Mental Models — General Introduction and Review of their Application to Human-Centred Security

Melanie Volkamer[1] and Karen Renaud[2]

[1] TU Darmstadt Darmstadt / CASED
Hochschulstraße 10, 64289 Darmstadt, Germany
melanie.volkamer@cased.de
http://www.secuso.cased.de
[2] University of Glasgow
18 Lilybank Gardens, Glasgow, United Kingdom
karen.renaud@glasgow.ac.uk
http://www.gla.ac.uk/schools/computing/

Abstract. The human-centred security research area came into being about fifteen years ago, as more and more people started owning their own computers, and it became clear that there was a need for more focus on the non-specialist computer user. The primary attitude fifteen years ago, in terms of how these new users were concerned, was one of exasperation and paternalism. The term "stupid user" was often heard, often muttered *sotto voce* by an IT specialist dealing with the aftermath of a security incident. A great deal of research has been published in this area, and after pursuing some unfruitful avenues a number of eminent researchers have started to focus on the end-user's perceptions and understandings. This has come from a realisation that end users are not the opponents, but rather allies in the battle against those carrying out nefarious activities. The most promising research direction currently appears to be to focus on mental models, a concept borrowed from the respected and long-standing field of Psychology and, in particular, cognitive science. The hope is that if we understand the end-user and his/her comprehension of security better, we will be able to design security solutions and interactions more effectively. In this paper we review the research undertaken in this area so far, highlight the limitations thereof, and suggest directions for future research.

1 Introduction

Over the last few years a number of different security mechanisms have been developed in order to protect users from different kinds of attacks eg. the SSL/TLS protocol. Some of these mechanisms have been formally proven to be secure and evaluated based on international security standards such as the Common Criteria or ISO 27001. However, a number of user studies [82, 95], as well as the prevalence of attacks [61] successfully targeting the human end-user, demonstrate that

many of these security mechanisms falter and fail as soon as the user is involved in the process. One big problem is the large number of security warnings users are confronted with. Users are habituated into ignoring these since they do not understand and thus perceive any risk [94].

This can either be attributed to the 'stupidity' of the user (not understanding what is secure and what is not) or the 'obtuseness' of the developers (not designing systems properly and not giving due consideration to the non-security related nature of the end-user's primary goal or task). As a solution, one could try to eliminate the end-user from the security mechanism's operation altogether. While this might work in a few cases (eg. virus scanners and firewalls), there are many contexts in which this is not advisable, for many reasons, as discussed in the paper 'Security Automation Considered Harmful?' [40]. For instance, in some cases eliminating the user could restrict functionality to such an extent that users will reject the mechanism (e.g. preventing users from visiting https web servers which do not possess extended validation certificates). There are also applications where user input is mandated by law. For example, the user has to be able to verify the correct processing of their vote when voting electronically.

Instead of eliminating the user altogether, one could try to force users to behave securely by defining corresponding policies or (long) lists of security and privacy rules and guidelines and try to compel the user to comply. In some cases users are punished for non-compliance, or at least threatened with sanctions [53]. This does not really work very well, at least in the way policies are designed nowadays [88]. For instance, policies often forbid actions that human nature almost compels eg. password sharing between colleagues in order to perform the primary goal/task effectively.

Wash and Rader [106] argue against all these options. It is far better, argue Wash and Rader and other researchers [106, 19, 5, 1], for developers to align necessary security-specific user interactions, educational endeavours and risk communication efforts with users' mental models and capabilities. Users' actions and decisions are directed by their mental models so it is crucial for designers and developers to know and understand *their* models when developing and designing security mechanisms. The design should be aligned with the end-users' mental models but not, as is nowadays often the case, purely with the developers' and designers' mental models and based on their assumptions about end-users' mental models and capabilities. Risk communication, particularly, can only be effective if it does not only depend on the nature of the risk but also on the alignment between the conceptual model embedded in the risk communication and the users' mental model(s) related to the context and reality of the risk. Risk communication is an important aspect of Human-Centered Security as it is implicit in each warning.

Mental models also influence trust and acceptance of technology: An incorrect mental model can make users mistrust insecure technologies [21]. This constitutes yet another reason for paying attention to end users' mental models.

In this paper we focus on aligning interactions and risk communication with the users' mental models. The first goal of this paper is to provide security re-

searchers with an introduction to mental models and lessons learned from other disciplines in which they have been successfully applied for many years. We also address mechanisms by which mental models are modified (either implicitly or explicitly). This is important because mental models are not static but rather change over time and differ between users or groups. The second objective is to provide an overview of existing security-related research on mental models in human-centred security. The mental models identified in the literature are summarized to inform the development of future security mechanisms and, in particular, security-related user *interactions* or the improvement of existing interactions including *risk communication*. Finally, we identify some limitations in the research, and the findings, and speculate about how these can be mitigated in future human-centred security mental model research.

2 Mental Models — Introduction and Overview

The term 'Mental Model' was first used by Craik [28] in 1943 in his book titled "The nature of explanation". Craik explained that a mental model was *a physical working model which works in the same way as the process it parallels*" (p. 4.2). Yet others described the concept of mental model before that, even though they used different terminology. Johnson-Laird [58] wrote a history of mental models and points out that Ludwig Boltzmann, writing in 1890, spoke about "constructing an image of the external world that exists only internally". The literature on mental models uses a range of terms and nomenclature, which makes the field somewhat challenging to investigate, with researchers using the following range of terms to refer to internal constructs of the world: analogy [29], metaphors [69, 56], perception [43, 97], theme [51], theory [98], internal concept [25] and reasoning [67]. For the purposes of this discussion we will use the term "mental model".

In 1983, two books with the words "Mental Model" in the title were published [57, 48]. Gentner and Stevens [48] argue that a study of mental models is beneficial since it helps us to understand human knowledge about the world. The first chapter in their book is written by Don Norman, who explains that mental models provide both a predictive and explanatory power for understanding our interaction with our environment, with other people and with technological artefacts. He also emphasises that the models are always evolving, are not necessarily accurate but that they do have to be functional.

Rouse and Morris [91] provide a definition of mental models that seems generic enough to encapsulate the meaning of all the different terminologies: "*mechanisms whereby humans generate descriptions of system purpose and form, explanations of system functioning and systems states, and predictions of future system states*" (p.49).

In the same paper, Rouse and Morris warn that any research into mental models, while it might deliver an understanding of the *what* of the mental model, cannot deliver an understanding of *how* the human uses it. For example, even if a person possesses a comprehensive mental model, and knows how to act

in a particular situation, it is particularly difficult to control for the impact of emotions [100]. Emotions are an imponderable which can make competent people (with perfect mental models) perform poorly. Consequently, even having identified mental models does not necessarily mean that they can be used to predict how people will deploy them in a particular situation.

The subsequent discussion will provide further insights into the intricacies of mental models while focusing on those aspects that are important for security researchers and developers to take cognisance of. This includes considerations of how mental models can be measured, developed and fostered.

2.1 Important properties and aspects of mental models

In 1985, Rouse and Morris [91] suggest that the two important aspects are (1) how the model is manipulated (ranging from implicit to explicit) and (2) how much discretion the human has in developing the model (ranging from none to full). They also identify a number of issues which beset mental model research. These are phrased as pertinent questions below, and a review of related literature provides a brief discussion of each:

To what extent is it possible to capture the details of mental models? Norman [80] warn that mental models are incomplete and unstable, and sometimes confused with each other. Also Jones *et al.* [59] say that mental models are "inconsistent representations" which "adapt to continually changing circumstances and to evolve over time". Given that this is true, how can one know that a study has delineated one particular mental model, in its entirety, without corruption from others? Researchers generally attempt to gain access to mental models by asking people to verbalise their understanding of a particular concept. Intuitively this seems a reasonable approach. Bostrom [15] *et al.* used interviews to determine whether the lay public have an understanding of the dangers of radon. They justify their use of interviews but acknowledge the potential for cognitive entrapment to occur. Unfortunately, Payne [83] found that when people were asked to verbalise their mental models of a particular construct, in their case bank machines, the *ad hoc* nature of the description was striking. Payne noted that people deployed mixed explanation types, where they mixed pre-existing conceptions with new insights and generally expressed something that was heterogeneous and it often seemed to evolve as they spoke. So eliciting verbalisations might well not be the best way of capturing details of mental models accurately. Also Richardson *et al.* [89] showed that attempts to elicit mental models ran the risk of distorting them.

Rowe and Cooke [92] carried out a comparison of four elicitation mechanisms: think-aloud, laddering[3], relatedness ratings and diagramming in order to determine the link between the quality of the resulting mental model representation

[3] The interviewer gives the participant a series of problem statements and asks them to identify four relevant aspects of the problem and then asks in-depth questions in order to explore linkages between answers the participant gives

and the troubleshooting ability of the participant. They found think-aloud to be particularly weak in this respect. They argue that laddering and rating appear to measure different aspects of the mental model, seeming to confirm the argument made by Staggers and Norcio [99] about the multi-faceted and dynamic nature of mental models including hierarchies of models or related models that are subtypes or specialisms of others. Rowe and Cooke also reported that the diagram quality was predictive of troubleshooting performance, suggesting that a sound mental model leads to a high quality diagram. Langan-Fox *et al.* [73] review a number of different techniques for eliciting mental models, and offer a methodology for choosing the optimal elicitation mechanism, depending on the research problem being investigated as well as the practical and theoretical considerations of the study.

Important to know: Accessing a mental model is challenging because one runs the risk of interfering with what one is attempting to measure. Even so, it is important to elicit mental models and, in the process, try to alter them as little as possible. Drawing diagrams where applicable seems to be the most promising approach.

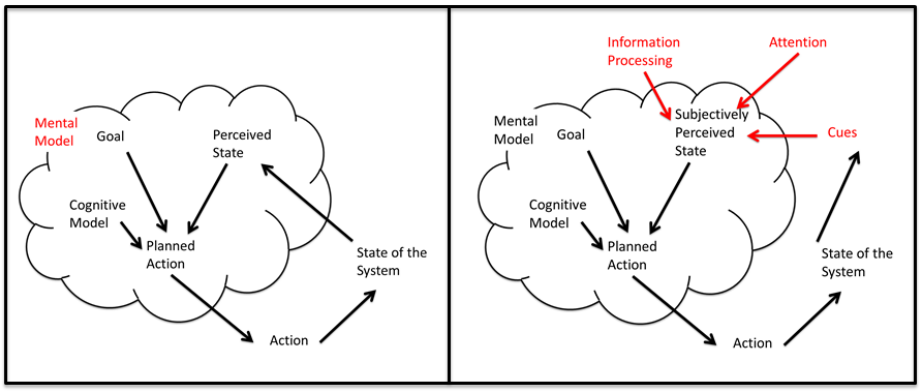

Fig. 1. Adapted from Richardson *et al.*'s [89] model

Do cues help to draw out the details of mental models? Since it is so challenging to elicit mental models it is worth considering the use of cues, both in elicitation and in assisting a user in correctly diagnosing a particular situation. Richardson *et al.* [89] apply a systems model approach[4] and claim the following primary mental model components: *intentions, perceptions, system structures* and *plans*. They see the mental model through the lens of a cybernetic loop, with the person interacting with the external system state and perceiving

[4] A conceptual model of a system which attempts to depict causatives and influences leading to outcomes.

changes in order to close the feedback loop (Figure 1 - left). They then extend their model, as shown on the right of Figure 1, to include the presence of cues, as mediating between the actual state of the system, and the perceived state of the system, showing the importance of cues that the human can interpret in order to understand the state of the system. They explain that cues are the signals that humans pay attention to, and argue that it is critical to understand how people use strategies and tactics are employed, and how such cues are interpreted. Dörner [35] explains that cues help to complete the feedback loop, thus helping people to build a more accurate mental model.

Even though cues are important it seems difficult to guarantee that cues will be efficacious. For example, visual cues are often missed, due to attentional limitations [96]. Even if someone notices a cue, sometimes they don't respond to it because they don't have the prior knowledge to understand it [87]. Even if they do have the necessary understanding, they may not react correctly due to emotional issues [39]. Finally, even if they notice the cue, know how to react, and their emotional state is such that they are able to react correctly, they may still misinterpret the cue if they do not trust its source [54].

Important to know: Cueing is simple and understandable in well-known contexts such as the theatre. When the cueing context is a human operating within a complex system, provision of a cue that will be noticed, interpreted and acted upon as intended is very difficult indeed.

How are mental models encoded in the brain (content /structure/ specificity)? Mental Models develop as humans experience life events [93], and are also impacted by the collective mental models of the group the person inhabits [59]. Such experiences can include exposure to media reports [70], culture [8] and education and training [104]. Moreover since human experience is as individual and unique as humans themselves, it is clear that mental models will differ from person to person depending on individual information processing strategies [90] even in people with similar backgrounds. The resulting mental model might well differ in representation format from person to person.

There is some speculation about whether they are analogical or spatial [83], or organized hierarchically [102] or as a cognitive collage [103]. Since the human brain can be considered one of the final, as yet uncharted, frontiers, it is possible that any attempt to pin down the precise mental model storage mechanism will be futile.

There is likely to be a difference in specificity too, especially when one compares novices and experts. Staggers and Norcio [99] provide some references to show that experts' mental models are more abstract and richer than those of novices [112]. They refer to the expert model as a *macro* model, that represents a higher level of functioning than novices' models. In addition to being more abstract, Thatcher and Greyling [102] also report that experts have more complete and detailed mental models. It is clear then that novices have an understanding of a particular problem situation, but do not possess sufficiently abstract mental models which enable them to extract core principles from such an understand-

ing. This means they cannot necessarily apply their models to solve problems which are different in specifics but generically similar.

Important to know: Mental models are based on past general and educational experiences. Moreover, it seems that the process which makes someone an expert also results in an abstracting process, so that they develop generic mental models which can be applied to a wide variety of contexts.

How should the development of correct mental models be encouraged? If one wishes to impart a mental model one always has to contend with the person's prior knowledge and experience. In terms of topic-specific prior knowledge there are three possibilities:

The topic is completely new to the learner. How best ought we to develop mental models in the minds of learners? It is important to ascertain people's understandings and assumptions, and to ensure that one matches educational delivery to these [110]. If the new material is not 'pitched correctly the lecturer runs the risk of either boring or confounding the listener. If this happens the intended mental model will not develop.

The learner has a correct but limited mental model of the topic. One should try to elicit some sense of what the person already knows [26] and then try to provide links to this knowledge so as to ease extension of the pre-existing mental model [101].

The learner has an incorrect mental model of the topic. Chapman and Ferfolja [23] outline the consequences of poor learning, and consequent imperfect mental models. They term the outcome as disruptive. It is especially damaging if further learning needs to build on a mental model that is faulty since the misunderstanding then impacts the entire learning process. Bain [7] explain that we need to prove that mental models are incorrect nefore they will realise that they need to change the way they see something. Having understood this, people can be helped to unlearn concepts, but it seems that this needs to be facilitated explicitly: one cannot merely present people who hold an incorrect understanding with the correct information [22]. This is time-consuming and effortful and so this situation should be avoided if possible.

One of the most common examples of incorrect models is demonstrated by shared, incorrect yet commonly held *folk models* [30]. Some examples include beliefs about diabetes by Bangladeshi British (incorrect eg. believing living in Britain caused diabetes) [49], shared models of infection in an English suburban community (incorrect eg. feed a cold, starve a fever) [55], understanding about heat loss from houses (incorrect eg. closing off rooms preserves energy) [65]. The danger of incorrect folk models is that they resist scientific checks which would highlight their flaws and are too easily generalised to situations they were never meant to be applied to [31]. A primary example of the resistance of folk models to change is the MMR controversy in the UK, where many parents believed that the vaccine would lead to autism, despite scientific evidence to the contrary [18]. It is difficult to challenge folk models, since they are so resistant to change.

People's mental models are formed by life experiences, and informed by educational activities. As security researchers and practitioners, we need to under-

stand how to design our educational endeavours so as to maximise the development of correct mental models. Morey and Frangioso[77] present six principles of effective learning: (1) acknowledging the person's existing mental models (2) fostering an understanding of the complexity of human-machine systems (3) challenge unthinking assumptions (4) listen to learners in order to understand where they are coming from (5) observe, assess, design and then implement (6) let learners teach others. These seem to deal with all of the possibilities mentioned above.

Important to know: In First Aid, the golden rule is "Do No Harm". In teaching, that should be our mantra too [17]. An effective teacher will first determine pre-existing assumptions, then challenge those that are faulty, then pitch the new material in order to match pre-existing knowledge, and finally facilitate peer to peer discussions to ensure that the new material is emphasised and remembered [23].

In conclusion. Mental models are obviously complicated multi-faceted mental entities, they are dynamic, based on individual experiences, are not easily described, combine many different modalities, are different between different groups and in particular between lay persons and experts, and are very difficult for researchers to capture. Finally, any attempt to capture the details of mental models are likely to change such models, and researchers have to be sure that they do not affect mental models during the process of measuring them. Moreover, when we become aware of an incorrect mental model we need to confront the holder(s) with the inconsistencies and imperfections of this model, so that we can guide them towards a correct model.

2.2 Example Applications of Mental Models

Mental model research has been deployed successfully in a variety of contexts.

Application areas. For example, Littman *et al.* [75] instructed programmers to develop two different kinds of mental models of software: the first was a systematic strategy and the second the as-needed strategy. Using the first, the programme attempts to trace the data flow all the way through the code where the latter focuses attention only on local code without consider antecedents. They found that the former strategy was by far the more effective in leading to better software maintenance. Pfeffer [84] explains how best to change the mental models of senior managers in organisations. Converse *et al.* [27] describe how best to foster shared mental models so that teams operate more effectively. Finally Nemire [79] relates how an incorrect mental model of how roller coasters operate led to a fatality, and argue for the need for specific educational efforts in order to address incorrect mental models.

Risk Communication vs. HCI. Fischoff *et al.* [44] write about the importance of aligning risk communication with people's perceptions, understandings and assumptions. They also explain that emotions and social processes play a

role in how people make use of mental models in making risk decisions. They acknowledge these confounding factors but argue for the need to get the cognitive aspects right, which *can* be controlled. An example where emotions and outrage overcame the best efforts of risk communicators is the MMR vaccine controversy [18].

In the human-computer interaction (HCI) field mental models have also been used to support design. Three examples serve to indicate the range of work in this area. Bates [11] designed an information searching interface that modeled a 'berry picking' metaphor which is better aligned to human information searching behaviour than traditional information retrieval interfaces. Khaslavsky [66] attempted to design an interface which was sensitive to cultural mental models but does not report on an evaluation of the interface. Finally, Donker *et al.* [34] determined that blind users had different mental models of an interface from sighted users, and designed an interface specifically for them.

3 Mental Models in Security Research

The only way to develop security mechanisms that effectively protect users against attackers is to align the design of security-related interactions, educational efforts and risk communication with users' mental models and capabilities. Users only protect themselves if their mental model includes some concept vulnerability to attacks. It might be necessary to attempt to adapt, extend or modify the target users' mental models in order maximise the possibility that users will act, and act correctly, to protect themselves from potential attacks.

The common knowledge about mental models presented in the previous section offers a suitable launching pad for a discussion of their use in security-related research. In particular, security researchers and developers of security mechanisms and security critical applications need to understand how end users mentally model their systems, and the security thereof (with or without taking additional security-related actions), and how they understand the effects of their actions on these systems.

As a consequence, security researchers have been conducting mental model research from different perspectives and using a variety of different methods to elicit their details. Many different methods have been applied, including: different types of interviews, user studies, and card sorting. Some researchers carry out the research in order to design security-related interactions more effectively, some to communicate risk, some to tailor educational efforts, and some to evaluate whether specific user interactions and interfaces lead to appropriate mental models. Such models are crucial in leading users to make informed security decisions.

Here we provide an overview of the most important research that has been conducted in the area of human-centred security, presenting findings and limitations. We also present and discuss research on mental model research which is intended to improve security interfaces, such as firewalls, anonymous credentials and Single Sign-On. We commence this discussion by presenting concrete

security-related models identified in literature and continue by summarizing the general findings.

3.1 Concrete Mental Models of Threats and Security Mechanisms

Jean Camp [19] identified mental models of security and privacy based on a literature review, i.e. mental models that are currently indirectly or directly being used to communicate security and privacy issues. These are:

- *Physical Security Model* (e.g. because of the lock metaphors)
- *Medical Model* (e.g. because of the phrase 'infected by a virus')
- *Criminal Model* (e.g. being arrested if you hack into a system)
- *Warfare Model* (e.g. because of intrusion detection and firewall tools)
- *Market Model* (e.g. because of people losing money)

Camp discusses these in detail, outlining their positive and negative impacts on user behaviour, with respect to security decisions. She explains that each of the models only covers parts of the security and privacy problems and thus only helps to protect against a subset of possible attacks. For instance, the *Physical Security Model* helps to protect the computer hardware but does not preserve privacy in terms of how much information one should provide using different services. Camp *et al.* [76] validated these five models using card sorting and interviews while distinguishing between experts and lay-persons. They found out that many of the studied security risks were either assigned to the Physical Security or the Criminal Models. However, they also conclude, that none of the five mental models "fit the understanding of the impression of the related risk" [76]. The Physical Security Model was also indirectly supported by results of Furman *et al.* , [46] as they show that end-users have most trust in banks (due to their physical protection) and and shops that they know in the physical world.

Rick Wash published a interview-based study [105] in which he indentifies folk models of security threats that are used by north American home computer users to decide what security software to use and which security advice to follow. He derived eight distinct folk models and explained how people used these to justify ignoring security advice. The mental model they ascribed to influenced whether or not they made backups, whether they installed anti-virus software and whether they were open to advice about their security-related actions or not. For example, some believed that one could only catch a virus if you visited a bad part of the Internet and thus did not think they were vulnerable. Some believed that they themselves were too insignificant to be worthy of a hacker's attentions, and thought they were therefore not at risk. Some also believed that hackers were only after large databases or companies and that they, as individuals, would not be attacked. Wash points out that efforts to explain how virus-protection software works to protect computers might encourage end-users to use it. Some advice was routinely ignored but not by all mental models. In summary, he identified the following models (four for risks and four for attacker types):

- *Virus — generally bad* but only high level understanding

- *Virus causes mischief* i.e. annoying problems with computer/data.
- *Virus intentionally downloaded in buggy software* computer will misbehave in time: e.g. crash or does not boot any more.
- *Virus supports crime* e.g. by stealing personal/financial information.
- *Hackers perceived as geeks* who want to impress friends.
- *Hackers are criminals* who target big fish (rich and important people).
- *Hackers support crime*, looking for large databases of info
- *Hackers are burglars* who steal personal/financial information

Kauer et al. [62] replicated this study in Germany. They found, in total, eleven folk models, with the eight from Rick Wash being mentioned with small differences. In addition, the authors identified: *Viruses are Governmental software*, *Hackers are Governmental Officials* referring to the Bundestrojaner and Staatstrojaner (engl. Federal Trojan horse) as well as *Hackers are Stakeholders with individual and opportunistic purposes* such as Anonymous[5] and the Chaos Computer Club[6]. With respect to the first two extra models, all participants said that the government only acted if there were suspicion of some crime. With respect to the latter model, participants did not fear being targeted, since these stakeholders are perceived to be more interested in media-effective targets. The government was also mentioned throughout the interviews conducted by Furman *et al.* [46]. Considering the question about whom or what users need to protect themselves from, their participants also mentioned colleagues and scammers (in addition to hackers, bad guys, and criminals).

Dourish *et al.* [36] also probed mental models of security using interviews. They identified four classes of threats:

- *Hackers* cause mischief and harm; vandalism; the least commonly identified threat.
- *Stalkers* get information online but can also continue their activities offline.
- *Spammers* advertise by means of unsolicited messaging which is perceived as a type of denial-of-service attack.
- *Marketers* invade individual privacy by surreptitiously collecting information about activities, purchasing patterns, etc.

Weirich and Sasse [108] also conducted interviews to explore perceptions of who they thought tried to get into other people's accounts and whom they targeted. The identified models for the first question were:

- *Kids*: only want to prove that they can do it, but do not cause serious harm;
- *Vandals*: plain mad, who cause serious harm;
- *Criminals*: only attack online-banking;
- *Vengeful people*: disagree with individual or organization;
- *Others*: industrial spies, terrorists, and jokers.

[5] `http://du-bist-anonymous.de/`
[6] `https://www.ccc.de/en/`

Friedmann *et al.* [45] identified mental models related to secure connections. The authors conducted 72 interviews (including drawings) and, in particular, asked how the participants to decided whether a connection was secure or not. They also asked what a secure connection meant to them. They identified five strategies with respect to the first question: (1) HTTPS; (2) lock/key icon; (3) point in transaction (main pages are usually not secured); (4) type of information requested, and (5) type of webpage. The last two are particularly interesting as this means that people assume that when sensible data is requested (like passwords or credit card information) the server ensures that the connection is secure. Similarly, they assume that organisations such as banks take due care. Regarding the second question (meaning of secure connection), the authors identified three mental models:

- *Transit*: protecting the confidentiality of information while it moves between entities.
- *Encryption*: specific mechanism for encoding/decoding messages.
- *Remote Side*: protecting data once it arrives at recipient. From the drawings, one can conclude that people with this mental model assume that the connection is either always secured, secured by default secured or unsecured since it there is no way for anyone to interfere with the transmitted message.

Benenson *et al.* [12] studied smartphone users with respect to privacy concerns and awareness. They identified two different mental models: the *Android* and the *iPhone* mental model. They confirmed that Android users did seem to be more privacy-concerned and -aware, as they significantly more often mentioned data privacy as an important factor for choosing an application. iPhone users thought that if applications needed the required access they would not necessarily ask. Moreover, iOS users were mostly unaware of application data usage. These results are explained by the different ways Apple and Google chose to inform users of applications' data usage. Furthermore, the researchers found that technical features were an important factor in informing smartphone choice. People are more likely to own an Android phone if they cared about technical features. People who were interested in technology were more likely to own an iPhone.

3.2 Mental Models to Improve Interactions in Concrete Applications (HCI)

HCI Sec researchers also studied mental models in the context of Interface design. We report here about five different types of applications.

Anonymous Credentials. There is a great deal of technical research on anonymous credential and privacy-friendly identity management. Wästlund *et al.* [107] evaluated three different interfaces and three different corresponding metaphors in a user study to explore the ideas behind anonymous credentials namely card-based, attribute-based and adapted card-based metaphors. The objective was to test whether the built mental models of anonymous credentials

are correct for any of these approaches. Therefore, they asked the participants to explain which information flows exist and who could violate anonymity. While the adapted card-based technique performed best, further improvements are necessary.

Firewalls. Raja *et al.* [85] base their research on firewall warnings on the physical mental model described in [19] after having identified a number of misconceptions with personal firewalls in [86] . They use this model to visualize the functionality of a firewall. Their comparison with the Comodo personal firewall shows that the visualization of a fireproof wall separating parts of a building helps users to develop better mental models of firewalls. While most users preferred the warnings designed by Raja *et al.*, some participants in their study mentioned that they would take these new warnings less seriously than warnings from Comodo.

Web Single Sign-On (SSO). Sun *et al.* [52] investigate, using interviews and drawings, end-users' perceptions of web SSO technology for authentication and found many misconcepts causing mistrust in the whole approach and raising privacy and security concerns. They re-designed the interfaces of a SSO solution and show that many more participants would use and trust this "new" approach.

Password Manager. Chiasson *et al.* [24] studied two different password managers and found that most of the identified problems were caused by inaccurate or incomplete mental models of the software's operation. For instance, most users do not even have a high-level understanding of how password managers function. The authors recommend more visibility (vs. more transparency) to enhance the usability of password managers.

Secure E-Mail products. Whitten and Tygar [109] conducted a cognitive walkthrough evaluation and a user study on PGP 5.0. From the cognitive walkthrough many problems were identified including those related to mental models and metaphors e.g. the lack of differentiation between private and public keys. The results from the user study showed that most users had difficulties with key management and with understanding the underlying PKI concept, leading to errors such as sending the secret unencrypted or encrypted with the wrong key.

3.3 General Findings

In this subsection, we provide the general findings on mental models in human-centered security independent from concrete mental models. We first report about a couple of general misconcepts and then about differences between different groups of users.

Problems with Diverse Antecendents. Computer users face different types of problems: common computer problems (e.g. buggy program or disk failure)

and security / privacy[7] problems. Gross and Rosson [51] conducted interviews to better understand end-user security management based on the hypothesis that people have difficulty distinguishing between these two kinds of problems. Their findings in [51] seemed to confirm this. However, in a later paper [50], Gross and Rosson showed, based on an online survey with 368 participants, that end-users did indeed make a distinction between these problems. Furthermore, they showed that end-users are more concerned with security and privacy problems than with general computer failure.

Focus on Confidentiality. Furman et al. [46] interviewed 40 people about their perceptions of security. They found that although the participants mentioned *confidentiality*, few mentioned the other two cornerstones of information security: *integrity* and *availability*. This was confirmed by other researchers. For example, in [45], the authors found out that people only considered confidentiality and encryption in their definitions of secure connections. When it comes to smartphone security, the situation is slightly different. When talking about these devices, users do actually mention availability [78] but more in terms of high mobility devices being more likely to be lost or stolen, thus referring to the availability of the device rather than the data on the device.

Others being Responsible. Gross and Rosson [51] attempted to understand end-user security management in companies and organizations. Most of their participants attributed responsibility for security to the IT staff or to the organization or both but most emphatically not themselves. Furman *et al.* [46] report a similar result in the home environment: people assign the responsibility to third parties such as the government, software companies, credit card companies, banks, or IT professionals. Also, Dourish *et al.* [36] showed that end-users were in favour of delegating responsibility, namely to technology, knowledgable colleagues, family members, room mates, organizations and institutions such as banks. According to an AOL/NCSA study[8] from 2005, most people felt that the responsibility laid with the government and with big companies (while at at least 15% felt that they themselves were responsible). Hence, it is not surprising that many studies [16, 46, 36] confirmed that users are rather relaxed when it comes to online banking as they assume the bank will take care (including ignoring warnings for such pages [16]).

In the smartphone context, King [68] shows that end-users trust applications because they assume that an in-depth evaluation has been carried out by Apple and Google before these are provided in their stores. This was confirmed by Kelley *et al.* [64] for Android users. The same observation was made in [81] within the context of electronic voting: people assume that observers will make sure that votes are properly tallied.

[7] Note, although the focus of this subsection is on security some researchers have studied both aspects together and thus those results are included here if they are not only privacy specific.

[8] `http://www.bc.edu/content/dam/files/offices/help/pdf/safety_study_2005.pdf`

An opposing finding is reported by Furnell *et al.* [47], who found that 90% of respondents (strongly) agreed with the phrase 'It is my responsibility to protect my computer from online attacks'. The difference might be caused by the question being concrete about one aspect while the other studies posed more broad-ranging open-ended questions. Moreover, it is possible that the direct question asked by Furnell *et al.* might have been misunderstood by the participants in the sense that it is not clear to them what is implicitly required; maybe they thought it just meant that they should install a virus scanner. This mismatch is a more general finding and is addressed in the next paragraph.

Lack of Awareness/Misunderstandings/Misapplication. Wash and Rader [106] explain that users often believe they *were* doing what was necessary to protect their computers. They argue that users were motivated to take necessary actions but obviously only against those attacks or threats they were aware of. Gross and Rosson [51] also found that their participants believed themselves to be able to sufficiently protect themselves, e.g. they know about Phishing, and know that they should check for urgency-related terms and typos in emails. Unfortunately, their knowledge is often out of date. Similarly, they are aware that software and passwords should be regularly updated and changed. However, regularly, to them, means six monthly. The authors of [6, 46, 47] drew similar conclusions. For instance, most of the participants in [46] could identify the evaluated trust and security seals/icons but did not understand what they meant. Different researchers [64, 41, 74] report that smartphone users, in particular Android users, express doubts about the permissions applications might have, but explain that they do not understand the explanations about permissions provided at installation time.

Researchers could also show that there are situations in which the end-users know very well how to behave but still decide, for arbitrary reasons, not to behave securely: (1) When one considers smartphone authentication, the participants in [78] mentioned that they knew they ought to use an authentication mechanism but stated that the mechanism available on the phone was inappropriate. They felt it was overly stringent if they wanted to access innocuous information such as the weather forecast. Similar findings were reported by [68] with respect to general security and privacy concerns on smartphones. (2) In [108], the authors identified social aspects in organizational environments in the context of password security as one reason: not being a nerd but being a team player; and not being paranoid. Similarly [78] reported that people explained that they did not lock their smartphones because they did not want to type passwords in front of their friends. To them, this made it look as if they wanted to hide something. Some also felt that such an action would send a message that that they did not trust their friends and this might compromise the relationship.

Misconcept of Hackers' Targets. Weirich and Sasse's [108] participants mentioned the following targets: security-conscious organisations, high-profile organizations, people with important information, and anyone who had annoyed an attacker. Wash [105] confirmed their findings with his different folk mod-

els as well as others. This suggests that many people consider themselves too insignificant to be attacked.

Dourish *et al.* [37] found that many of the non-specialist participants they interviewed in their study reported feeling that their security efforts were futile. They reported that they were unable to protect themselves from the efforts of faceless and nameless attackers. Also participants in the interviews of [108] argued that hackers could always find a way in despite their efforts. Wash [105] reports that some of his participants felt powerless to protect themselves from the efforts of hackers and therefore did not take any action to protect their computers. Others believed that all they had to do was to make it harder for hackers to get into their accounts than into those of other home computer owners, what Wash refers to as the 'speed bump' theory. These findings were confirmed by [36]: people referred to the 'unknown other who will always be one step ahead'.

Differences between novices and experts. Several security researchers [76, 4, 47] confirmed the results reported by Staggers and Norcio [99] that experts' mental models are more abstract and richer than those of novices (see also Section 2.1). For instance, Asharpour et at. [4] showed that the mental model of security risks strongly correlate with their level of expertise in security. Liu et at. [4] even tried to measure the distance between the mental models of experts and lay-persons. Experts are defined in this paper as those who know all the technical definitions of the security-related words. In a later paper, the authors [4] changed the definition to "One who has at least five years expertise in security as a researcher, student or practitioner" (so called security specialist)[9]. The authors reported differences between the security specialists and 'the others'.

Bravo-Lillo *et al.* [16] report a mental model study on how people (experts and lay persons) decide whether to ignore or follow security warnings. Experts are defined by having completed at least one year's security or privacy course or at least a one year security or privacy project. They report very clear differences between novice and expert users and thereby confirmed previously mentioned research. For example, they found that experts actively looked for vulnerabilities and considered multiple factors when they encountered a potentially risky situation. Novices, on the other hand, performed fewer security checks. Novices can either relate a warning to viruses or they consider that there is actually no problem. Furthermore, novices tended to assess the safety of an action after they performed it whereas experts tended to be more cautious and assessed the safety of a task (considering recent actions, sensitivity of information and consequences) before they embarked on it. They also found that novices were more likely to trust in the ability of large corporations to protect them. Accordingly, they expected online banking to be secure because banks traditionally have good security. They confirmed the findings of [111, 45, 63] that novices made decisions based on the look and feel of a website. On the other hand, experts agreed that bank websites with warnings were usually not trustworthy. They also identified

[9] The second definition is far more appropriate since it represents a higher level of understanding in terms of Bloom's [2] well-known taxonomy of the cognitive domain

a misconcept with respect to file storage. Novices believed that storing a file was more dangerous than opening it They thought that opening the file provided them with a safe preview but saving the file on their desktop was akin having a time bomb on their computer: the file posed a danger due to its presence.

Bartsch and Volkamer [9] study in a qualitative card sorting study how lay and expert users assess risks connected to Web sites. Their results indicate the diversity of mental models, both between the two groups and between individuals, particularly related to their preferences (e.g. concerning privacy or financial consequences). Bartsch and Volkamer conclude that it is not enough to distinguish between experts and lay persons.

Cultural & Demographic Differences. Most of the research on mental models has been conducted in the USA. Moreover, there are only a few studies considering cross-cultural issues in mental models of security and privacy. For instance, Diesner *et al.* [33] studied mental models of data privacy and security in India – a country without data protection laws at this point in time. They conducted interviews with 29 participants and used Network Text Analysis and map analysis techniques including Auto map [20, 32] to evaluate the transcripts and NetDraw [14] to visualize the results. They found that personal information, identity, and knowledge were the central contents while security-related terms are not central in people's minds. Later, Kumaraguru *et al.* [72] conducted an exploratory study in India and in the USA to compare privacy perceptions and concerns in the two countries. They suggest that Indians and people from the US differ with respect to level of concern about privacy, and people from the USA are more privacy-aware when it comes to new technologies. Kauer *et al.* [62] replicated Wash's folk model study [105] in Germany. They identified some differences. In particular, they considered the government and the 'Bundestrojaner' as possible threats as well Anonymous and the Chaos Computer Club. These were not mentioned by participants in Wash's study.

There is very little research into demographic differences in the context of mental models. Dourish *et al.* [36] identified some differences between younger and older participants in terms of who might threaten them on the Internet: young people are more likely to identify big organizations as threats than older ones. Sheng *et al.* [95] studied the susceptibility of different demographic groups with respect to phishing. The results may lead to the conclusion that women are, in general, more susceptible than men and that people between the ages of 18 and 25 will be more susceptible than the older computer users.

Different Environments. The authors of [51] concentrate on the organizational environment. They do not state this explicitly but it seems as if there is a mental model considering IT Staff as being responsible for security within organisations. Thus, it could be expected that people have different mental models about security at home and at work but this is not specifically addressed in this paper.

Muslukhov *et al.* [78], in common with other researchers, noticed differences between the mental models of *home computer* and *smartphone* security. People

assume their smartphone to be a less secure device on which to store sensible data. They attribute this to the mobility of smartphones which made them more likely to be lost or stolen. In general, their main concerns are the smartphone being lost or stolen combined with losing data such as their address book or someone dialing expensive numbers on their account.

Trust and adoption. Researchers who studied privacy critical applications have in common that they (indirectly) showed that misconception can lead to distrust and thus lead to end-users not adopting the proposed technology (e.g. in [107, 52, 60]). Thereby the confirmed the research from Castelfranchi and Falcone [21] who studied the relation between mental model in trust in general.

4 Conclusion

To conclude this overview paper, we present some pertinent limitations of mental model research and suggest directions for future research.

4.1 Limitations

This review of mental model research, with a particular focus on human-centred security, has revealed uncertainty about the following:

Validity of findings: the findings of current mental model research may not be applicable in other contexts, because existing research has been carried out with limitations on:

- ***Methodology:*** Most of the studies on mental models are based on self reported data about security behaviour which might or is very likely to be inaccurate (participants want to be seen as more security and privacy aware and concious than they actually are) and depends on the context.
- ***Heterogeneity:*** most of the studies have been carried out within one country within a limited area. We could not find any comparative study which contrasts or compares the findings from one country to that of others. Clearly the findings of such homogeneous studies need to be replicated in different contexts in order to confirm, validate, and if possible generalize their findings.
- ***Time:*** much of the research seems rather dated in 2013. Mental models are extremely dynamic, and findings from some years ago are now of questionable validity.

Design methodology: no attempt has been made to design interfaces to accommodate an understanding of mental models. This might be because these will probably differ slightly across user populations, and change during the lifetime of the system.

Measuring mental models: The question of how to identify individual mental models still remains an open question. A number of techniques have been deployed, as explained in Section 2, but none has so far been identified as being the best mechanism. Furthermore, it would be necessary to identify the mental models of a particular user before start using a security mechanism; and this should not take too much time. In addition, it is unclear how to handle the dynamic characteristics of mental models.

Identifying experts: There is no clear way to distinguish novices from experts. It is important to be able to do this, since one might want to use different metaphors to train them [56] or provide them with different interfaces [3]. One cannot simply ask people if they are expert, since humans are notoriously bad at judging their own abilities [38, 71]. One could ask people how long they have been using a particular technology, but people make use of their time differently, so time, being a weak indicator, is not a reliable predictor. The traditional way of assessing knowledge is to set a task to determine whether the person is able to complete it. That is probably an infeasible approach outside an educational setting, and, moreover might suffer from the same problems mentioned in the discussion related to accessing mental models in Section 2.

4.2 Future work

To move forward we need to ascertain how to utilise this knowledge about the different aspects of mental models. A number of aspects are particularly prescient:

1. How do we *identify* which mental model(s) a particular user ascribes to in order to adopt risk communication, education and interfaces accordingly? Does it mean that we have to ask the user a set of questions before launching an application? Should the system learn about the mental model(s) by 'observing" the user's behaviour as he/she uses the system? How do we detect changes in the user's mental model(s)?
2. On the other hand, it might be possible to design *generic* risk communication techniques, educational efforts and interfaces. It could be that we ought to design in such a way that is independent of the different concrete mental model(s) users possess.
3. How do we *predict* user behaviour based on our understanding of their mental models? There is some research proposing different approaches, such as Blythe and Camp in [13] who model Wash's folk models [105] in software agents according to Gentner and Stevens [48] in a type similar to STRIPS [42] or in terms of causalities [10] or in terms of a graph [16]. Clearly more research is needed to consolidate the findings and generate guidelines for informing design.

4.3 In Closing

This paper has presented an overview of mental model research with particular application to human-centred security. We do not claim this review to be

exhaustive but it does give a flavour of the applicable research in the area and highlights areas that require more attention.

Acknowledgement

Karen Renaud carried out this research while on a visit to Darmstadt and funded by a KIVA grant (Kompetenzentwicklung durch interdisziplinäre Vernetzung). The authors would like to thank Steffen Bartsch for his valuable input.

References

1. A. Adams and M. A. Sasse. Users are not the enemy. *Communications of the ACM*, 42(12):40–46, 1999.
2. L. Anderson, D. Krathwohl, P. Airasian, K. Cruikshank, R. Mayer, P. Pintrich, J. Raths, and M. Wittrock. A taxonomy for learning, teaching, and assessing. In L. Anderson and D. Krathwohl, editors, *A Revision of Bloom's Taxonomy of Educational Objectives, Complete Edition*, pages 212–218. Longman, 2001.
3. W. Appelt, E. Hinrichs, and G. Woetzel. Effectiveness and efficiency: the need for tailorable user interfaces on the web. *Computer networks and ISDN systems*, 30(1):499–508, 1998.
4. F. Asgharpour, D. Liu, and L. J. Camp. Mental models of security risks. In S. Dietrich and R. Dhamija, editors, *Financial Cryptography and Data Security, 11th International Conference, FC 2007, and 1st International Workshop on Usable Security, USEC 2007, Scarborough, Trinidad and Tobago, February 12-16, 2007. Revised Selected Papers*, volume 4886 of *Lecture Notes in Computer Science*, pages 367–377. Springer, 2007.
5. F. Asgharpour, D. Liu, and L. J. Camp. Mental models of computer security risks. In *WEIS: Workshop on the Economics of Information Security*, Carnegie Mellon University, 7-8 June 2007.
6. K. Aytes and T. Connolly. Computer security and risky computing practices: A rational choice perspective. *Journal of Organizational and End User Computing (JOEUC)*, 16(3):22–40, 2004.
7. K. Bain. *What the best college teachers do*. Harvard University Press, 2011.
8. M. Bang, D. L. Medin, and S. Atran. Cultural mosaics and mental models of nature. *Proceedings of the National Academy of Sciences*, 104(35):13868–13874, 2007.
9. S. Bartsch and M. Model. Effectively communicate risks for diverse users: A mental-models approach for individualized security interventions [to appear]. In *Informatik Jahrestagung*.
10. S. Bartsch, M. Volkamer, H. Theuerling, and F. Karayumak. Contextualized web warnings, and how they cause distrust. In *6th International Conference on Trust & Trustworthy Computing*, pages 205–222, 17-19 June, London, United Kingdom, 2013.
11. M. J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5):407–424, 1989.
12. Z. Benenson, F. Gassmann, and L. Reinfelder. Android and iOS users' differences concerning security and privacy. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 817–822, New York, NY, USA, 2013. ACM.

13. J. Blythe and L. J. Camp. Implementing mental models. In *IEEE Symposium on Security and Privacy Workshops*, pages 86–90. IEEE Computer Society, 2012.
14. S. P. Borgatti, M. G. Everett, and L. C. Freeman. *UCINET for Windows: Software for social network analysis*. Analytic Technologies, Harvard, 2002.
15. A. Bostrom, B. Fischhoff, and M. G. Morgan. Characterizing mental models of hazardous processes: A methodology and an application to radon. *Journal of Social Issues*, 48(4):85–100, 1992.
16. C. Bravo-Lillo, L. F. Cranor, J. S. Downs, and S. Komanduri. Bridging the gap in computer security warnings: A mental model approach. *Security & Privacy*, 9(2):18–26, 2011.
17. M. Buchmann. Teaching knowledge: The lights that teachers live by. *Oxford Review of Education*, 13(2):151–164, 1987.
18. D. C. Burgess, M. A. Burgess, and J. Leask. The mmr vaccination and autism controversy in united kingdom 1998–2005: Inevitable community outrage or a failure of risk communication? *Vaccine*, 24(18):3921–3928, 2006.
19. L. J. Camp. Mental models of privacy and security. *Technology and Society Magazine, IEEE*, 28(3):37–46, 2006.
20. K. Carley and M. Palmquist. Extracting, representing, and analyzing mental models. *Social Forces*, 70(3):601–636, 1992.
21. C. Castelfranchi and R. Falcone. Trust is much more than subjective probability: Mental components and sources of trust. In *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, pages 10–pp. IEEE, 2000.
22. J.-G. Cegarra-Navarro, S. Eldridge, and A. L. Gamo Sánchez. How an unlearning context can help managers overcome the negative effects of counter-knowledge. *Journal of Management & Organization*, 18(2):231–246, 2012.
23. J. A. Chapman and T. Ferfolja. Fatal flaws: the acquisition of imperfect mental models and their use in hazardous situations. *Journal of Intellectual Capital*, 2(4):398–409, 2001.
24. S. Chiasson, P. C. van Oorschot, and R. Biddle. A usability study and critique of two password managers. In *Proceedings of the 15th conference on USENIX Security Symposium - Volume 15*, USENIX-SS'06, Berkeley, CA, USA, 2006. USENIX Association.
25. S. R. Clegg. Ten propositions concerning security, terrorism and business. *Global Business and Economics Review*, 10(2):184–196, 2008.
26. D. Conrad. Building knowledge through portfolio learning in prior learning assessment and recognition. *Quarterly Review of Distance Education*, 9(2):139–150, 2008.
27. S. A. Converse, J. A. Cannon-Bowers, and E. Salas. Team member shared mental models: A theory and some methodological issues. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 35, pages 1417–1421. SAGE Publications, 1991.
28. K. J. W. Craik. *The nature of explanation*. Cambridge University Press, 1967.
29. Z. R. Dagher. Review of studies on the effectiveness of instructional analogies in science education. *Science education*, 79(3):295–312, 1995.
30. R. G. d'Andrade. *The development of cognitive anthropology*. Cambridge University Press, 1995.
31. S. Dekker and E. Hollnagel. Human factors and folk models. *Cognition, Technology & Work*, 6(2):79–86, 2004.
32. J. Diesner and K. M. Carley. Automap1.2 - extract, analyze, represent, and compare mental models from texts. Technical report, CMU, 2004.

33. J. Diesner, P. Kumaraguru, and K. M. Carley. Mental models of data privacy and security extracted from interviews with Indians. *55th Annual Conference of the International Communication Association (ICA), New York*, May 26-30, 2005.
34. H. Donker, P. Klante, and P. Gorny. The design of auditory user interfaces for blind users. In *Proceedings of the second Nordic conference on Human-computer interaction*, pages 149–156. ACM, 2002.
35. D. Dörner. On the difficulties people have in dealing with complexity. *Simulation & Gaming*, 11(1):87–106, 1980.
36. P. Dourish, J. Delgado De La Flor, and M. Joseph. Security as a practical problem: Some preliminary observations of everyday mental models. In *Proceedings of CHI 2003 Workshop on HCI and Security Systems*, Fort Lauderdale, Florida, 5-10 April 2003.
37. P. Dourish, R. E. Grinter, J. D. De La Flor, and M. Joseph. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing*, 8(6):391–401, 2004.
38. D. Dunning, K. Johnson, J. Ehrlinger, and J. Kruger. Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3):83–87, 2003.
39. J. A. Easterbrook. The effect of emotion on cue utilization and the organization of behavior. *Psychological review*, 66(3):183, 1959.
40. W. K. Edwards, E. S. Poole, and J. Stoll. Security automation considered harmful? In *Proceedings of the 2007 Workshop on New Security Paradigms*, pages 33–42. ACM, 2008.
41. A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: user attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, pages 3:1–3:14, New York, NY, USA, 2012. ACM.
42. R. E. Fikes and N. J. Nilsson. Strips: a new approach to the application of theorem proving to problem solving. In *Proceedings of the 2nd international joint conference on Artificial intelligence*, IJCAI'71, pages 608–620, San Francisco, CA, USA, 1971. Morgan Kaufmann Publishers Inc.
43. B. Fischhoff. Risk perception and communication unplugged: Twenty years of process1. *Risk analysis*, 15(2):137–145, 1995.
44. B. Fischhoff, A. Bostrom, and M. J. Quadrel. Risk perception and communication. *Annual review of public health*, 14(1):183–203, 1993.
45. B. Friedman, D. Hurley, D. C. Howe, E. Felten, and H. Nissenbaum. Users' conceptions of web security: A comparative study. In *CHI'02 extended abstracts on Human factors in computing systems*, pages 746–747. ACM, 2002.
46. S. M. Furman, M. F. Theofanos, Y.-Y. Choong, and B. Stanton. Basing cybersecurity training on user perceptions. *Security & Privacy, IEEE*, 10(2):40–49, 2012.
47. S. Furnell, P. Bryant, and A. D. Phippen. Assessing the security perceptions of personal internet users. *Computers & Security*, 26(5):410–417, 2007.
48. D. Gentner and A. L. Stevens. *Mental models*. Lawrence Erlbaum, Hillsdale, New Jersey, 1983.
49. T. Greenhalgh, C. Helman, and A. M. Chowdhury. Health beliefs and folk models of diabetes in british bangladeshis: a qualitative study. *BMJ: British Medical Journal*, 316(7136):978, 1998.
50. J. B. Gross and M. B. Rosson. End user concern about security and privacy threats. In L. F. Cranor, editor, *SOUPS*, volume 229 of *ACM International Conference Proceeding Series*, pages 167–168. ACM, 2007.

51. J. B. Gross and M. B. Rosson. Looking for trouble: understanding end-user security management. In *Proceedings of the 2007 Symposium on Computer Human interaction For the Management of information Technology*, page 10. ACM, 2007.
52. S. Gupta and R. P. Bostrom. Theoretical model for investigating the impact of knowledge portals on different levels of knowledge processing. *International Journal of knowledge and Learning*, 1(4):287–304, 2005.
53. M. Harris and S. Furnell. Routes to security compliance: be good or be shamed? *Computer Fraud & Security*, 2012(12):12–20, 2012.
54. R. Helm and A. Mark. Implications from cue utilisation theory and signalling theory for firm reputation and the marketing of new products. *International Journal of Product Development*, 4(3):396–411, 2007.
55. C. G. Helman. "feed a cold, starve a fever"folk models of infection in an english suburban community, and their relation to medical treatment. *Culture, Medicine and Psychiatry*, 2(2):107–137, 1978.
56. Y. Hsu. The effects of metaphors on novice and expert learners performance and mental-model development. *Interacting with Computers*, 18(4):770–792, 2006.
57. P. N. Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness*, volume 6. Harvard University Press, 1983.
58. P. N. Johnson-Laird. Mental models and thought. In K. J. Holyoak and R. G. Morrison, editors, *The Cambridge handbook of thinking and reasoning*, pages 185–208. Cambridge University Press, 2005.
59. N. A. Jones, H. Ross, T. Lynam, P. Perez, and A. Leitch. Mental models: an interdisciplinary synthesis of theory and methods. *Ecology and Society*, 16(1):46, 2011.
60. F. Karayumak, M. Kauer, M. M. Olembo, T. Volk, and M. Volkamer. User study of the improved Helios voting system interface. In *1st Workshop on Socio-Technical Aspects in Security and Trust (STAST), 2011*, pages 37–44. IEEE Digital Library, 2011.
61. Kaspersky. The evolution of phishing attacks: 2011-2013., 2013. `http://media.kaspersky.com/pdf/Kaspersky_Lab_KSN_report_The_Evolution_of_Phishing_Attacks_2011-2013.pdf`.
62. M. Kauer, S. Günther, D. Storck, and M. Volkamer. A comparison of American and German folk models of home computer security. In *Human Aspects of Information Security, Privacy, and Trust*, pages 100–109. Springer, 2013.
63. M. Kauer, F. Kiesel, F. Ueberschaer, M. Volkamer, and R. Bruder. The influence of trustworthiness of website layout on security perception of websites. In *Current Issues in IT Security 2012*, number 18, pages 215–220. 5th MPICC Interdisciplinary Conference on Current Issues in IT Security, Freiburg i Breisgau, Germany, May 7-11, 2012., Duncker & Humblot, 2012.
64. P. G. Kelley, S. Consolvo, L. F. Cranor, J. Jung, N. Sadeh, and D. Wetherall. A conundrum of permissions: installing applications on an android smartphone. In *Proceedings of the 16th international conference on Financial Cryptography and Data Security*, FC'12, pages 68–79, Berlin, Heidelberg, 2012. Springer-Verlag.
65. W. Kempton. Variation in folk models and consequent behavior. *American Behavioral Scientist; American Behavioral Scientist*, 1987.
66. J. Khaslavsky. Integrating culture into interface design. In *CHI 98 Cconference Summary on Human Factors in Computing Systems*, CHI '98, pages 365–366, New York, NY, USA, 1998. ACM.
67. T. Kindberg, A. Sellen, and E. Geelhoed. Security and trust in mobile interactions: A study of users perceptions and reasoning. In *UbiComp 2004: Ubiquitous Computing*, pages 196–213. Springer, 2004.

68. J. King. How come I'm allowing strangers to go through my phone? - Smartphones and privacy expectations;, 2013. `http://jenking.net/mobile/`.
69. R. Klimoski and S. Mohammed. Team mental model: Construct or metaphor? *Journal of management*, 20(2):403–437, 1994.
70. R. B. Kozma. Will media influence learning? Reframing the debate. *Educational technology research and development*, 42(2):7–19, 1994.
71. J. Kruger. Lake wobegon be gone! the "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of personality and social psychology*, 77(2):221, 1999.
72. P. Kumaraguru, L. F. Cranor, and E. Newton. Privacy perceptions in India and the United States: An interview study. *In The 33rd Research Conference on Communication, Information and Internet Policy (TPRC)*, September 2005.
73. J. Langan-Fox, S. Code, and K. Langfield-Smith. Team mental models: Techniques, methods, and analytic approaches. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 42(2):242–271, 2000.
74. J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 501–510, New York, NY, USA, 2012. ACM.
75. D. C. Littman, J. Pinto, S. Letovsky, and E. Soloway. Mental models and software maintenance. *Journal of Systems and Software*, 7(4):341–355, 1987.
76. D. Liu, F. Asgharpour, and L. Camp. Risk communication in security using mental models, 2008. Usable Security Website: http://usablesecurity.org/papers/liu.pdf.
77. D. Morey and T. Frangioso. Aligning an organization for learning-the six principles of effective learning. *Journal of Knowledge Management*, 1(4):308–314, 1997.
78. I. Muslukhov, Y. Boshmaf, C. Kuo, J. Lester, and K. Beznosov. Understanding users' requirements for data protection in smartphones. In *Data Engineering Workshops (ICDEW), 2012 IEEE 28th International Conference on*, pages 228–235. IEEE, 2012.
79. K. Nemire. Case study: The wrong mental model can kill you. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 51, pages 554–558. Sage Publications, 2007.
80. D. Norman. Some observations on mental models. In D. Gentner and A. Stevens, editors, *Mental Models*. Erlbaum, Hillsdale, New Jersey, 1983.
81. M. M. Olembo, S. Bartsch, and M. Volkamer. Mental models of verifiability in voting. In V. T. Steve Schneider, James Heather, editor, *4th International Conference on e-Voting and Identity (VoteID13)*, volume 7985 of *Lecture Notes in Computer Science*, pages 142 – 155. Springer, July 2013.
82. G. L. Orgill, G. W. Romney, M. G. Bailey, and P. M. Orgill. The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. In *Proceedings of the 5th conference on Information technology education*, pages 177–181. ACM, 2004.
83. S. J. Payne. A descriptive study of mental models. *Behaviour & Information Technology*, 10(1):3–21, 1991.
84. J. Pfeffer. Changing mental models: HR's most important task. *Human Resource Management*, 44(2):123–128, 2005.
85. F. Raja, K. Hawkey, S. Hsu, K.-L. Wang, and K. Beznosov. Promoting a physical security mental model for personal firewall warnings. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1585–1590, New York, NY, USA, 2011. ACM.

86. F. Raja, K. Hawkey, P. Jaferian, K. Beznosov, and K. S. Booth. It's too complicated, so I turned it off! Expectations, perceptions, and misconceptions of personal firewalls. In *Proceedings of the 3rd ACM workshop on Assurable and usable security configuration*, pages 53–62. ACM, 2010.
87. A. R. Rao and K. B. Monroe. The moderating effect of prior knowledge on cue utilization in product evaluations. *Journal of Consumer Research*, pages 253–264, 1988.
88. K. Renaud. Blaming noncompliance is too convenient: What really causes information breaches? *Security & Privacy, IEEE*, 10(3):57–63, 2012.
89. G. P. Richardson, D. F. Andersen, T. A. Maxwell, and T. R. Stewart. Foundations of mental model research. In *Proceedings of the 1994 International System Dynamics Conference*, pages 181–192, 1994.
90. I. T. Robertson. Human information-processing strategies and style. *Behaviour & Information Technology*, 4(1):19–29, 1985.
91. W. B. Rouse and N. M. Morris. On looking into the black box: Prospects and limits in the search for mental models. *Psychological bulletin*, 100(3):349, 1986.
92. A. L. Rowe and N. J. Cooke. Measuring mental models: Choosing the right tools for the job. *Human resource development quarterly*, 6(3):243–255, 1995.
93. D. E. Rumelhart and D. A. Norman. *Representation in memory*. Cognitive Science Laboratory, Center for Human Information Processing, University of California, San Diego, 1983.
94. S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The emperor's new security indicators. In *Security and Privacy, 2007. SP'07. IEEE Symposium on*, pages 51–65. IEEE, 2007.
95. S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs. Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 373–382, New York, NY, USA, 2010. ACM.
96. D. J. Simons and D. T. Levin. Change blindness. *Trends in cognitive sciences*, 1(7):261–267, 1997.
97. P. Slovic. Perception of risk. *Science*, 236(4799):280–285, 1987.
98. J. L. Spears and H. Barki. User participation in information systems security risk management. *MIS quarterly*, 34(3):503–522, 2010.
99. N. Staggers and A. F. Norcio. Mental models: concepts for human-computer interaction research. *International Journal of Man-machine studies*, 38(4):587–605, 1993.
100. B. M. Staw and S. G. Barsade. Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses. *Administrative Science Quarterly*, pages 304–331, 1993.
101. K. S. Taber. Mediating mental models of metals: Acknowledging the priority of the learner's prior learning. *Science Education*, 87(5):732–758, 2003.
102. A. Thatcher and M. Greyling. Mental models of the internet. *International journal of industrial ergonomics*, 22(4):299–305, 1998.
103. B. Tversky. Cognitive maps, cognitive collages, and spatial mental models. In *Spatial Information Theory A Theoretical Basis for GIS*, pages 14–24. Springer, 1993.
104. S. Vosniadou and W. F. Brewer. Mental models of the earth: A study of conceptual change in childhood. *Cognitive psychology*, 24(4):535–585, 1992.
105. R. Wash. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, page 11. ACM, 2010.

106. R. Wash and E. Rader. Influencing mental models of security: a research agenda. In *Proceedings of the 2011 workshop on New security paradigms workshop*, NSPW '11, pages 57–66, New York, NY, USA, 2011. ACM.
107. E. Wästlund, J. Angulo, and S. Fischer-Hbner. Evoking comprehensive mental models of anonymous credentials. In J. Camenisch and D. Kesdogan, editors, *iNetSeC*, volume 7039 of *Lecture Notes in Computer Science*, pages 1–14. Springer, 2011.
108. D. Weirich and M. A. Sasse. Pretty good persuasion: a first step towards effective password security in the real world. In *Proceedings of the 2001 workshop on New security paradigms*, NSPW '01, pages 137–143, New York, NY, USA, 2001. ACM.
109. A. Whitten and J. Tygar. Why Johnny Can't Encrypt. In *Proceedings of the 8th USENIX Security Symposium. Vol. 99*, page 1. McGraw-Hill, 1999.
110. D. T. Willingham. *Why don't students like school: A cognitive scientist answers questions about how the mind works and what it means for the classroom.* Wiley. com, 2009.
111. M. Wu, R. C. Miller, and S. L. Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 601–610. ACM, 2006.
112. N. Ye and G. Salverndy. Expert-novice knowledge of computer programming at different levels of abstraction. *Ergonomics*, 39(3):461–481, 1996.