’s implicit eye movements to verify their identity. The possession factor refers to authenticating a user by showing they “have” a token, a key, a card, or similar.
While authentication protects privacy indirectly by limiting access to confidential content, some approaches aim at directly protecting private content from attackers. Here, gaze can be leveraged in two ways: a) Actively protecting the user’s privacy by, for example, hiding content the user is not looking at, or b) raising the user’s awareness of shoulder surfers by detecting the gaze direction of bystanders.

**GAZE-BASED PRIVACY PROTECTION**

While authentication protects privacy indirectly by limiting access to confidential content, some approaches aim at directly protecting private content from attackers. Here, gaze can be leveraged in two ways: a) Actively protecting the user’s privacy by, for example, hiding content the user is not looking at, or b) raising the user’s awareness of shoulder surfers by detecting the gaze direction of bystanders.

**Active Visual Privacy Protection**

Eiband et al. [39] showed that while most shoulder surfing resilience research focused on authentication, the vast majority of observed content is text, photos and videos. This means that we need methods to protect the visual privacy of users. Brudy et al. [19] proposed several methods to protect users of public displays from shoulder surfing. The gaze direction of the user and the bystanders were detected using a Kinect device. Privacy protection was done either by moving or hiding content, or by blacking out sensitive content such as personal emails. Ali et al. [7] proposed a slightly similar privacy protection application for detecting bystanders, but they only detected the presence of faces.

Similar systems, like EyeSpot [77] and Private Reader [125], were proposed for privacy protection on mobile devices. In EyeSpot [77], the content that the user is gazing at is visible to them, while the rest is masked either by a black filter overlay, a crystallized mask, or fake content. In the usability analysis of the different filter types, the authors found that the size of the visible spot impacts the reading speed significantly, and that the crystallized filter is more usable compared to the blackout one and fake text in terms of reading speed. The authors found no significant impact of the filters on neither the perceived mental workload nor text comprehension. However, participants favored the crystallized mask as it allowed them to see contextual information such as chat bubbles in chatting apps. Private Reader [125] similarly enhances privacy by rendering the portion of the text that is gazed at. The authors studied the impact on text comprehension and workload, and found that their method reduces comprehension and induces higher workload on attackers compared to the users.

While the aforementioned systems relied on eye gaze to selectively hide or render certain content, other works leveraged the inconspicuous nature of eye gaze to allow privacy-aware interactions. For example, iType [99] allows users to type text on mobile devices using their gaze. Another example is EyeVote [84], which allows users to anonymously vote on public displays without revealing their choices to bystanders. Several systems were proposed for transferring content from public devices to personal ones using eye gaze because it makes it more difficult for bystanders to know which content the user is interested in [110, 158–160].

While these systems were not built with the aim of privacy protection, privacy-aware interactions were a byproduct of using gaze input.

**Raising Awareness of Shoulder Surfers in Real Time**

In addition to active privacy protection, Brudy et al. [19] also proposed mechanisms for raising the awareness of public display users about bystanders who might be shoulder surfing them. They experimented with flashing the borders of the display when a bystander gazes at the display while it is in use by someone else, and visualizing the passerby’s gaze direction and/or body orientation when it is in use.

Zhou et al. [176, 177] proposed multiple interfaces that raise the user’s awareness of shoulder surfers through visual and auditory notifications. Similarly, Saad et al. [137] proposed different methods to communicate the presence of shoulder surfers to users by using face recognition. Despite not being based on gaze estimation, these works discuss this as a future step for improving the accuracy of detection and increasing the applicability of their concepts.
Privacy Protection by Assessing Bystander Gaze

Gaze-based privacy protection is a promising application area. However, research done in this area so far is disconnected. Some works investigated how gaze can be used to inform which on-screen content to hide from shoulder surfers [19, 77, 125]. In these works, the presence of shoulder surfers was assumed. Other works studied best practices to inform the user of the presence of shoulder surfers [137, 176, 177]. None of those works studied how to detect the presence of shoulder surfers.

A straightforward idea could be to detect the gaze direction of bystanders to estimate if they are shoulder surfing the user. However, this solution comes with several implications. First, from a technical perspective, wider angle lenses (e.g., fisheye lenses) need to be used, which, however, distort the edges of photos. This raises the challenge of detecting the eye contact despite the distortions to the shoulder surfer’s face. Second, being able to detect the gaze direction of surrounding bystanders brings its own privacy concerns: Is a user’s device allowed to detect the gaze direction of bystanders? How can bystanders consent to that? If consent can be retrieved, would shoulder surfers consent to processing their gaze behavior knowing that this might deter their attacks?

Previous work already discussed how the pervasiveness of eye tracking raises privacy and ethical concerns [73]. Here, we raise an additional concern of how protecting the privacy of the user might result in compromising the privacy of bystanders if we assume that every bystander is an attacker. Addressing this would not only improve privacy protection mechanisms, but can also improve authentication mechanisms. In particular, some observation resilient explicit authentication mechanisms are very effective against shoulder surfing but cannot be used on a daily basis due to, for example, requiring long entry times [79]. If the system is aware of the presence of bystanders, it could then require the user to authenticate using a more secure mechanism. This would improve the overall user experience as the user will be using a more usable mechanism.

IMPROVING SECURITY BASED ON GAZE BEHAVIOR

We discussed the use of gaze to support authentication and privacy protection. In addition, tracking the user’s gaze behavior could help understand their attitude towards security and detect insecure behavior. A number of attempts to build mechanisms for supporting or improving security have been proposed based on the understanding gained from observing and analyzing user gaze behavior when performing security tasks.

Prior work used eye tracking to study the effectiveness of security indicators on web browsers and the ability of users to detect phishing websites. Here, gaze has not only been used to understand user behavior [9, 29] but also as a mean to improve behavior to prevent such attacks [113]. For example, Arianezhad et al. [9] reported correlations between security expertise and gaze durations at security indicators. While their results highlight the correlation, Miyamoto et al. [113] developed mechanisms to build on this knowledge by proposing a web browser extension preventing users from providing input in web forms until they gazed at the browser’s address bar.

Steinfeld et al. used eye tracking to explore users’ attitudes towards privacy policies [155]. They revealed users’ tendency to read the policy when presented by default, while when given the option to sign their agreement without reading the policy, they tend to skip it. Pfeffel et al. [122] used eye tracking to explore how users decide if an email is phishing email or real.

Rather than for input, some researchers analyzed gaze during authentication with the aim to nudge users towards adopting more secure behavior. Mihajlov et al. [112] explored how much time is spent by users in different registration fields. Similarly, Katsini et al. [69, 71] and Fidas et al. [42] explored where users’ attention is drawn and how it is associated with graphical password choices. They used this knowledge to design mechanisms that nudge users towards better password decisions [71, 72]. In graphical authentication, eye tracking data has been used for building dictionaries of hot-spots [96], i.e., frequently selected – and thus insecure – positions and for creating cognition-based user models to provide personalized adaptations of authentication schemes [70, 128].

Understanding Gaze Behavior in Security Tasks

There is relatively little research in this area. Gaze behavior reflects cognitive processes, visual attention, and other user attributes [109] which can be used to identify vulnerabilities of security systems and design improved solutions.

Understanding gaze behavior can help improve security. For example, similar to previous work [9, 29, 113], eye tracking can be used to detect whether or not users examined an email’s sender to deter users from accessing links in phishing emails. Another approach is to use gaze to detect fear or sense of urgency [2], which are among the emotions social engineers try to instill in phishing and vishing attacks [130]. Analyzing gaze behavior can also help improve usability and memorability. For example, the user’s pupillary response can reflect if the cognitive load induced by recalling passwords is too high, indicating that the scheme’s memorability can be improved. Similarly, frequent scanpaths might indicate confusion, which can in turn indicate that the usability of a system or a task (e.g., installing security updates) should be improved.

A drawback of continuous gaze monitoring, even if done with the intention of improving security, is that the tracked gaze data can have negative privacy implications. For example, the widespread use of security applications that leverage gaze data can be a gateway for adversaries to spread malware exploiting the user’s gaze data for profiling. We discuss this in detail in “Challenge 2: Privacy Implications of Eye Tracking”.

RESEARCH DIRECTION 7

Privacy Protection by Assessing Bystander Gaze

Gaze-based privacy protection is a promising application area. However, research done in this area so far is disconnected. Some works investigated how gaze can be used to inform which on-screen content to hide from shoulder surfers [19, 77, 125]. In these works, the presence of shoulder surfers was assumed. Other works studied best practices to inform the user of the presence of shoulder surfers [137, 176, 177]. None of those works studied how to detect the presence of shoulder surfers.

A straightforward idea could be to detect the gaze direction of bystanders to estimate if they are shoulder surfing the user. However, this solution comes with several implications. First, from a technical perspective, wider angle lenses (e.g., fisheye lenses) need to be used, which, however, distort the edges of photos. This raises the challenge of detecting the eye contact despite the distortions to the shoulder surfer’s face. Second, being able to detect the gaze direction of surrounding bystanders brings its own privacy concerns: Is a user’s device allowed to detect the gaze direction of bystanders? How can bystanders consent to that? If consent can be retrieved, would shoulder surfers consent to processing their gaze behavior knowing that this might deter their attacks?

Previous work already discussed how the pervasiveness of eye tracking raises privacy and ethical concerns [73]. Here, we raise an additional concern of how protecting the privacy of the user might result in compromising the privacy of bystanders if we assume that every bystander is an attacker. Addressing this would not only improve privacy protection mechanisms, but can also improve authentication mechanisms. In particular, some observation resilient explicit authentication mechanisms are very effective against shoulder surfing but cannot be used on a daily basis due to, for example, requiring long entry times [79]. If the system is aware of the presence of bystanders, it could then require the user to authenticate using a more secure mechanism. This would improve the overall user experience as the user will be using a more usable mechanism.

IMPROVING SECURITY BASED ON GAZE BEHAVIOR

We discussed the use of gaze to support authentication and privacy protection. In addition, tracking the user’s gaze behavior could help understand their attitude towards security and detect insecure behavior. A number of attempts to build mechanisms for supporting or improving security have been proposed based on the understanding gained from observing and analyzing user gaze behavior when performing security tasks.

Prior work used eye tracking to study the effectiveness of security indicators on web browsers and the ability of users to detect phishing websites. Here, gaze has not only been used to understand user behavior [9, 29] but also as a mean to improve behavior to prevent such attacks [113]. For example, Arianezhad et al. [9] reported correlations between security expertise and gaze durations at security indicators. While their results highlight the correlation, Miyamoto et al. [113] developed mechanisms to build on this knowledge by proposing a web browser extension preventing users from providing input in web forms until they gazed at the browser’s address bar.

Steinfeld et al. used eye tracking to explore users’ attitudes towards privacy policies [155]. They revealed users’ tendency to read the policy when presented by default, while when given the option to sign their agreement without reading the policy, they tend to skip it. Pfeffel et al. [122] used eye tracking to explore how users decide if an email is phishing email or real.

Rather than for input, some researchers analyzed gaze during authentication with the aim to nudge users towards adopting more secure behavior. Mihajlov et al. [112] explored how much time is spent by users in different registration fields. Similarly, Katsini et al. [69, 71] and Fidas et al. [42] explored where users’ attention is drawn and how it is associated with graphical password choices. They used this knowledge to design mechanisms that nudge users towards better password decisions [71, 72]. In graphical authentication, eye tracking data has been used for building dictionaries of hot-spots [96], i.e., frequently selected – and thus insecure – positions and for creating cognition-based user models to provide personalized adaptations of authentication schemes [70, 128].

Understanding Gaze Behavior in Security Tasks

There is relatively little research in this area. Gaze behavior reflects cognitive processes, visual attention, and other user attributes [109] which can be used to identify vulnerabilities of security systems and design improved solutions.

Understanding gaze behavior can help improve security. For example, similar to previous work [9, 29, 113], eye tracking can be used to detect whether or not users examined an email’s sender to deter users from accessing links in phishing emails. Another approach is to use gaze to detect fear or sense of urgency [2], which are among the emotions social engineers try to instill in phishing and vishing attacks [130]. Analyzing gaze behavior can also help improve usability and memorability. For example, the user’s pupillary response can reflect if the cognitive load induced by recalling passwords is too high, indicating that the scheme’s memorability can be improved. Similarly, frequent scanpaths might indicate confusion, which can in turn indicate that the usability of a system or a task (e.g., installing security updates) should be improved.

A drawback of continuous gaze monitoring, even if done with the intention of improving security, is that the tracked gaze data can have negative privacy implications. For example, the widespread use of security applications that leverage gaze data can be a gateway for adversaries to spread malware exploiting the user’s gaze data for profiling. We discuss this in detail in “Challenge 2: Privacy Implications of Eye Tracking”.

RESEARCH DIRECTION 8

Understanding Gaze Behavior in Security Tasks

There is relatively little research in this area. Gaze behavior reflects cognitive processes, visual attention, and other user attributes [109] which can be used to identify vulnerabilities of security systems and design improved solutions.

Understanding gaze behavior can help improve security. For example, similar to previous work [9, 29, 113], eye tracking can be used to detect whether or not users examined an email’s sender to deter users from accessing links in phishing emails. Another approach is to use gaze to detect fear or sense of urgency [2], which are among the emotions social engineers try to instill in phishing and vishing attacks [130]. Analyzing gaze behavior can also help improve usability and memorability. For example, the user’s pupillary response can reflect if the cognitive load induced by recalling passwords is too high, indicating that the scheme’s memorability can be improved. Similarly, frequent scanpaths might indicate confusion, which can in turn indicate that the usability of a system or a task (e.g., installing security updates) should be improved.

A drawback of continuous gaze monitoring, even if done with the intention of improving security, is that the tracked gaze data can have negative privacy implications. For example, the widespread use of security applications that leverage gaze data can be a gateway for adversaries to spread malware exploiting the user’s gaze data for profiling. We discuss this in detail in “Challenge 2: Privacy Implications of Eye Tracking”.
FUTURE CHALLENGES
In the following, we highlight some of the most pressing issues that, in our view, should be addressed in the near future of gaze-based security applications.

Challenge 1: Accuracy and Speed Trade off
Implicit gaze-based security applications require highly accurate gaze estimates to be truly implicit and work without the user’s intervention. To collect highly accurate gaze data, calibration is necessary [109]. For a long time, eye trackers required users to be very still and even required them to use chin rests [33]. While modern eye trackers afford to allow users to move around to an extent, they often need to be recalibrated every time the user’s or the setup’s state change significantly. But calibration introduces an overhead to the interaction process, and it is perceived to be tedious, unnatural and time consuming [164]. There is a lot of research directed at making calibration more of an implicit rather than an explicit procedure by, for example, making it part of the interaction process while reading text or watching videos [81, 123, 156]. Previous work that studied implicit calibration addressed general use cases but not implicit authentication. This leaves room for future work on how to calibrate in a way to optimize the performance of implicit authentication. This requires first understanding the trade-off between calibration time and accuracy in implicit gaze-based authentication.

In contrast, some explicit gaze-based security applications do not require accurate gaze data. For example, many explicit schemes employ calibration-free gaze input methods like gestures [32, 79] and Pursuits [27, 83] which can perform accurately even when using inaccurate gaze data. These techniques require no calibration, as a result of which users can start the authentication process faster. However, calibration-free gaze input techniques often require longer entry times compared to other modalities. For example, in CueAuth [83], users spent 26.35 seconds to authenticate using Pursuits, while touch input required only 3.73 seconds. Achieving a balance between authentication time and calibration time – in particular, while considering the authentication context – is an important direction for future work to maintain fast authentication.

Challenge 2: Privacy Implications of Eye Tracking
The eye tracking technology itself can be a threat to privacy. For example, a user’s mobile device with eye tracking enabled could track the eyes of bystanders without their consent. This raises multiple questions. How can bystanders be made aware that a particular user’s device can track their eyes? How can their consent be retrieved? And how can their privacy be protected if they do not wish their eyes to be tracked? Like many ubiquitous technologies [95], eye tracking can reveal many private user attributes [100] such as emotional valence [121], mind wandering [162], personality traits [58], and women’s hormonal cycles [94]. Another important challenge is to securely store and process the gaze data without leaking it to third parties. This becomes more problematic if the tracking device uploads eye images to the cloud for processing rather than estimating gaze on the fly.

The privacy implications of pervasive eye tracking were discussed in recent work [73], and a few solutions to address this were proposed. For example, PrivacEYE [154] is a system which integrates a mechanical shutter into a wearable eye tracker. The shutter is activated when a bystander’s face is in the camera’s view. This protects the privacy of bystanders, and assures them that they are not being tracked. Another line of work applied differential privacy to gaze data by introducing noise to gaze data to prevent user identification without compromising the data’s utility. Steil et al. [153] applied their differential privacy approach on gaze interfaces in virtual reality, while Liu et al. [101] applied theirs on heat maps. Future HCI research should 1) investigate the privacy implications of pervasive eye tracking, and 2) develop mechanisms for protecting the privacy of not only the users but also everyone in the tracking range such as bystanders.

CONCLUSION
In this paper, we summarize previous work on gaze-based security applications and classify the utility of gaze into: 1) authentication, 2) privacy protection, and 3) understanding gaze behavior in security tasks. We identified eight promising research directions based on gaps that we found in the literature. For example, we see great promise in taking evaluations of explicit and implicit authentication mechanisms to real world settings, there are usability and security benefits of blending authentication with every day tasks, and there is a lack of work on gaze-based privacy protection and gaze behaviour analysis during security tasks. Furthermore, we identified two challenges that are important to address in order to make full use of gaze in security applications and require further research. Namely, there seems to be a trade off between the accuracy and speed of gaze-based security solutions. The trade off is particularly impacted by the need for calibration. A second challenge is that pervasive eye tracking itself can be a threat to privacy of the users of the technology and those around them.

ACKNOWLEDGEMENTS
This work was supported, in part, by the Royal Society of Edinburgh (Award number 65040), and the Deutsche Forschungsgemeinschaft (DFG), Grants AL 1899/2-1 and 1899/4-1.

REFERENCES


[99] Zhenjiang Li, Mo Li, Prasant Mohapatra, Jinsong Han, and Shuaiyu Chen. 2017a. iType: Using Eye Gaze to Enhance Typing Privacy. In IEEE INFOCOM 2017 - IEEE Conference on Computer Communications. IEEE, USA, 1–9. DOI: http://dx.doi.org/10.1109/INFOCOM.2017.8657233


