A Decade of Development of Mental Models in Cybersecurity and Lessons for the Future

Robert Murimi[1], Sandra Blanke[2], and Renita Murimi[3]

[1,2,3] University of Dallas, Irving TX 75062, USA
{[1]rkmurimi,[2] sblanke,[3] rmurimi}@udallas.edu

Abstract. Mental models are essential in learning how to adapt to new and evolving circumstances. The landscape of best practices in cybersecurity is a constantly changing area, as the list of best practices evolves in response to the increasing complexity and scope of threats. In response, users have adapted to the threats and corresponding countermeasures with mental models that simplify the complex networked environments that they inhabit. This paper presents an overview that spans over a decade of research in mental models of users when dealing with cybersecurity threats and corresponding security measures in different kinds of environments. The lessons from over a decade of research in mental models for cybersecurity offer valuable insights about how users learn and adapt, and how their backgrounds and situational awareness play a critical role in shaping their mental models about cybersecurity.

Keywords: Mental Model, Cybersecurity, User Engagement.

1 Introduction

A recent report from ThreatLocker found that 85% of cybersecurity breaches in 2021 were due to human errors [1]. This finding comes on the heels of similar reports that document that if the human were eliminated, 19 out of 20 breaches could have been eliminated [2]. One significant factor that contributes to the human error in cybersecurity breaches is complexity. Users are faced with increasingly complex decisions, such as the need to understand evolving advice on password hygiene, scanning emails for suspicious headers and content, and refraining from clicking on benign-looking download buttons that carry malicious payloads. At the same time, the stakes keep getting riskier. Whereas once the clicking of links and buttons might have resulted in individual data loss at a lower level of risk, threat vectors carrying ransomware have the potential to result in massive data breaches that cause damage to the entire organizational infrastructure. Simultaneously, technology applications have diversified rapidly. For example, Web3, the third generation of the Web, which is built upon blockchain technologies has required users to learn about new concepts, terminologies, and applications.

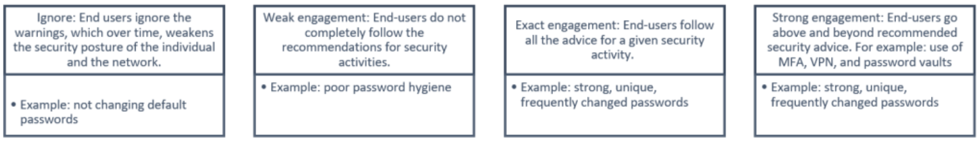
One way in which people adapt to the complex environments around them is through the development of mental models. In [3], the authors define a mental model as a "dynamic, planned action setting must be composed of (at least) these four elements:

intentions, perceptions, system structures, and plans." Mental models seek to simplify the complexity and extract elements that are useful to the users in order to understand, adapt, and engage with the tools and platforms of these environments, and develop with experience [4]. Mental models are dynamic and are based on individual experiences [5], as they influence people's decisions and actions [6]. However, certain challenges to mental models exist. Mental models have the potential to alter behavior, but not necessarily for the better since they can be incomplete or inaccurate [7]. An incorrect mental model can make users mistrust insecure technologies [8]. Further, since mental models exist in the mind, one cannot directly measure which mental model is active in a person's mind and extent of the model's performance.

An effective mental model that keeps up with the advances in cybersecurity requires a human-centered foundation, with consideration of the technology, situational awareness and human behavior. In cybersecurity, secure decisions rely on users having correct mental models of security issues. In general, computer users have difficulty anticipating cybersecurity threats, or perhaps may take incorrect actions and end up making themselves less secure. While cybercrime has a diverse array of outcomes, such as data theft, fraudulent financial transactions, stolen intellectual property, or software or hardware destruction, hacking often plays a significant role in the execution of cybercrime [9]. Organizations routinely roll out improved threat detection and intrusion prevention systems to bolster cybersecurity, but hackers have begun targeting humans in addition to hardware or software.

The issue of how to increase understanding of, and compliance with, security measures has been one of the hallmarks of cybersecurity research and has necessitated some attention to the psychology of users. In response to a warning or suggested activity, a typical user has four ways to engage with the warning or activity. These ways lie on a spectrum of engagement, as shown in Fig. 1, which illustrates how users interact with password hygiene advice. At one end of the spectrum, a user might *ignore the warning* or security activity and ignore the warnings about changing default passwords. For example, generic SSL/TLS (Secure Socket Layer/Transport Layer Security) warnings does little to encourage or discourage these mental models, so users are relatively free to adopt whatever mental model they like. As found in [10, 11], users' responses to TLS warnings are relatively consistent: they usually ignore them. A slightly more involved form of engagement is *weak engagement*, where the user might change the default password but only substitutes it for a weak password. Further along, a more cautious user would engage *exactly* as recommended by the warnings and security recommendations and use passwords that are strong, unique, and frequently changed. At the extreme end of the spectrum, motivated users with *strong engagement* would go above and beyond the recommendations for password hygiene and leverage additional tools for password security such as vaults and VPNs. For each of these levels of engagement, the driving factor is the mental model of the user. It should be noted that the modes of engagement vary determined by multiple factors such as expert guidance on best practices, cultural norms, risk-averse behavior, and familiarity with technology. For example, the National Cyber Security Center (NCSC) issued password guidance in 2015 that no longer endorsed frequent change of passwords, citing the burden imposed on users in changing and remembering the new passwords [12].

This paper offers a review of mental model research in cybersecurity over the past decade, where users have developed ways (ignore, weak, exact, and strong) of interacting with security recommendations and security activities in their networks. Although we refer to a "decade" of research, we have included research extending as far back as the early 2000s in order to include some seminal work referring to the role of mental models in how we think about secure online practices. The research that we profiled for this paper was derived from a Google Scholar search of the terms, "mental models" and "cybersecurity". The rest of this paper is organized as follows: Section 2 presents two categories of mental models, folk metaphors and formal methods, to classify the ways in which users engage with the technologies. Section 3 presents findings in mental model research of interfaces that people interact with and how that influence cybersecurity. Section 4 provides cybersecurity-related mental model research in specific platforms, technologies, user types, social factors, tools, or applications. The implications of this research for emerging technologies are discussed in Section 5, and Section 6 concludes the paper.



Ignore: End users ignore the warnings, which over time, weakens the security posture of the individual and the network.

- Example: not changing default passwords

Weak engagement: End-users do not completely follow the recommendations for security activities.

- Example: poor password hygiene

Exact engagement: End-users follow all the advice for a given security activity.

- Example: strong, unique, frequently changed passwords

Strong engagement: End-users go above and beyond recommended security advice. For example: use of MFA, VPN, and password vaults

- Example: strong, unique, frequently changed passwords

Fig. 1. User engagement with cybersecurity activities with examples of password hygiene.

2 Folk Metaphors and Formal Models

Since a mental model is a representation in working memory, the mental models of experts and non-experts vary. The mental models of two types of users – those who had informal exposure to cybersecurity topics and those who had formal exposure were studied in [13]. Starting with a hypothesis of inverse relationship between cybersecurity knowledge and perceived security, the authors constructed mental models of end users' cybersecurity knowledge. They found that Users with formal cybersecurity exposure offered longer responses and more domain-specific words than users with informal exposure. Thus, prior formal exposure was shown to produce different mental models between the two groups of users. This has important implications for cybersecurity training and awareness campaigns, where end users may possess varying levels of domain-specific expertise in cybersecurity topics. While the role of mental models in cybersecurity is undisputable, there is a wide range of models in existing literature. Categorizing them and systematically analyzing them will be the theme of the rest of this paper.

To start with, this section approaches mental models in terms of their underlying philosophy. The ones derived from analogies to common social phenomena such as medical infections or warfare are referred to as folk models, while those derived from engineering disciplines are referred to as formal models. This categorization is not absolute, since there are some studies whose findings about mental models would not neatly fit in either or have significant overlap between the two.

Table 1. Summary of mental models in cybersecurity

Article	Mental model developed/tested	Type	Focus of mental model(s)
[14]	Physical, criminal, medical, warfare, market	Folk	Security risks
[16, 17]	Crime, burglar, mischief, vandal, buggy, medical	Folk	Common security activities
[16]	Stupid, education, understand-how-users-think	Folk	Home users
[19]	Brave new world model, fatalistic, little value, maintenance, not-my-job, optimistic model, reputation, verification model	Folk	Privacy and security
[20]	Error and blocking state model	Formal	Formal methods
[37]	Control methods, OODA (Observation, Orientation, Decision, Analysis), and HITL	Formal	Cognition
[65]	Incorrect, incomplete, partially correct, complete	Formal	firewalls
[81]	Access control, black box, cipher, iterative encryption	Formal	Encryption technologies
[48, 49]	Usable security models	Formal	Usability
[78]	Human factors ontology (HUFO)	Formal	Trust
[81]	Access control, black box, cipher, iterative encryption	Formal	Encryption Technologies
[87]	Incomplete, inaccurate	Formal	Cryptocurrencies

2.1 Folk Models

Among the earliest research in folk mental models for cybersecurity is the work in [14]. Here, the authors proposed a framework where there are predominantly five kinds of mental models for communicating complex security risks. The models take the form of analogies or metaphors to other similar situations: physical security, criminal behavior, medical infections, warfare and market failure. The use of metaphors for mental model nomenclature was used to make security less virtual and more tactile in order to increase risk awareness. Associating virtual risks with more tactile risks—wild animals, disease, crime and war has been shown to increase sensitivity to and awareness of risks [15]. Comparison of responses by experts and non-experts to these metaphors showed that non-experts found physical and criminal model to be the most accessible mental model for cybersecurity risk communication. Additional work in [16] implemented mental models from [14] by using agents that simulate human behavior within a network security testbed. The simulation exercise specifically analyzed four security activities: using antivirus software, caution in visiting websites, making regular backup, and updating patches regularly. The findings suggested that mental models are not necessarily self-consistent.

Other early work in [17] investigated the mental models of home computer users in the context of computer security. Their findings about users' mental models were categorized as follows: buggy (due to software flaws), mischief (due to mischief-mongers), crime (intended to obtain sensitive information), burglar (stealing financial data), vandal (causing damage for showing off), and big fish (targeting rich or important individuals for attacks). The authors noted that majority of home computer users have little computer security knowledge and most of the decisions they make about computer security are guided by how they think about computer security, which may not be technically correct to lead to a desirable security behavior. In other words, sometimes even "wrong" mental models produce good security decisions. Thus, this work called for additional research to investigate the connection between mental models and actual security behaviors, since not all mental models lead to positive security behaviors. A key finding of this research was that non-expert users such as home computer users could still navigate computer security-related issues efficiently, thus eliminating the constraint that nontechnical users must become more like computer security experts to properly protect themselves. The authors argued for two action items to successfully change people's mental models. First, research needs to identify how people form these mental models, and how these mental models can be influenced. Second, the relationship between these mental models and associated security behaviors needs to be analyzed to identify which mental models are good for home computer users. The authors recommended three approaches to assist users in computer security. First, the *stupid* approach attempted to create security solutions by removing the user from decision making process since they are seen as the weakest link [18]. This approach was not recommended due to its one-size-fits-all stance to security problems in spite of the fact that people use computers for such a variety of different purposes. Second, the *education* approach was suggested that allows users the freedom to choose and provide them with appropriate training to enable them to make good security choices. However, this approach was constrained by the fact that home computer users are rarely interested in learning the details of how security software works. Finally, the *understand-how-users-think* approach involves working to understand how computer users think about security, and how they make security decisions. This approach leads to an understanding of user thinking (mental models) in order to understand user behavior. A key finding of their paper was that mental models are neither correct nor incorrect, rather they result in different potential benefits from associated behaviors. Thus, the emphasis should not be on teaching people "correct" mental models, but on finding ways to encourage models that lead to valuable security behaviors.

Work in [19] explores how people use mental models to understand and experience privacy and security as they engage in online and computer actions and activities in their daily lives. Here, the authors emphasize the need for effective cybersecurity training and education for end- to navigate the online environment safely. However, effective cybersecurity training requires that we first understand how people think about and experience online privacy and security in their daily lives. In their findings, the authors identified multiple mental models: the brave new world model, the fatalistic model, the little value model, the maintenance model, the not-my-job model, the optimistic model, the reputation model, and the verification model. These models were identified as coping mechanisms for users to adapt to their rapidly changing technological environment. Their findings also suggested that these mental models were sometimes partially formed, and that multiple mental models were used simultaneously indicating the lack of a single overarching model among users. Among the various mental models that were identified, the authors found that the brave new world model was the most common model among users.

2.2 Formal Models

Formal methods in mental models are derived from range of disciplines, including control systems, human factors, and state models. Mental models of users at the intersection of formal methods and human factors engineering [20] were analyzed in the form of a proposed framework that combines the error and blocking state architecture of [21] with human factors engineering. This framework builds upon folk models of computer security threats initially conducted by [22] and uses error and blocking states to reveal insights about user and system model mismatches. The use of formal methods coupled with human factors engineering offers powerful tools for discovering the hard-to-anticipate failure modes that threat actors leverage through social engineering and other attack strategies in which humans are the targets.

Building upon the mental model frameworks developed in [16, 23, 25] suggested that the use of mental models in cybersecurity serves two purposes. First, mental models could result in “strong intervention”, which states that mental models are necessary in order to understand the internet security situation [16]. Second, mental models could also result in “weak intervention”, which states that mental models supplement understanding of the Internet security situation [23]. The difference between the strong and weak claims is that the strong claim predicts that understanding and performance in cybersecurity situations is improved by the use of mental models, whereas the weak claim predicts only that understanding and performance in cybersecurity situations will be changed using mental models. The findings of [24] indicated that while there was little evidence for the strong intervention hypothesis, the use of any mental model or metaphorical framing of the context improved people’s understanding of internet security situations. They also found that while the weak intervention hypothesis predicted systematic changes in performance across different mental models, there was no specific change in performance. In general, there was a better overall performance when using cybersecurity context than when employing any mental models.

Two approaches to risk communication: the human information-processing approach (C-HIP) and the mental model approach in risk communication (MMARC) were analyzed from the perspective of approaches for risk communication in cybersecurity

[25]. The C-HIP approach has been taken by researchers in cybersecurity to study user perception of security risks and how risk perception influences risk-taking behavior [26]. This approach characterizes the human as a communication system, with risk-communication information from a source delivered to the receiver, who processes the information in various stages [27]. In the C-HIP approach, it is assumed that effective communication must trigger attention of the receiver, achieve comprehension, and influence decision making and behavior of the receiver [28]. This model not only considers the traditional information-processing stages of humans, but also accounts for social-cognitive and cognitive-affective components. Work in [29] used C-HIP as an investigation tool and designed a survey to understand users' attitudes towards software warnings and updates, to explain users' hesitation in applying software updates.

While memory and comprehension are key components in the C-HIP model [30-32], the MMARC approach emphasized the importance of understanding and comparing the mental models of experts and non-experts, and drafting and evaluating risk-communication messages. MMARC has been widely applied to studies concerning risk communication, from communications for tourists [33], flash flood risks [34] and medical risk [35]. The concept behind the MMARC approach is that security designs and educational efforts should align with users' mental models that direct their decisions and actions [5, 36]. This approach highlights a five-step process - developing the expert model, eliciting the public model, and comparing it with the expert model, conducting confirmatory surveys among the broader population to determine the prevalence of the public model, drafting risk-communication messages based on the knowledge gap between experts and non-experts, and evaluating effectiveness of the messages.

The role of cognition in cybersecurity activities was studied in [37], where the authors noted that although technology plays an important role in cybersecurity, it is humans who play key roles in cyberspace as attackers, defenders, and users. The authors proposed a cognitive security model with three layers: Knowledge, Information and Cognitive. Their model supports the process of modeling mental maps, the generation of knowledge, and the fusion and handling of large datasets. The proposed cognitive security model includes control techniques, OODA (Observation, Orientation, Decision, Analysis), and HITL (Human in the loop). OODA was employed to infer patterns from the analysis of the datasets by generating mental models based on profiles of attacks, threats, and user behaviors. The outcomes of OODA would be used to influence the situation awareness of the organization and associated activities for maintaining secure cybersecurity posture. HITL integrates human interaction with the technological solutions of the three layers, and informs the generation of false alerts and other indicators of performance. The use of this proposed model affords the security analyst the ability to integrate experience and knowledge for a dynamic analysis of the organization's cybersecurity posture and the support of decision-making processes.

As shown in Table 1, not all models fall neatly within folk or formal categories. For example, a basic model of cyberspace consisting of three layers: technical, socio-technical, governance to create a foundation with three corresponding mental models was proposed in [38]. An extension with eight additional mental models, crown jewels, kill chain, situational awareness, risk assessment model, risk response model, institutional model, direct and indirect social contract model, and triple bow tie model, was further suggested. (See [38] for an explanation of each of these models.) In the basic model with three layers, each layer has its own security requirements. The inner layer of the cyberspace model concerned all IT that enables the cyber activities. The middle layer of socio-technical cyber activities are activities executed by people and smart IT for accomplishing their personal, business, or societal goals. Finally, the outer layer concerns the governance layer of rules and regulations that should be put in place to properly organize the two underlying layers, including their security.

3 Interface Design and Perceptions of Security

People interact with technology of any kind using interfaces, and this holds true for activities related to cybersecurity as well. Thus, interface design is a crucial element in the design of secure systems and networks. This section explores existing work in how mental models of interfaces influence the activities of users, and the importance of interface design in altering the perception of security.

3.1 Interface Design

Human factors engineering was studied in [39] for exploring the human dimension of cybersecurity by employing a human factor integration (HFI) framework. Defining HFI as a “systematic process for identifying, tracking and resolving the wide range of human related issues in the development of capability”, the authors used HFI to consider the multiple ways in which the human can differentially affect the security of a system. The key mandate of the HFI process is to characterize and address the risks to a system generated by the humans. The HFI framework is divided into seven domains: social & organizational factors, manpower, personnel, human factors engineering, system safety, training, and health hazard assessment. Noting that the complexity of current technical systems is a major source of vulnerabilities, the authors suggest that introduction of technical security measures in such complex environments can lead to unforeseen human consequences including a reduction of effectiveness or efficiency. Similar to the work in [40, 41] that counters the notion of humans being the weakest link in the fight against cyberattacks, their findings indicate that HFI has the potential to prioritize and act upon the greatest risks originating from the human factor while also adhering to time and budget constraints.

The focus on the human as the introducer of risk in networks is countered in [41], where the authors addresses the misconstrued notion that the “user is the enemy”. Many factors such as lack of clarity, uncertain consequences, use history, system expertise, shared responsibility, and conventional reliance on technical experts for providing cybersecurity have been identified as contributing factors to this notion. However, labeling users as the problem is not a solution—this displacement of responsibility takes a costly toll on our economy and on our safety. The direct relationship between usability and security dictates that the users cannot be left out of this equation of solving cybersecurity issues. Changing user expectations mean that all users, and not just experts, now have an active role to play in digital security. By shifting focus and start looking at users as the greatest hope for system security, interface designers can explore how compromising behavior can be designed out of the system. By considering the user's profiles in terms of what they know, how they use the system, and what their needs are, designers will be better positioned to empower them in their digital security roles. The findings of [41] pointed to the need for appropriate mental models, increased transparency, and effective interface design for situational cybersecurity awareness. Additional work on the use of mental models for correct interpretation of displayed information is in [42], where the emphasis is on the design of cyber security dashboards that leverage data visualization and consequently guide policymaking. Existing literature on dashboard design suggests the construction of different types of dashboards for different people, which might be due to the dominance of different mental models in different user groups. Based on expert interviews, the authors demonstrated that there exists a difference in the perceptions of cyberattacks by different categories of users such as managers, operators, and analysts.

Aligning designing guidelines for usable cybersecurity systems to standardized security controls has been studied in [43]. Here, the authors propose the design of context-independent guidelines, that do not focus on specific areas within cybersecurity such as authentication, access control, encryption, firewalls, secure device pairing and secure interaction. The rationale here is that broad context-independent guidelines can be adapted readily to various domains, and offer scope for innovation and customization while also being responsive to core problems and evaluation methods.

Interface design in applications using anonymous credentials was studied in [44]. Here, the authors explored ways in which mental models of data minimization can be evoked on end users for online applications. Existing literature in mental models of anonymous credentials is sparse both due to their novelty and complexity, which has further complicated the design of easily understandable interfaces end users. The authors found that users have grown accustomed to believing that their identity cannot remain anonymous when acting online, and so they lack the right mental model to understand the working of anonymous credentials or how it can be used to protect their privacy. In their study on mental models of anonymous credentials, the authors explored different user interface approaches for anonymous credentials based on three different metaphors: card-based, attribute-based, and adapted card-base approaches. The authors found that successful adoption of novel technologies such as anonymous credentials requires a comprehension of their advantage and disadvantages, and that inducing adequate mental models is a key issue in successful deployment. Out of the three metaphors, the authors found that the adapted card-based approach was the closest comprehensive mental model for anonymous credential application by helping users understand that attributes can be used to satisfy conditions without revealing the value of the attributes.

The importance of effective risk communication using mental models that incorporated human-centered security was studied in [5]. The authors extend the narrative that end-users cannot be blamed for being the weakest link in cybersecurity, due to the number of warnings that the end-users receive on a daily basis. The high volume of warnings coupled with the inability to understand the nature of these warnings leads users to ignore these warnings, which eventually turns into a habit since users do not perceive the risk of ignoring these warnings. To rectify this malformed habit of ignoring warnings, the author proposes that the design of warnings should be aligned with the mental models of end users, and not just those of the developers and designers. The decision fatigue imposed by warnings was also studied in [45], where the authors found that users were overwhelmed by the constant need to be alert requiring them to make more decisions than they could process. The authors suggested the need for simpler user interfaces that could aid users in mitigating the decision fatigue caused by complex security advice.

While most research in mental models is focused on the development of mental models, work in [46] investigated the assessment of mental models. Here, the authors developed an interface called Sero using the concept of mapping. Concept mapping enables advanced assessment techniques for mental models, with their capacity to blend recall, recognition, and reasoning techniques in the context of a nonlinear assessment. In this manner, concept mapping allows for more efficient data collection than interviews do, and presents an advantage over writing-based assessments and facilitates self-monitoring. authors note that although other techniques such as think-aloud protocols, narrative text, causal diagrams, pretest–posttest comparison, and lunar phases concept inventory (LPCI) represent other modes of assessment, they often fall short of the necessary requirements for the assessment of mental models and are not feasible for practical implementation.

Software designers and security architects continue to face issues in developing a model that is both secure and usable. This is because security and usability seem to be conflicting in their goals. The main challenge in designing any usable security is finding a balance between protecting the system from unauthorized disclosure and cognitively designing the system to conform to the user's expectations and satisfaction. In [47], the authors developed a holistic meta-model that combines security, usability, and mental models. This meta-model applies knowledge nurturing, mobilization and sharing concepts in its development. The authors found that the degree of usable security depends on the ability of the designer to capture and implement the user's tacit knowledge. The authors advocate for user interface design that aligns with the user's mental and conceptual models and is consistent with the user's expectations for functionality of the system. Mental models help to bridge the incongruity gap between the security and usability expectations of users, and thus, are a crucial foundational element of cybersecurity.

Recognizing that not all security measures are friendly for users, [48] called for the need for usability tests that are different from the conventional software testing activities with a focus on rooting out any impediments that might affect user experience. Specifically, usability testing calls for a focus on user perceptions, characteristics, needs, and abilities as essential inputs for effective and robust system design. In their follow up to this article published almost a decade later [49], the authors noted that user resistance, ineffectiveness of password-based security measures, and high volume of breaches indicate a pressing need for revisiting usable security. Their recommendations for usable security include the need for improved user experience, codification of best practices, graceful recovery procedures, and thinking of users not as adversaries but as part of the solution in cybersecurity.

Interface design for risk communication was also studied in [50], where the authors analyzed the mental models of security experts and non-experts. Typical risk communication consists of a message that has been formulated by security experts to warn the non-experts of the looming threats. The gap between mental models of security experts who create the risk communication, and the mental models of non-experts who are expected to act upon the risk communication can decrease the efficacy of the risk communication. The authors note that the purpose of risk communication is not to convey a perfect truth to the users, but rather prompt them to take appropriate action in defending their systems against risk. Their findings showed that mental models based on physical security were appropriate for the non-experts but not the medical infection mental model, while the opposite was found to be true for experts leading the authors to suggest that risk communications should be driven by mental models of non-expert users.

3.2 Perceptions of Security

While user interface design has garnered a lot of interest in motivating users to perform desirable activities to improve their risk posture, interface elements also tend to be ignored by users leading to increased risk to the users' security and privacy. Work in [51] investigates the motivation behind the reason the users chose to follow or not follow common computer security advice. The impact of security advice on users in four well-known areas was studied: keeping software up to date, using password managers, using 2FA (two-factor authentication), and changing passwords frequently. The authors used a cost-benefit framework in their study, which was supplemented by risk perception and social motivation constraints. Their findings indicated that risk perception is central to security behavior while social motivation is much stronger and independent of instrumental motivation. Unlike other studies that separated experts from non-experts, their work simply compared those who followed the advice versus those who did not. The authors found that social considerations were largely trumped by individualized rationales, and each group viewed their decision as the rational one.

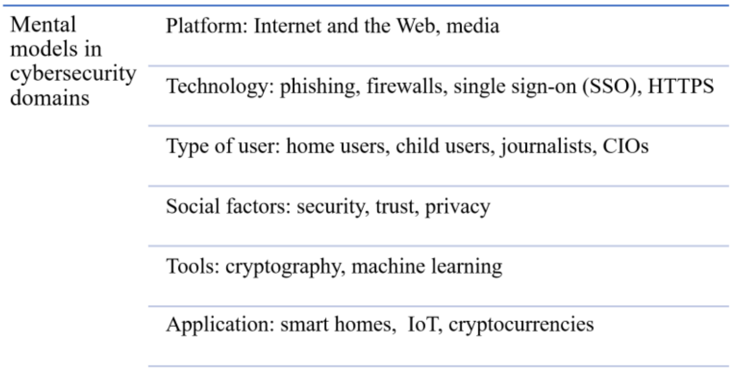
Negative perceptions of security were explored in [52]. The authors explored how security advocacy can attempt to overcome negative perceptions that security is scary, confusing, and dull. Although cyber threats are evolving, users are falling behind in defending their systems and networks. Users often fail to implement and effectively use basic cybersecurity practices and technologies, due in part to negative feelings about security. Their findings called for security advocates, who must first establish trust with their audience and address concerns by being honest about risks while striving to be empowering to overcome these negative perceptions.

Cybersecurity perception and behavior differed between experts and non-experts [53], where the authors found that part of the challenge of cybersecurity has been understanding the ways in which different groups of people think about and interact with cybersecurity. By examining the similarities and differences between experts and non-experts and identifying what characteristics influence their attitudes and behaviors, the authors provided insights into how to help non-experts understand and protect themselves online. In their study, the authors did not observe experts and non-expert as two separate groups, where participants could fall anywhere between non-experts and experts. Their findings showed that non-experts did not have a solid mental model related to cybersecurity, and they drew on multiple mental models that were ill-formed and which only partially helped them understand and navigate cybersecurity. On the other hand, experts used different mental models and tended to be proactive in their online security practices.

Additional work in [54] explores the reasons why non-experts choose not to protect themselves from cyber threats by investigating the role of catalogued warning messages. In their evaluation, the authors organized their study around five elementary components of Technology Threat Avoidance Theory (TTAT) - perceived susceptibility, perceived severity, perceived effectiveness, perceived costs, and self-efficacy. The authors observe that non-experts who choose to not protect themselves have several reasons as to why they do not take the warning seriously such as a view that the threats as probably not real or not harmful, a view of threat countermeasures as probably not effective, costly, and difficult to implement, while also thinking of the task as not their job. Their findings point to the need to describe users' actions, include information about threat probability, use color to represent threat severity, include information about threat consequences, provide users with specific instructions about how to avoid the threat, directly contrast potential losses from the attack with estimates of how much time will be required to implement the recommended actions to prevent the attack, and provide users with information about what their response accomplished once they respond to the warning message.

4 Mental Models in Specific Cybersecurity Domains

This section offers perspectives in mental model development in several domains. Fig. 2 shows the taxonomy of the various domains, which are further elaborated upon below. Each of these domains focuses on specific platforms, technologies, user types, social factors, tools, or applications.



Mental models in cybersecurity domains	Platform: Internet and the Web, media
	Technology: phishing, firewalls, single sign-on (SSO), HTTPS
	Type of user: home users, child users, journalists, CIOs
	Social factors: security, trust, privacy
	Tools: cryptography, machine learning
	Application: smart homes, IoT, cryptocurrencies

Fig. 2. Taxonomy of mental models in cybersecurity domains

4.1 Platform

In this subsection, we review existing work in mental models of cybersecurity that investigate how platforms such as the Internet, Web, and media influence mental model development.

Mental Model of the Internet. An analysis of cybersecurity mental models by delving into mental models of the underlying Internet itself was performed in [55], where the work examined users' mental models of how the internet works and their privacy and security behavior in today's Internet environment. The authors investigated users with computer science or related technical background against users with no technical or computational background in the field. The findings were consistent with other previous studies, where users with no technical background (otherwise referred to as non-experts in previous studies) had simpler mental models that omitted Internet levels, organizations, and entities in their design. Users with more articulated technical models perceived more privacy threats, possibly driven by their more accurate understanding of where specific risks could occur in the network. In observing the experts, the authors identified patterns in their conceptual models of the network and awareness of network related security and privacy issues. The authors suggest that user perceptions vary as a function of their personal experiences and technical education level. Users' technical knowledge partly influences their perception of how their data flows on the Internet. However, their technical knowledge does not seem to directly correlate with behaving more securely online. The authors further suggest that regardless of their technical knowledge, participants seem to have made most of their privacy-related decisions based on their experiences and cues. The authors found that there is mixed and indirect evidence of whether an accurate mental model and more advanced Internet knowledge are associated with more secure online behavior.

Additional work on risk assessment of expert and non-expert users of the Web is in [56]. Building upon work in [17], the authors explored how mental model approach can be combined with individualization of security interventions. Here, the authors use card sorting to qualitatively study how users (expert and non-expert) perceive risks on web sites. The authors propose four strategies on how to effectively improve security interventions through individualization. The first strategy deals with emphasizing unknown risks, where an emphasis on unknown risk is suitable for behavioral data. In the second strategy, the mental model is enhanced by concrete implications to make the communication more effective. The third strategy is related to the perception of risk communication. The emphasis is on making the communication relatable to the users to make it effective. Finally, the fourth strategy increases the granularity of mental models, where the level of detail of the user's mental model can identify the gaps in knowledge, the concreteness of current knowledge, and the individual's perception of the risk. The findings suggest that the granularity of the mental models needs to go beyond a lay–expert-user dualism. Their findings of users' mental models of security interventions support the notion that for most comprehensive and effective risk communication, security interventions need to be individualized.

Role of Media. The role of media in forming mental models was studied in [7], where the authors examined the relationship between computer security and fictional television and films. Participants in their study were shown six clips from television series and films depicting computer security topics. The authors found that merely exposing users to the concept of computer security may improve their understanding or awareness. However, inaccurate, and exaggerated portrayals could also harm development of healthy mental models. This is because people's ability to correctly recognize evidence of security breaches depends on their idea of what security incidents look like. Users draw conclusions about what is (not) realistic about computer security in fictional media using a variety of heuristics, most of which are either entirely non-technical or only partially grounded in technical understanding. The findings of their work indicated that security researchers and educators should take the effects of fictional portrayals into account when trying to teach users about security concepts and behaviors, supporting the findings of [57, 58] that fictional media can be a major source of security information for users.

Work in [59] analyzes three issues that users face when dealing with cyber security. These issues are related to users' conceptualization of passwords, antivirus protection, and mobile online privacy. Although the security industry provides users with ample security advice to help them stay informed about the latest threats and the best security practices, many users remain vulnerable because of noncompliance with security policies and the recommended security advice. While most of the security communications focus on the action level, the authors suggest a supplementary approach that uses metaphors and graphical explanation to facilitate users' understanding of new security concepts. Specifically, the work proposed an online interactive comic series called Secure Comics for this purpose. The authors identified three challenges unique to usable security – (1) users are typically interested in security, (2) security systems are complex and abstract, and (3) users have poor mental models of security. The goal of the interactive comic series was to help to motivate learners' interest in the challenges mentioned, and emphasized the role of educational efforts supplementary to technical, legal, and regulatory approaches for a holistic solution to securing computer systems.

4.2 Technology

Cybersecurity is a vast domain comprising of a plethora of technologies, and users accordingly develop mental models of the different technologies that they interact with. In this subsection, we review literature in mental models of various technologies (both beneficial and detrimental) encountered in cybersecurity such as phishing, firewalls, single sign on (SSO) and secure HTTP.

Mental Models of Phishing Security. The mental models of experts and novices in relation to the prevention of phishing attacks was studied in [60], where the authors generated ten terms (updates, anti-malware, training, red team, warnings, passwords, software, authentication, encryption, and blacklist), which the participants were asked to rate the strength of the pair of each concept. The authors used Pathfinder, a statistical software tool that represents pairwise proximities in a network, to determine the relatedness of the pairing and consequently the relationship between expertise and performance [61-64]. The authors found that mental model of experts was more complex than novices in the prevention of phishing attacks.

Mental Models of Firewalls. In [65], the authors explored the users' mental models of Vista Firewall (VF), where they investigated changes to the users' mental models and the users' understanding of the firewall settings after working with both the basic VF interface and the authors' developed prototype. Their findings acknowledged the tension between complexity of the interface and the security of the system leading to the conclusion of the need for transparency. That is, if the security of the application changes because of underlying context, then the changes should be revealed to the users. The authors categorized their participants' mental models based on their drawings of VF functionalities. In the *incorrect* mental model, users had an incorrect basic understanding of the inner workings of a firewall, while in the *incomplete* mental model, users had a correct basic understanding of the firewall operation without context of network location and connection. Users with a *partially correct* model had a correct basic understanding of the firewall operation, with either the context of network location or connection. The *complete* mental model resulted in context of both network location and connection. The authors suggested that designers should consider the impact of contextual factors when designing the user interface of any security application, and users should be provided with contextual information for understanding application functionality. Mismatch between users' computer-centric perspective of their security and the firewall's changing security states could lead to dangerous errors, which in turn could change the state of the firewall. The authors suggested that providing an interactive tutorial for the firewall may help provide a platform for users to learn about the firewall and the impact of network context on firewall configuration.

Mental Models of SSO. Single Sign-On (SSO) has the potential to increase security and usability of the systems it covers by making authentication for the user of multiple applications convenient and easy to use. While previous studies by [18, 67] have documented the issues with password authentication citing human cognitive limitations as the main problem, work in [66] analyzed the complexities of aligning mental models in an SSO. They found that there was a mismatch in the users' understanding of the operation of the SSO system and its actual operation. By correcting the user's mental models of SSO to match the actual functionality of the application is effective in getting users to successfully enroll in SSO and to better perceive what SSO is doing for them. In addition, by having matching models, the users could more easily maneuver the process when new applications are added into the single sign-on, or when they are asked to perform routine maintenance on their passwords. The authors found that users have different mental models, hence it is important for HCI designers to consider this diversity when designing effective and usable SSO systems.

Mental Models of HTTPS. The mental models of end users and administrators concerning HTTPS and the types of attackers that HTTPS protects against were studied in

[68]. The authors found that mental models of both users and administrators are developed based on the protocols and user interfaces with which they interact. They also found that many non-expert participants significantly underestimate the level of protection that HTTPS offers, whereas administrators generally have a good understanding of what HTTPS can or cannot protect against. In addition, most administrators had little conceptual knowledge of how the protocol works but were familiar with the different steps of establishing a communication. The authors found that there were differences between mental models of HTTPS between the two groups. Administrator mental models were generally protocol-based and correct even if sparse, on the other hand, the mental models of end users were sometimes not only sparse but simply wrong or non-existent.

4.3 Type of user

The preceding discussion shows that the bulk of existing work has focused on the mental models of experts and non-experts. This section reviews current work in the cybersecurity-related mental models of additional categories of users – children, journalists and chief information officers (CIOs). The differing demographics (children, journalists, CIOs) showcase the breadth of existing cybersecurity mental model research among different types of users, and is not a comprehensive summary of all kinds of users. For example, studies on the privacy and security-related mental models among the elderly [69], incident responders [70], ER medicine [71] and MTurk workers [72] offer a wide range of cybersecurity-related mental models from the perspective of different types of users.

Mental Models of Privacy and Security in Children. The increased usage of technology by children has attracted the attention of researchers in studying their mental models of privacy and security. Work in [73] examined how children ages 5-11 manage their privacy and security when they are online. The authors found that children have a reasonable understanding of some privacy and security components, but children ages 5-7 had some gaps. Children have some strategies to manage privacy and security online but rely heavily on their parents for support. Parents use mostly passive strategies to mediate their children's device use, and they largely defer addressing privacy and security concerns to the future. The authors recommended several strategies among them, building apps and websites for children ages 5-11 that incorporate learning opportunities that children can encounter through regular use of the service, development of educational resources to help children understand that other actors are involved in online activities and that these actors affect people's privacy and security online, create educational resources that are more focused on how the Internet functions may complement these resources by helping children better understand how and why certain online activities raise privacy and security concerns. The authors suggest that resources should promote direct engagement between parents and children and should focus on helping children grasp why certain practices protect privacy and security online and by doing this, parents may benefit from guidance on how to help children develop good privacy and security practices before they reach adolescence. Further work in [74] showed that children had adequate mental models of passwords, where they understood that passwords provided access control and offered privacy and protection.

Mental Models of Information Security Among Journalists. Most of both legal and technological security risks to journalists and sources in recent years have centered on digital communications technology. Journalists' mental models of information security were studied in [75]. The motivation for this study was in the shield laws, where journalists had operated with the understanding that their communications with sources were effectively protected from government interference. These shield laws prevented law enforcement from using the legal system to compel journalists to reveal their sources. The authors found that among reporters and editors, the need for security precautions was dependent on both coverage area and reporters' lack of first-hand experience with security incidents. These two areas indicate that participants' perceived security risk was primarily related to how sensitive or visible one's subject of reporting may be to powerful actors, rather than the vulnerabilities of the technologies through which that reporting is done. One security strategy referenced by the participants, was the use of face-to-face conversation as a security strategy. This strategy of avoiding the use of technology as a privacy or security measure has been previously categorized as a privacy-enhancing avoidance behavior [76]. The authors describe journalists' mental models of information security as "security by obscurity" to reflect journalists' thinking about security risk and avoidance in relation to digital communications technology. This mental model treats as "secure" any type of journalism that is sufficiently "obscure" to not be of interest to powerful actors, such as nation-states. However, the authors noted two prerequisites to the security by obscurity mental model, that being "obscure" as a journalist or journalistic organization is possible, and two, that being lower profile in this way offers a measure of security.

Mental Models of CIOs. The mental models of individuals who are responsible for managing the security culture of the organization were studied in [77]. Four mental models that were based on four growth phases were used to explain the evolution of security management. In the growth phase, CIOs recognize the security needs and acquire and implement security tools. In the integration phase, as the size of the installed base of tools increases, integration problems start to appear. In the formalization phase, required resources including security mechanisms are mobilized to be ready when they are needed. In the involvement phase, people deal with security mechanisms actively and so, effective design is necessary to prevent these security mechanisms from turning into constraints since security measures that are perceived as obstacles cannot be maintained over time. Their findings showed that a combination of all the efforts - technical, integration, formal and involving – were essential to lead the firm towards the desired security level.

4.4 Social factors

The implications of effective (or ineffective) cybersecurity are anchored on social factors such as trust, privacy, and security. This subsection looks at research in mental models of these aforementioned social factors in relation to cybersecurity.

Trust. A trust-based human factors ontology (HUFO) for cybersecurity was studied in [78]. Since humans are part of virtually all networks either as users, defenders, or attackers, they are capable of introducing risk into the network even if they are not attackers. The HUFO cybersecurity model proposed by the authors considers humans as risk factors and as risk mitigators, and incorporates metrics that go beyond the classic CIA (confidentiality, Integrity, Availability) framework. Another trust-based holistic risk assessment framework for users, defenders, and attackers is in [79]. While the primary focus in their risk assessment framework was on defenders, it also aimed to identify the differences in characteristics between trust in defenders, trust in users, and trust in attackers. An interesting perspective was provided of how attacks are perceived by attackers and defenders. The authors noted that attacks were easier to design, create, and launch from origins of the attackers' choosing, while cyber defense efforts instead were challenged with predicting and detecting attacks. In their framework, the authors suggested that trust in the human factors is composed of two main categories: inherent characteristics, that which is a part of the individual, and situational characteristics, that which is outside of the individual. Their proposed trust-based framework also accounted for differing mental models, risk postures and inherent biases. The motivation for this framework was derived from the 1996 Presidential Congressional Commission Framework for Risk Management, which incorporated standards from the environmental and human health risk assessment. Their trust-based holistic assessment framework comprised of context and problem formulation, assessment of systems, humans, risks and threats, as well as agility and decision-making options.

Privacy. Users' mental models about privacy of data flow was studied in [80] by analyzing the perception of privacy using three applications: Endomondo – an application that supports users in being active in all sports, MobilePay – a payment application which works between peers and in many stores, and Roskilde festival application - an application that users use to get information about the festival and the best places to eat and visit during this festival. For each of these applications, the authors studied interface design from the user's perspective, and also alignment with GDPR privacy laws. Their findings indicated that the participants did not see privacy to be a significant challenge. For example, the participants did not see anything wrong with sharing private data like heartbeat and GPS locations, but were wary about unwanted exchange of data where the service provider might preserve the collected data for use after the transaction. The work concluded that the mental models could be used to discuss privacy challenges and as basis for suggesting notifications and consent forms relating privacy.

4.5 Tools

Tools such as encryption and machine learning have come to define the burgeoning complexity of cybersecurity. This subsection presents a summary of mental model research in cybersecurity in these two areas.

Mental Models of Encryption. Cryptography is a pillar of modern cybersecurity, offering the capabilities of confidentiality, availability and non-repudiation. Work in [81] explores mental models of encryption by exploring the three facets of users' perception of encryption: what it is, how it works, and its role in their lives. The rationale for this work was derived from the potential of encryption technologies to circumvent the possibility of user error by transparently incorporating encryption into software and thus bypassing the user entirely, where the authors identified four mental models of encryption. First, the access control model was offered as the basic model of encryption which provides only the most minimal abstraction of access control. The black box model was developed as an extension of existing credential-based access. In the next level of abstraction labeled the "cipher" model, participants had a clear sense of how the "transformation" process of encryption works. Finally, the iterative encryption model was the most advanced, where participants explicitly described the encryption process as an iterative one involving multiple passes over the source data to produce the encrypted output. Although these four mental models of encryption vary in detail and complexity, their ultimate role is in access control. The authors suggest that perception of encryption as access control can be useful in the right contexts since a majority of participants adhered to this mental model and offered a base model upon which other models could be layered. The authors noted that improving the how of encryption was not sufficient in itself to drive adoption, rather the how and why were necessary in developing and using effective mental models. To aid in the how and why, users would need to understand the potential benefits of encryption to self and society at large. Since cryptography tends to be mathematically and computationally complex, the authors recommend that rather than offering users the technical details of encryption an effective model should focus on aligning designs and communication efforts with the functional models that the users already possess. Their findings called for security and warning indicators to be carefully designed, with an aim of being noticed, trusted, and validated by users.

Additional work in mental models of encryption has been performed in [82] explores users and non-users' mental models of end-to-end encryption of communication tools. Specifically, the authors investigated the use of secure communication tools to empower people to resist surveillance. In this study, the authors studied mental models of a hypothetical encrypted communication tool to avoid introducing bias about well-known tools. Prior work by [83] has shown that incorrect mental models are a key obstacle to the adoption of secure communication tools and other privacy-enhancing technologies. The authors found that end-to-end encrypted tools are widely used but not accurately understood, which leads to users unknowingly selecting insecure communication tools in situations where they most require privacy. The authors suggested that it is important to communicate the security properties of E2E encrypted communication tools, since users might know about the choice of specific encrypted tools for security-sensitive environments. Other important factors in adoption of encrypted communication tools were the size of the user base, and focused training sessions that reached out to specific categories of users such as activists and dissidents who would most benefit from the use of these encrypted tools.

Mental Models of Security of Machine Learning Systems. Work in [84] explored the perception of vulnerabilities in machine learning (ML) applications. Specifically, the authors conducted a first study to explore mental models of adversarial machine learning (AML). The authors focused on developers' mental models of the ML pipeline and potentially vulnerable components. Despite ML being increasingly used in industry, very little is known about ML security in practice. They identified four characteristic ranges that described practitioners' mental models in AML. The first range was concerned with the intertwined relationship of adversarial machine learning (AML) and standard security. Here, the distinction between AML and security was not clear, security threats were often taken for granted, and practitioners were less aware of AML attack scenarios. The second range was concerned with the structural components and functional relations. By crafting inputs, an attacker might deduce architectural choices within the functional structure, whereas on the other hand a key parameter from the model could be accessed unlawfully. The third range was concerned with variations in the pipelines, attacks, and defenses, where the attacker either injected specific inputs or general malicious input to the application. Finally, the fourth range corresponded to the varying levels of technical depth, where the industrial practitioners perceived ML-specific threats and defenses at a varying level of technical depth.

4.6 Applications

In this subsection, mental models of specific applications such as smart homes, IoT devices, and cryptocurrencies are surveyed.

Mental Models of Smart Homes and IoT Security. Although there has been a growing interest in smart homes technologies, the level of acceptance varies among potential users leading to varying explored the end users' perceptions, expectations, and concerns regarding smart homes [85]. The authors found that users' understanding of smart homes was superficial and their mental models seemed to be affected by the technologies advertised in the media. While users appreciated the convenience offered by smart homes, their major concerns were burglary, hacker attacks, data theft, data abuse, and collection and storage of sensitive data like bank details. Comparing the mental models of experienced users' and non-experienced users', the authors found that most mental models for smart homes had two foci. First, users were concerned about their own technical deficiencies and anticipated issues due to perceived dependence on the technology or had some difficulties with the device thus rendering the device useless. Second, users were concerned about the privacy and security of smart homes. A key recommendation of this study was access control, such as limiting Internet access to a predetermined subset of devices. Work in [86] explored how users interact with the security features of IoT devices, and found that users' mental model errors and biases, as well as their perceptions of complex conflicting system features were significant drivers of their interaction with these devices.

Mental Models of Security in Cryptocurrency Systems. The rise of blockchain-enhanced frameworks has prompted the development of Web3 technologies, which include diverse individual applications such as cryptocurrencies, non-fungible tokens, token economics, and decentralized digital solutions. Among these applications, cryptocurrencies (namely Bitcoin) was the first blockchain application to gain prominence. Work in [87] explored mental models of security and privacy risks in cryptocurrencies. The authors studied two mental models – the incomplete mental model, where participants knew most of the details such as the requirement of a sender to sign the transaction with a generated key, and knowledge of addresses as the payment destination. The inaccurate mental model on the other hand explored misconceptions of cryptocurrency systems such as anonymity, cryptographic keys, and fees. In the inaccurate model, some participants assumed that cryptocurrencies were based on central management entities or direct end-to-end connections between users. The authors found although many misconceptions did not jeopardize users' security or privacy, major misconceptions related to the functionality and management of cryptographic keys are not compensated by the cryptocurrency tools. The authors observed that wallet interfaces shaped the way participants perceived the blockchain location (centralized vs. decentralized), its functionality (persistent, transparent), and the users' role within the cryptocurrency system. The authors recommend that modifications of the interface of cryptocurrency management tools can prevent security and privacy threats caused by incorrect mental models. Inaccurate mental model can lead to devastating loss in term of monetary value. The authors suggest that cryptocurrency tools should perform encryption by default and inform the users about this safety measure, while advocating for designers to add cues and visualizations to explain to the users which security measures (e.g., encryption) are implemented so that users can make informed trust decisions.

5 Discussion

This paper has provided a brief overview of existing research in mental models that dictate our cybersecurity activities. These models span the spectrum of various tools, technologies, platforms, user types, and applications. Below, we discuss some implications of the role of standardization, social factors, workforce development, and emerging technologies such as Web3 in influencing users' mental models of cybersecurity.

Standardization: Drawing upon metaphorical nomenclature as well as formal methods, researchers have found that people think about cybersecurity in different ways. While the distinct ways of thinking about cybersecurity are a function of an individual's expertise and background, it also poses a challenge for the designers and developers of security tools. Further, a single user might use different mental models for different applications or activities. For example, a mental model about password security might be different from a mental model about blockchain keys. One approach to addressing the challenge of a plethora of models would be curation and standardization efforts, similar to the MITRE Common Vulnerabilities and Exposure (CVE) database that is maintained by a consortium of organizations in academia, industry, and government sectors as a central point of information dissemination about security vulnerabilities. Similar efforts in identification and development of mental models, with inputs with diverse disciplines such as computer science, cyberpsychology, criminology, economics, marketing, information systems, and others would help to identify widely used mental models among different types of users for differing applications.

Social factors: Mental models are an outcome of several attributes of an individual, including but not limited to experience with the artefact, linguistic interpretation of cybersecurity advice, and socio-cultural factors. For example, the findings of [88] indicate that the big five personality traits were an important predictor of cybersecurity-related behavior. Further, studies of technology usage show distinct trends in adoption and usage behavior such as Twitter usage in three major cities around the world [89], demographic patterns of Facebook usage [90], and interactive radio formats for audience engagement in Africa [91]. Further research is required to identify patterns in adoption of cybersecurity-related activities around the world and finding their antecedents in socio-cultural cues.

Workforce development: Cyber workforce development is a pressing need. As identified in [92], the importance of social fit in the highly complex and heterogenous cyber workforce is crucial in building secure networks. The authors identified six assumptions for the future of cybersecurity workforce development - the requirement for systemic thinkers, team players, a love for continued learning, strong communication ability, a sense of civic duty, and a blend of technical and social skill. Since cybersecurity is a critical component enabling social aspects of human behavior in networks [93], cybersecurity professionals must understand both the technical and human aspects of interaction with networks.

Web3: The rise of blockchain-fueled applications such as cryptocurrencies, non-fungible tokens, and digital records for provenance have ushered in new technologies for users. The social media platforms of Web2 have also undergone a parallel evolution with newer platforms such as Discord, Notion, and others that are heavily leveraged by the Web3 ecosystem. Blockchain-based applications have steadily moved into mainstream discourse, starting with their initial foray from into cryptocurrencies such as Bitcoin and Ethereum. Since then, Ethereum has become a blockchain platform enabling the development of distributed applications (dApps). Additionally, blockchain architectures such as public, private and consortium allow varying levels of participation, security, and privacy, which also require the development of a new vocabulary and associated mental models to understand the evolving terminology [94]. These new technologies need new mental models to process concepts related to public key cryptography and the security of private keys, the visibility of wallet addresses, hashing, and terms such as immutability and consensus which form the foundation of blockchain systems. The security threats confronting these technologies are diverse and evolving in complexity as well [95], some examples of which include ransomware strains that attack IoT devices [96] and niche supply chains [97], and botnets which were once at the forefront of spam spread are now actively used in mining cryptocurrencies [98].

6 Conclusions

Users engage differently with the security recommendations, and their activities range from ignoring, weakly engaging, exactly engaging, or strongly engaging with these recommendations. Each of the models surveyed in our paper offers users the ability to ignore (not engage) or engage to some degree (weak, exact, or strong) with the security advice. Ultimately, it is users' mental models that determine their level and intensity of engagement. This paper discussed mental models in cybersecurity from over a decade of research that explored how people perceived threats and associated security measures when interacting with technologies of different kinds. The findings of this survey indicate that mental models are diverse, and that no single mental model can be used to describe a definitive way of thinking about cybersecurity. This points to the need for continuous research to discover mental models that users develop in response to their situational awareness, newer technologies, and threats, so that security countermeasures can be designed in response to these mental models. Further, the design and implementation of security measures has to be driven by inputs from a variety of stakeholders, since the measures are often developed by experts for non-experts. Users leverage their cybersecurity mental models as a primary defense mechanism in understanding and responding to the protection of their data and networks, and so mental model research in cybersecurity will need rigorous modeling with an increased understanding of how non-experts navigate environments that are increasing in technological complexity.

References

1. Threatlocker: 12 steps to protect against ransomware. https://www.threatlocker.com/12-steps-to-protect-against-ransomware/, last accessed 2022/5/16.
2. IBM Cyber Security Intelligence Index Report (2021). https://www.ibm.com/security/threat-intelligence/, last accessed 2022/5/16.
3. Richardson, G. P., Andersen, D. F., Maxwell, T. A., Stewart, T. R.: Foundations of mental model research. In Proceedings of the 1994 International System Dynamics Conference (1994).
4. Rowe, A.L., Cooke, N.J., Hall, E.P., Halgren, T.L.: Toward an online knowledge assessment methodology: Building on the relationship between knowing and doing. Journal of Experimental Psychology: Applied, 3-47 (1996).
5. Volkamer, M., Renaud, K.: Mental models–general introduction and review of their application to human-centered security. In Number Theory and Cryptography, 255-280, Springer, Berlin, Heidelberg (2013).
6. Morgan, G., Fischoff, B., Bostrom, A., Atman, C. J.: Creating an expert model of the risk. Risk Communication: A Mental Models Approach, 34-61 (2002).
7. Fulton, K. R., Gelles, R., McKay, A., Abdi, Y., Roberts, R., Mazurek, M. L.: The effect of entertainment media on mental models of computer security. In Proceedings of the Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019), 79-95 (2019).

8. Castelfranchi, C., Falcone, R.: Trust is much more than subjective probability: Mental components and sources of trust. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (2000).
9. FBI. 2016 Internet crime report. https://www.fbi.gov/news/stories/ic3-releases-2016-internet-crime-report, last accessed 2022/5/16.
10. Akhawe, D., Felt, A. P.: Alice in warning-land: A large-scale field study of browser security warning effectiveness. In Proceedings of the 22nd USENIX Security Symposium, 257-272 (2013).
11. Porter-Felt, A. P., Reeder, R. W., Almuhimedi, H., Consolvo, S.: Experimenting at scale with google chrome's SSL warning. In Proceedings of the SIGCHI conference on human factors in computing systems, 2667-2670 (2014).
12. NCSC. The problems with forcing regular password expiry**.** https://www.ncsc.gov.uk/blog-post/problems-forcing-regular-password-expiry#:~:text=The%20NCSC%20now%20recommend%20organisations,of%20long%2Dterm%20password%20exploitation, last accessed 2022/5/16.
13. Cotoranu, A., Chen, L. C.: Applying Text Analytics to Examination of End Users' Mental Models of Cybersecurity. AMCIS 2020 Proceedings,10 (2020).
14. Camp, L. J.: Mental models of privacy and security. IEEE Technology and society magazine, 28(3), 37-46 (2009).
15. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. Science, 211(4481), 453-458 (1981).
16. Blythe, J., Camp, L. J.: Implementing mental models. In 2012 IEEE symposium on Security and privacy workshops, 86-90 (2012).
17. Wash, R., Rader, E.: Influencing mental models of security: a research agenda. In Proceedings of the 2011 New Security Paradigms Workshop, 57-66 (2011).
18. Adams, A., Sasse, M. A.: Users are not the enemy. Communications of the ACM, 42(12), 40-46 (1999).
19. Prettyman, S. S., Furman, S., Theofanos, M., Stanton, B.: Privacy and security in the brave new world: The use of multiple mental models. In Proceedings of the International Conference on Human Aspects of Information Security, Privacy, and Trust, 260-270 (2015).
20. Houser, A., Bolton, M. L.: Formal mental models for inclusive privacy and security. In Proceedings of SOUPS 2017.
21. Degani, A., Heymann, M.: Formal verification of human-automation interaction. Human factors, 44(1), 28-43 (2002).
22. Wash, R.: Folk models of home computer security. In Proceedings of the Sixth Symposium on Usable Privacy and Security, 1-16 (2010).
23. Wash, R., Rader, E.: (2015). Too much knowledge? security beliefs and protective behaviors among united states internet users. In Proceedings of SOUPS (2015).
24. Brase, G. L., Vasserman, E. Y., Hsu, W.: Do different mental models influence cybersecurity behavior? Evaluations via statistical reasoning performance. Frontiers in psychology, 8, 1929 (2017).
25. Chen, J.: Risk communication in cyberspace: a brief review of the information-processing and mental models approaches. Current opinion in psychology, 36, 135-140 (2020).
26. Agrawal, N., Zhu, F., Carpenter, S.: Do You See the Warning? Cybersecurity Warnings via Nonconscious Processing. In Proceedings of the 2020 ACM Southeast Conference, 260-263 (2020).
27. Proctor, R. W., Vu, K. P. L.: Human information processing: an overview for human-computer interaction. The human-computer interaction handbook, 69-88 (2007).

28. Breakwell, G. M.: Risk communication: factors affecting impact. British medical bulletin, 56(1), 110-120 (2000).
29. Fagan, M., Khan, M. M. H., Buck, R.: A study of users' experiences and beliefs about software update messages. Computers in Human Behavior, 51, 504-519 (2015).
30. Wogalter, M. S., Laughery, K. R., Mayhorn, C. B.: Communication-human information processing stages in consumer product warnings. In Human Factors and Ergonomics in Consumer Product Design, CRC Press, 41-67 (2011).
31. Wogalter, M. S.: Communication-human information processing (C-HIP) model. in Forensic Warning Analysis. In: Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T., Fujita, Y. (eds) Proceedings of the 20th Congress of the International Ergonomics Association, Advances in Intelligent Systems and Computing, 821 (2019).
32. Conzola, V., Wogalter, M.: A communication–human information processing (C–HIP) approach to warning effectiveness in the workplace. Journal of Risk Research, 4(4), 309-322 (2001).
33. Aliperti, G., Nagai, H., Cruz, A. M.: Communicating risk to tourists: A mental models approach to identifying gaps and misperceptions. Tourism Management Perspectives, 33, 100615 (2020).
34. Lazrus, H., Morss, R. E., Demuth, J. L., Lazo, J. K., Bostrom, A.: "Know what to do if you encounter a flash flood": Mental models analysis for improving flash flood risk communication and public decision making. Risk Analysis, 36(2), 411-427 (2016).
35. Stevenson, M., Taylor, B. J.: Risk communication in dementia care: family perspectives. Journal of Risk Research, 21(6), 692-709 (2018).
36. Norman, D. A.: Some observations on mental model models. Hillsdale, NJ (1983).
37. Andrade, R. O., Yoo, S. G.: Cognitive security: A comprehensive study of cognitive science in cybersecurity. Journal of Information Security and Applications, 48, 102352 (2019).
38. Van den Berg, J. (2019, July). Grasping cybersecurity: A set of essential mental models. In European Conference on Cyber Warfare and Security, 534 (2019).
39. Nixon, J., McGuinness, B. Framing the human dimension in cybersecurity. EAI Endorsed Transactions on Security and Safety, 1(2), (2013).
40. Still, J. D.: Cybersecurity needs you! Interactions, 23(3), 54-58 (2016).
41. Hernandez, J. The human element complicates cybersecurity. Defense Systems. https://defensesystems.com/cyber/2010/03/the-human-element-complicates-cybersecurity/189831/, last accessed 2022/5/16.
42. Maier, J., Padmos, A., Bargh, M. S., Wörndl, W.: Influence of mental models on the design of cyber security dashboards. In Proceedings of VISIGRAPP (3: IVAPP), 128-139 (2017).
43. Nurse, J. R., Creese, S., Goldsmith, M., Lamberts, K. Guidelines for usable cybersecurity: Past and present. In Proceedings of the 3rd International Workshop on Cyberspace Safety and Security, 21-26 (2011).
44. Wästlund, E., Angulo, J., Fischer-Hübner, S.: Evoking comprehensive mental models of anonymous credentials. In Proceedings of the International Workshop on Open Problems in Network Security (pp. 1-14). Springer, Berlin, Heidelberg (2011).
45. Stanton, B., Theofanos, M. F., Prettyman, S. S., Furman, S.: Security fatigue. IT Professional, 18(5), 26-32 (2016).
46. Moon, B., Johnston, C., Moon, S.: A case for the superiority of concept mapping-based assessments for assessing mental models. In Proceedings of the 8th International Conference on Concept Mapping, Medellín, Colombia: Universidad EAFIT (2018).
47. Mohamed, M., Chakraborty, J., Dehlinger, J.: Trading off usability and security in user interface design through mental models. Behaviour & Information Technology, 36(5), 493-516 (2017).

48. Theofanos, M. F., Pfleeger, S. L.: Guest Editors' introduction: Shouldn't all security be usable? IEEE Security & Privacy, 9(2), 12-17 (2011).
49. Theofanos, M.: Is Usable security an oxymoron? Computer, 53(2), 71-74 (2020).
50. Asgharpour, F., Liu, D., Camp, L. J.: Mental models of security risks. In Proceedings of the International conference on financial cryptography and data security, 367-377, Springer, Berlin, Heidelberg (2007).
51. Fagan, M., Khan, M. M. H. To follow or not to follow: a study of user motivations around cybersecurity advice. IEEE Internet Computing, 22(5), 25-34 (2018).
52. Haney, J. M., Lutters, W. G.: " It's Scary… It's Confusing… It's Dull": How cybersecurity advocates overcome negative perceptions of security. In Proceedings of the Fourteenth Symposium on Usable Privacy and Security ({SOUPS}), 411-425 (2018).
53. Theofanos, M., Stanton, B., Furman, S., Prettyman, S. S., Garfinkel, S.: Be prepared: How US government experts think about cybersecurity. In Proceedings of the Workshop on Usable Security (USec), Internet Society (2017).
54. Jones, K. S., Lodinger, N. R., Widlus, B. P., Namin, A. S., Hewett, R.: Do warning message design recommendations address why non-experts do not protect themselves from cybersecurity threats? a review. International Journal of Human–Computer Interaction, 1-11 (2021).
55. Kang, R., Dabbish, L., Fruchter, N., Kiesler, S.: “My data just goes everywhere:” User mental models of the internet and implications for privacy and security. In Proceedings of 2015 SOUPS 39-52 (2015).
56. Bartsch, S., Volkamer, M.: Effectively communicate risks for diverse users: A mental-models approach for individualized security interventions. INFORMATIK 2013–Informatik angepasst an Mensch, Organisation und Umwelt (2013).
57. Abu-Salma, R., Redmiles, E. M., Ur, B., Wei, M.: Exploring user mental models of end-to-end encrypted communication tools. In Proceedings of the 8th USENIX Workshop on Free and Open Communications on the Internet (2018).
58. Ruoti, S., Seamons, K.: Johnny's journey toward usable secure email. IEEE Security & Privacy, 17(6), 72-76 (2019).
59. Zhang-Kennedy, L., Chiasson, S., Biddle, R.: The role of instructional design in persuasion: A comics approach for improving cybersecurity. International Journal of Human-Computer Interaction, 32(3), 215-257 (2016).
60. Zielinska, O. A., Welk, A. K., Mayhorn, C. B., Murphy-Hill, E.: Exploring expert and novice mental models of phishing. In Proceedings of the human factors and ergonomics society annual meeting, 59(1), 1132-1136 (2015).
61. Day, E. A., Arthur Jr, W., Gettman, D.: Knowledge structures and the acquisition of a complex skill. Journal of Applied Psychology, 86(5), 1022 (2001).
62. Dorsey, D., Campbell, G.E., Foster L.F., Miles, D.E. Assessing Knowledge Structures: Relations with Experience and Post training Performance, Human Performance, 12:1, 31-57 (1999).
63. Goldsmith, T. E., Johnson, P. J., Acton, W. H.: Assessing structural knowledge. Journal of educational psychology, 83(1), 88 (1991).
64. Rowe, A. L., Cooke, N. J.: Measuring mental models: Choosing the right tools for the job. Human Resource Development Quarterly, 6(3), 243-255 (1995).
65. Raja, F., Hawkey, K., Beznosov, K. Revealing hidden context: improving mental models of personal firewall users. In Proceedings of the 5th SOUPS (2009).
66. Heckle, R., Lutters, W. G., Gurzick, D.: Network authentication using single sign-on: the challenge of aligning mental models. In Proceedings of the 2nd ACM Symposium on Computer Human interaction For Management of information Technology, 1-10 (2008).

67. Van der Veer, G., Melguize, M.: Mental Models. In J.A. Jacko and A. Sears (eds.), The Human Computer Interaction Handbook, 52-80, Mahwah, NJ: Lawrence Associates (2003).
68. Krombholz, K., Busse, K., Pfeffer, K., Smith, M., von Zezschwitz, E.: "If HTTPS were secure, I wouldn't need 2FA"-end user and administrator mental models of https. In Proceedings of the 2019 IEEE Symposium on Security and Privacy, 246-263 (2019).
69. Fritsch, L., Tjostheim, I., Kitkowska, A.: I'm not that old yet! the elderly and us in HCI and assistive technology. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI) (2018).
70. Floodeen, R., Haller, J., Tjaden, B.: Identifying a shared mental model among incident responders. In Proceedings of the 2013 Seventh International Conference on IT Security Incident Management and IT Forensics (2013).
71. Stobert, E., Barrera, D., Homier, V., & Kollek, D.: Understanding cybersecurity practices in emergency departments. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (2020).
72. Shillair, R.: Talking about online safety: A qualitative study exploring the cybersecurity learning process of online labor market workers. In Proceedings of the 34th ACM International Conference on the Design of Communication (2016).
73. Kumar, P., Naik, S. M., Devkar, U. R., Chetty, M., Clegg, T. L., Vitak, J.: 'No telling passcodes out because they're private' understanding children's mental models of privacy and security online. In Proceedings of the ACM on Human-Computer Interaction, 1(CSCW), 1-21 (2017).
74. Choong, Y. Y., Theofanos, M. F., Renaud, K., Prior, S.: "Passwords protect my stuff"—a study of children's password practices. Journal of Cybersecurity, *5*(1) (2019).
75. McGregor, S. E., Watkins, E. A.: "Security by obscurity": journalists' mental models of information security. Quieting the Commenters: The Spiral of Silence's Persistent Effect, 33 (2016).
76. Caine, K. E.: Supporting privacy by preventing misclosure. In Proceedings of the CHI'09 Extended Abstracts on Human Factors in Computing Systems, 3145-3148 (2009).
77. Sarriegi, J. M., Torres, J. M., Santos, J.: Explaining security management evolution through the analysis of CIOs' mental models. In Proceedings of the 23rd International Conference of the System Dynamics Society, Boston (2005).
78. Oltramari, A., Henshel, D. S., Cains, M., Hoffman, B.: Towards a Human Factors Ontology for Cyber Security. Stids, 26-33 (2015).
79. Henshel, D., Cains, M. G., Hoffman, B., Kelley, T.: Trust as a human factor in holistic cyber security risk assessment. Procedia Manufacturing, 3, 1117-1124 (2015).
80. Sørensen, L. T.: User perceived privacy: mental models of users' perception of app usage. Nordic and Baltic journal of information and communications technologies, 1, 1-20 (2018).
81. Wu, J., Zappala, D.: When is a tree really a truck? exploring mental models of encryption. In Proceedings of 14th ({SOUPS} 2018), 395-409 (2018).
82. Abu-Salma, R., Sasse, M. A., Bonneau, J., Danilova, A., Naiakshina, A., Smith, M. Obstacles to the adoption of secure communication tools. In Proceedings of the IEEE Symposium on Security and Privacy, 137-153 (2017).
83. Renaud, K., Volkamer, M., Renkema-Padmos, A. Why doesn't Jane protect her privacy? In Proceedings of the International Symposium on Privacy Enhancing Technologies Symposium, 244-262 (2014).
84. Bieringer, L., Grosse, K., Backes, M., Krombholz, K.: Mental Models of Adversarial Machine Learning. arXiv preprint arXiv:2105.03726 (2021).

85. Zimmermann, V., Bennighof, M., Edel, M., Hofmann, O., Jung, J., von Wick, M. “Home, smart home”–exploring end users’ mental models of smart homes. Mensch und Computer 2018-Workshopband (2018).
86. Yarosh, S., Zave, P.: Locked or not? mental models of IoT feature interaction. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2993-2997 (2017).
87. Mai, A., Pfeffer, K., Gusenbauer, M., Weippl, E., Krombholz, K.: User mental models of cryptocurrency systems-a grounded theory approach. In Proceedings of the Sixteenth Symposium on Usable Privacy and Security ({SOUPS}, 341-358 (2020).
88. Shappie, A. T., Dawson, C. A., Debb, S. M.: Personality as a predictor of cybersecurity behavior. Psychology of Popular Media, 9(4), 475 (2020).
89. Adnan, M., Leak, A., Longley, P.: A geocomputational analysis of Twitter activity around different world cities. Geo-spatial Information Science, 17(3), 145-152 (2014).
90. Gil-Clavel, S., Zagheni, E.: Demographic differentials in Facebook usage around the world. In Proceedings of the International AAAI Conference on Web and Social Media, 13, 647-650 (2019).
91. Srinivasan, S., Diepeveen, S.: The power of the “audience-public”: Interactive radio in Africa. The International Journal of Press/Politics, 23(3), 389-412 (2018).
92. Dawson, J., Thomson, R.: The future cybersecurity workforce: going beyond technical skills for successful cyber performance. Frontiers in psychology, 9, 744 (2018).
93. Garvin, D. A., Wagonfeld, A. B., Kind, L.: Google’s Project Oxygen: Do Managers Matter? Boston, MA: Harvard Business School Review (2013).
94. Yao, W., Ye, J., Murimi, R., Wang, G.: A survey on consortium blockchain consensus mechanisms. arXiv preprint arXiv:2102.12058 (2021).
95. Carlin, D., Burgess, J., O'Kane, P., Sezer, S.: You could be mine (d): the rise of cryptojacking. IEEE Security & Privacy, 18(2), 16-22 (2019).
96. Yaqoob, I., Ahmed, E., ur Rehman, M. H., Ahmed, A. I. A., Al-Garadi, M. A., Imran, M., Guizani, M.: The rise of ransomware and emerging security challenges in the Internet of Things. Computer Networks, 129, 444-458 (2017).
97. Jarjoui, S., Murimi, R., Murimi, R.: Hold My Beer: A case study of how ransomware affected an australian beverage company. In Proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (2021).
98. Murimi, R.: Use of botnets for mining cryptocurrencies. Botnets, CRC Press 359-386 (2019).