Effectively Communicate Risks for Diverse Users: A Mental-Models Approach for Individualized Security Interventions

Steffen Bartsch, Melanie Volkamer

CASED, TU Darmstadt
Hochschulstr. 10
64298 Darmstadt
steffen.bartsch@cased.de
melanie.volkamer@cased.de

Abstract: Security interventions – such as Web warnings – currently do not work. One approach to remedy the situation is to make the communication of risks in the interventions more understandable and motivating. Mental models that users have of security have been studied to accomplish these aims, primarily to better align the intervention with the mental model of the users. However, the users' mental models are currently foremost understood in broad groups such as of lay and expert users – while risk communication literature proposes to individualize the communication. To explore how the mental-models approach can be combined with individualization, we analyze in a qualitative card-sorting study how lay and expert users assess risks connected to Web sites in this paper. Our study indicates 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). Based on these results, we propose four strategies on how to effectively improve security interventions through individualization.

1 Introduction

A current strand of research in information security considers the behavior of users interacting with security mechanisms to follow a "bounded rationality": When users behave in insecure manners, this is often due to a lack of motivation because they neglect consequences or externalities [AS99, Her10]. One way to influence the behavior to better align it with the user's actual goals is to better communicate the risks involved – for example, when warning against accessing a potentially fraudulent Web site. However, the problem with warnings in Web browsers is that they do not work [SEA+09]. Regarding risk communication, literature indicates that Web warnings are not understood, leading to suboptimal behavior [BLCDK11, DHC06]. We thus focus on warnings in Web browsers as a case in point for improving risk communication in security interventions in this paper[1].

[1] Generally, it should be the goal to reduce warnings to a minimum to prevent habituation effects. However, risks still need to be communicated when the adequate behavior cannot be chosen automatically and the decision

A recent approach in information security to influence behavior is to take the mental models of security that users have into account [Cam09, Was10]. For example, Wash and Rader [WR11] suggested to influence the mental models of users to improve security decisions and Bravo-Lillo et al. [BLCDK11] applied mental models to security warnings. However, this prior research treats users as broad groups – for example, distinguishing lay from expert users[2]. On the other hand, De Keukelaere et al. [DYT+09] showed how individualization of warnings can improve the effectiveness. Therefore, our goal in this paper is to combine the mental-models approach with the individualization of security interventions for more effective risk communication.

For this goal, we study the assessment of Web risks as an operationalization of risk perception and mental models at a higher level of detail than in prior research. We qualitatively study through card sorting how users perceive risks on Web sites both on a structural level (*What types of concepts* do the users perceive as risks?) and regarding content (*What threats and consequences* do they perceive as risks?). Since prior studies showed general differences between lay and expert users (c.f. [ALC07]), we compare lay users' perceptions with those of experts in addition to analyzing the differences on an individual level. In our study, lay and expert users show different structures of risk perceptions and differences in what they consider risks. For example, lay users more often rely on page types and abstract risks, while experts consider concrete consequences. Moreover, our research shows that the *individual* lay and expert users differ widely in these respects.

Our results support the notion that for the most comprehensible and effective risk communication, security interventions need to be individualized. We build upon the research on mental models by Wash and Rader [WR11] and propose concrete approaches to elicit and adapt the individual user's mental model. In particular, we describe strategies on how to make mental-models–based individualized security interventions practically feasible.

2 Risk communication and information security

Risk communication – the communication of potential harmful events, consequences, and probabilities – has been studied in diverse disciplines. For medical risk communications, for example, to enable profound decisions on treatment options, Rothman and Kiviniemi [RK99] state that making risks concrete is more successful in creating awareness and influencing behavior. Cognitive psychology indicates that it is important that people are able to "simulate" or imagine the antecedents and consequences of risks [KT82]. For medical risks, consequences (symptoms) that are easier to picture increase the awareness, as do testimonials of affected individuals when there is an identification with those [RK99].

In information security, risk communication has received attention from the research com-

needs to be delegated to the user. In case of Web warnings, the adequate decision often depends on too many factors of which many are not available to the system. For example, can the system decide that a user should be blocked from accessing an insecure banking site even if he only plans to check on interest rates?

[2] We refer to "lay users" as individuals not professionally related to security. We contrast these to "expert users," but emphasize that the lay–expert dualism only applies to security-related expertise and try to avoid valuation of either group.

munity regarding security warnings. Particularly, for Web security, it was shown that existing warnings are often ineffective [DTH06, ECH08]. However, the content of warnings can be improved to make users more responsive to the warnings [BvOP+09]. For example, Kauer et al. [KPV+12] found that individuals are less likely to ignore warnings when they perceived personal risks, corresponding to the experience from medical risk communication. To achieve this, De Keukelaere et al. [DYT+09] proposed to individualize warnings to make them more effective, adapting them to the users' expertise and prior experience. We follow this approach in this paper.

3 Users' perceptions and mental models of security risks

One promising way of individualizing risk communication to make it more effective is by taking the users' perception and understanding of security risks into account. Friedman et al. [FHH+02] found that security concerns differ between users' backgrounds in a study with participants from rural and urban areas. In a survey on online risk perception, Garg and Camp [GC12] analyzed how users perceive different kinds of risks and found that severity and temporal (when it will occur) factors were the most important factors. Onarlioglu et al. [OYKB12] confronted 164 users in an online study with various online threats on a test platform to explore their reactions. They argue that users can mitigate threats even without full understanding by using their intuition. However, users rely on simple cues which may be easily forged, such as the length of a URL.

Users' perception of risks is influenced by their models of how the environment reacts to actions – *mental models* [JL80]. Mental models have been employed to explain and improve the comprehensibility and influence behavior in various fields. For instance, in the risk communication for medical drugs, treatment advice in package inserts were less effective when they did not correspond to consistent mental models [JST88]. For computer security, Camp [Cam09] studied what mental models users have, considering, for instance, a physical security model, through the analogy of locks and doors, but also a warfare model. Asgharpour et al. [ALC07] found that there are differences between the mental models of expert and lay users. For security warnings, Bravo-Lillo et al. [BLCDK11] developed general mental models of expert and lay users on security warnings. They showed the differences between the two groups, with the lay-user mental model being less comprehensive and in part wrong. For lay users ("folks"), Wash [Was10] studied their mental models ("folk models") of general computer security. These include different models of the concepts of viruses or malware and of those of hackers or aggressors. Wash and Rader [WR11] argue that we need to accept that users have different mental models, but can influence the models for more secure behavior. Morgan et al. [MFBA02] suggest to elicit mental models of expert and lay users to identify those beliefs of lay users that most need correction for effectively influencing behavior.

Even though the mental-model approach seems promising for improving the risk communication in information security, it is still unclear how to effectively leverage the approach' potential. To propose actionable approaches, we extend this research on users' risk perception by combining it with the individualization of the risk communication in security

interventions (cf. Section 2). In particular, we not only consider the differences between the broad groups of expert and lay users, but emphasize the differences between individuals to these ends.

4 Study on the perception of Web risks

To explore novel practical approaches for individualizing the risk communication by applying the mental-models approach, we study and compare the perception of Web risks by lay and expert users. We employ the assessment of risks and risk perceptions as an operationalization of mental models since we expect that the mentioning of specific risks indicates the knowledge of the underlying causalities.

4.1 Study design

4.1.1 Procedure

To explore the perceptions of Web-related risks, we conducted a card-sorting experiment [RM97]. In card sorting, participants are confronted with a set of cards, in this case Web sites, and asked to group the cards into categories based on specified criteria. Card sorting has already been used in HCISec successfully – for example, to analyze how individuals restrict access in Social Networks [KBM+11] and to explore the general mental models of security of expert and lay users [ALC07]. Our goal was to motivate the participants to talk about the factors that influence their categorization. Thus, we were only interested in the process of the categorization and did not predetermine the categories to prevent bias (open card sort [ALC07]).

Specifically, we asked expert and lay users to imagine to have user accounts at 67 Web sites, which were presented to them as the cards to be sorted in the form of printed DIN-A5 screenshots of the pages ("picture sorting" giving visual clues [RM97, p. 83]). The task was to categorize the Web sites according to the consequences of their account on the respective Web site being compromised. When Web sites were unknown, the participants could inquire details and were provided general information on the type of page, the potential activities, and further information on specific request[3]. The participants were encouraged to think aloud while completing the task. Lastly, they completed an exit questionnaire for demographics.

We selected the 67 Web sites (within the range of 30 to 100 cards as suggested in literature [Spe09]) from the Top-500 most visited sites in Germany according to Alexa.com. From the first 60 of the Top-500, we removed excessive duplicates in types of Web sites (e.g. information sites, email providers). We then added missing types from the Top-500 (e.g. job portals), to arrive at a broad range of Web sites.

[3] We were aware of the potential bias that the provided information could introduce and were careful only to provide the requested information.

4.1.2 Participant sampling

We recruited seven expert and seven lay users for the experiment from personal contacts and did not offer a reward for participation. We defined experts as individuals professionally related to information security. The personal recruitment strategy allowed us to achieve a broad coverage of security experts (average age 37, youngest 28, oldest 52), including system administrators, security researchers, and security consultants.

The seven lay-user participants were recruited to cover a broad range of average users, not professionally related to information security, but not representative, with an average age of 23.1 (youngest 22, oldest 25) – 2 female and 5 male. Five lay users stated in the exit questionnaire to use the Internet several times a day, and all lay users had already conducted transactions at online shops. The lay users were evenly distributed between regularly and never using online banking and social networking sites. Similarly broad was the distribution concerning computer literacy: Confronted with computer problems, three stated that they "receive help from others," two "help others but receive help from others for difficult problems", and two "help others also with difficult problems." Thus, while limited in breadth of age groups, the sampling achieved a broad range of general expertise.

4.1.3 Analysis

The experiment was audio-recorded and the categorization process was qualitatively analyzed. Bryman [Bry88] argues that qualitative research methods result in "rich, deep" data (p. 94), as needed for the analysis of a broad range of arguments for categorizing Web sites by risks. We chose a Grounded-Theory approach [GS67] for the analysis, because of its strength in systematically identifying and categorizing concepts in rich data, and since it has been successfully employed in HCISec [ALC08]. The open coding of categorization argumentation resulted in 795 raw quotes that we then consolidated (axial coding) to eight primary and 52 secondary concepts. For example, this assessment of the TV guide tvinfo.de:

> *[They have] further information then for me... if someone sets wrong preferences for me, this will come back to me, and I will find out quickly that the information is not sent to my email address anymore*

was coded regarding *Type of page* factors: *Content: Information* (the content of the Web site is information); and *Risk-related* factors: *Consequence: Modification/Concrete* (modified preferences), *Scenario: Concrete* (someone sets wrong preferences), and *Further risk: Concrete* (will be found out quickly). *Content*, *Consequence*, *Scenario*, and *Further risk* denote the primary concepts in this quote. For internal validity, the coding was conducted by one of the authors. The coded quotes were then quantitatively analyzed by the number of mentions of each concept as described below.

The experiment was conducted in German, with the coding in English and quotes translated for inclusion in this paper.

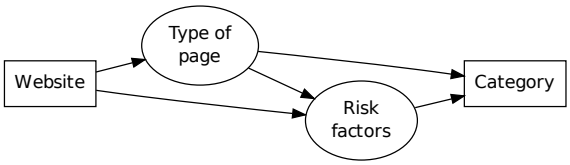

Figure 1: High-level argumentation patterns

Concept	Examples	Lay		Expert		$p < 0.05$
Type-of-page arguments		225	**58%**	134	**38%**	Yes
Activity	Publishing, shopping	97	25%	68	19%	No
Institution	Bank, forum	78	20%	50	14%	No
Context	Job, leisure	28	7%	12	3%	No
Content	Information, media	22	6%	4	1%	No
Risk-factor arguments		172	**45%**	236	**67%**	Yes
Data	Personal contacts, credit card #	101	26%	126	36%	No
Consequence	Financial loss, annoyance	63	16%	148	42%	Yes
Scenario	How to misuse data	75	19%	148	42%	Yes
Further risk	Relevance, probability	90	23%	142	40%	No
Total		385		354		

Table 1: Primary concepts used in the categorization of Web sites

4.2 Expert- and lay-user perceptions

Since prior research shows different perceptions and different mental models between expert and lay users, we expected that we can reproduce these findings. Extending prior work, we particularly focused on the concepts that the two groups employed (how they argued) and what factors were frequently mentioned:

4.2.1 Argumentation patterns

We analyzed the arguments that participants stated while categorizing the Web sites for their consequences. The argumentation followed different patterns (cf. Figure 1): One option was to argue based on the features of the page ("Type of page") in one or more of four different language constructs (cf. Table 1: *Activities* that the user may conduct on the Web site, the *Institution* that it represents, its *Context* of use, or its *Content*). Second, participants used risk factors for the categorization (the type of personal *Data* stored/requested by the Web site, the *Consequence* of an attack, *Scenarios* of how the adversary could misuse the account, and/or moderating *Further risk* factors). As indicated by the arrows in Figure 1, the line of argument could only involve either Type-of-page or Risk factors, or it could include both to explain the categorization. For example, for the quote in Section 4.1.3, the argumentation would include Type-of-page (content) and Risk factors (Consequence, Scenario, Further risk).

Table 1 shows for each primary concept how often an argument of the respective type was

Risk concept	Lay	Expert	$p < 0.05$
Data-related arguments	101 59%	126 53%	No
Concrete	41 24%	86 36%	No
Consequence arguments	63 37%	148 63%	Yes
Concrete	34 20%	112 47%	Yes
Scenario descriptions	75 44%	148 63%	Yes
Concrete	22 13%	114 48%	Yes
Further risk factors	90 52%	142 60%	No
Concrete	3 2%	65 28%	Yes

Table 2: Concreteness for different risk concepts

used to explain the categorization of a Web site[4]. Expert and lay users significantly differ in their argumentation. Experts more frequently used the risk-factor arguments, particularly the specific consequence, the risks, and the scenarios, than lay users (as the expert did in the example in Section 4.1.3). Lay users, in contrast, more often relied on the Type-of-page factors of a Web site without explicitly considering risk factors – for example, only the possible activities (*"Eventim, that's where one may buy, order tickets"*).

4.2.2 Comprehensiveness and concreteness of risk factors

Not only did lay users less often discuss risk factors than experts; when they did, they did so less structured and less concrete. Experts generally tended to approach the categorization in a more structured manner, considering the consequences more comprehensively: They more often did not rely only on one factor to decide on the categorization, but, for example, considered several possible threats. This is reflected in the quantities in Table 1 in the higher total number of risk-factor arguments for experts for the same number of cards. Similarly, experts rather formulated concrete scenarios and named the concrete consequence (*"modifies my preferences"*) or affected personal data (*"bank account data put there"*), and the concrete evaluation of specific risk factors (*"I will find out quickly"*), instead of only mentioning solely a general risk level such as *"I'd classify it as comparatively bad"* when categorizing Web sites (cf. concreteness in Table 2).

4.2.3 Differences regarding mentioned types of data and consequences

A further difference between the argumentation of expert and lay users were the priorities that they appeared to assign to specific types of data that might be affected and types of consequences that might occur when their account is compromised. Table 3 shows the most common types of data and consequences that participants referred to when they mentioned risk factors while categorizing Web sites. For example, lay users more frequently mentioned financial data when arguing with risk factors. Conversely, lay users

[4]Since each raw quote could be coded as multiple Type-of-page or Risk factors, sums and proportions do not add up. We applied a Welch Two Sample t-test on the individuals' proportions and noted in the last column for which proportion the differences between expert and lay users are significant, i.e. the null hypothesis was rejected because of $p < 0.05$.

Risk concept	Examples	Lay		Expert		$p < 0.05$
Data-related arguments		101	**59%**	126	**53%**	No
Personal	Real name, contacts	42	24%	44	19%	No
Financial	Credit card #, bank account #	32	19%	19	8%	No
Behavior	Visited pages, movement profile	8	5%	43	18%	No
Consequence arguments		63	**37%**	148	**63%**	Yes
Emotional	Annoyance, mobbing	22	13%	21	9%	No
Financial	Money loss, unauth. bank transfer	16	9%	52	22%	No
Impersonation	Posts, messages on behalf	10	6%	63	27%	Yes

Table 3: Data-related and consequence arguments used in categorization when considering risk factors

less frequently than experts discussed the data related to their behavior that is accessible through their accounts (for example, search history). Lay users also discussed specific consequences from a compromised account significantly less often, particularly the impersonation that may occur when adversaries write messages in their name.

4.2.4 Discussion

As expected from prior research on risk perception and mental models in security (Section 3), expert and lay users took very different approaches to categorize the Web sites. While we did not directly elicit the mental models for the individual participants, their argumentation patterns and emphasis on certain aspects of Web risks indicate different mental models. Beyond prior research, we found that the argumentation patterns of lay users were more concerned with the type of page, whereas experts more frequently considered risk factors. The differences in argumentation patterns underscore that communication of Web risks needs to *adapt to how the risk is perceived* (cf. Section 2).

Moreover, when categorizing Web sites, lay users considered some consequences, such as impersonation, less frequently, indicating a lower awareness of risks of those types. This is in line with prior research that states differences in the awareness of risks between lay and expert users (Section 3). Extending prior research, we found that lay users more often resorted to abstract discussions of risk factors, instead of concrete consequences or scenarios, also indicating a lack of expertise that is necessary to actually formulate concrete and detailed arguments. If the goal in risk communication is to close the knowledge gap between expert and lay users, the communication needs to *emphasize unknown risks*, but also build upon existing abstract notions of risks and extend these notions to *complete the risk picture for concreteness*.

4.3 Individual differences in risk perception

Not only did the two groups of expert and lay users differ in their risk perception as expected from prior research (Section 3), but individual participants also showed a broad

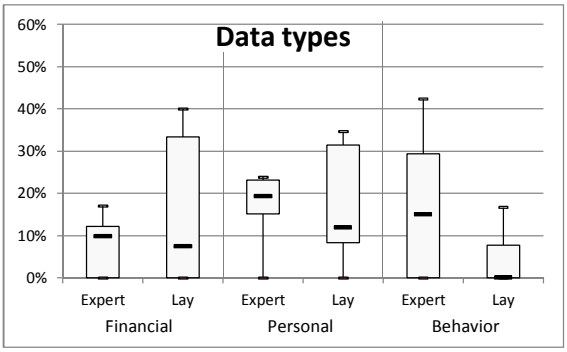

Figure 2: Proportions of data types used by the participants when considering risk factors

spectrum of perceptions and heterogeneous argumentation patterns. We selected a number of particularly striking examples of these spreads and illustrate them through box plots of the individual proportions in Figure 2 – detailing the data from Table 3.

One example for the broad spectrum can be found in the types of data that participants referred to as being at risk. While the mean for using financial-data–related arguments is 15% of the risk-related arguments for lay users (Table 3), we see how broad the spread within the 25/75 percentile is in the box plot, ranging from 0 to 33%. Thus, there appears to be a broad spread in how significant entering, for instance, a credit card number is for lay users. We see a similar pattern for behavioral data in the case of expert users. This spread might be related to the level of privacy awareness of the individual expert.

The spreads become even more pronounced when we consider cases of arguing with certain types of data in which the means of lay and expert users are similar, but show widely different spreads between participants within the groups. For example, in the case of arguing through personal data being at risk, the means of the proportions are close for lay and expert users (24%/19%, Table 3), but the box plots show that lay users are much more heterogeneous in this respect than experts.

In addition to the patterns in the box plots for the main types of data that participants used to argue about the risk while categorizing Web sites, we can observe similar patterns for the different primary risk concepts and types of consequences as shown in Figure 3.

4.3.1 Discussion

For experts, the differences in perceptions may be explained by the differences in their professional focus (i.e. system administration vs. security consultancy vs. security research). It is not surprising that their range of backgrounds that we recruited them for influenced their risk perception. Security researchers are confronted with different security problems, often of more theoretical nature, than system administrators or security consultants. Since the spread in argumentation is also wide for lay users, forming homogeneous mental models for expert and lay users appears to be more difficult than prior research indicated.

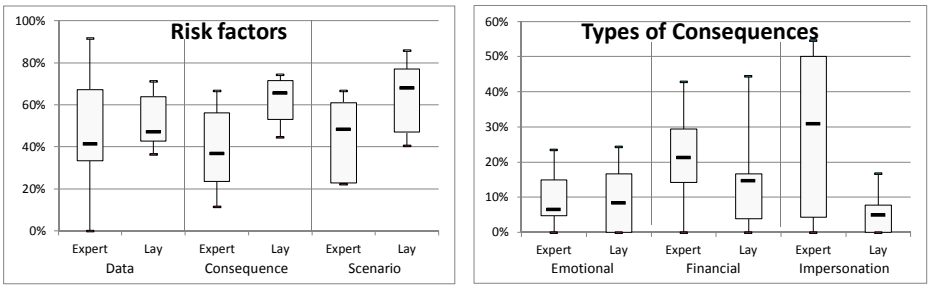

Figure 3: Box plots of the proportions of types of risk factors and consequences used when arguing about risks

These findings indicate that we need to go beyond the current approaches of mental models in warnings (Section 3) that attempt to build comprehensive mental models of expert and lay users to develop general warnings. Conversely, the *mental models need to be elicited at a higher granularity* if they should improve the effectiveness of the risk communication through individualization – for example, for subgroups of certain expertise and preferences (e.g., concerning privacy).

5 Individualization for effective risk communication

As described in Section 3, prior research on applying the mental-models approach in security primarily focused on eliciting mental models of broad groups (lay/expert users) and how we need to adapt the mental models to improve security behavior. Our study was intended to explore options of combining the mental-models approach with the individualization of security interventions for effective risk communication (Section 2). Individualization then not only concerns the external factors (e.g., type of Web site, objective risks) and the current situation (intention of the user), but also the *individual's* mental model of the specific security risks. Individualized security interventions will, in turn, not only improve the warning as a one-time security intervention, but also help to adapt the mental model as intended by Wash and Rader [WR11] to improve security behavior.

Building upon prior research on risk communication (Sections 2) and mental models (Section 3), we gather and extend here the conclusions from our study on how to apply mental models. We formulate four strategies for individualization from our results:

Emphasizing unknown risks Individual risk communication can, on the one hand, help us point out the risks that a user might otherwise overlook in the intuitive or conscious risk assessment because certain causalities are missing in their mental model. In our study, we find the approach of emphasizing unknown risks suitable for example for behavioral data, which might be at risk but was considered by lay users significantly less often than by expert users. We can thus enrich and complete the mental model by pointing out the risks

that users have overlooked.

Complete the risk picture for concreteness Second, we found that numerous risks were only considered at an abstract level by lay users. In individualized risk communication, we can strive to complete the mental model also by connecting to the abstract notions and adding the concrete implications and thus make the communication more effective (cf. Section 2). For example, the abstract idea of potential financial consequences could be enriched by that the attacker would transfer money from the user's account.

Adapt communication to how risk is perceived Third, our study showed that individuals differed in how they argued about the risks – for example, lay users significantly more often argued through the type of the page than experts. Risk communication should account for these differences, since literature shows that the communication is more effective if users can relate to it (cf. Section 2). This extends the previous strategy on connecting the communication to known abstract risks by also taking into account how the individual generally perceive risks – for example, through the context of use of a Web site, or whether it is primarily a business or leisure-time context.

Increase granularity of mental models Fourth, for all above strategies, we need to construct the user's mental model of the situation and an ideal mental model that might be deduced from expert knowledge. The level of detail of the user's mental model needs to be at a sufficient level of detail to identify the gaps in knowledge, the concreteness of current knowledge, and the individual's perception of the risk. To these ends, the granularity of the mental models need to go beyond a lay–expert-user dualism as indicated from the spread between the individuals in our study. It will also often not suffice to elicit a general mental model of security, but needs to be more focused on the specific application area, such as Web browsing in this paper.

We see two main challenges to these strategies that we will address in future work:

5.1 Eliciting the applicable mental model

Eliciting the individual mental models to enable individualized risk communication appears to be challenging because of the effort necessary to construct each mental model. However, since the level of detail only needs to suffice for identifying gaps and aspects to connect to, it should not be necessary to elicit full mental models from each individual user. Instead, we will pursue to elicit a range of typical mental models for a specific context that can then be matched to the individual user.

We see two approaches to identifying which mental model to apply for the specific user: explicit and implicit elicitation. For explicit elicitation, the user needs to provide enough data points manually – for example, in form of completing a questionnaire or quiz on causalities. In this case, the key point is how much data points are necessary and how

much are acceptable for users. We are currently developing an efficient questionnaire and will assess how context-specific the elicitation needs to be. We will also look into applying user typologies, such as based on the user's privacy perception (e.g. the Privacy Preference Scale [MAR74]) and the attitudes of users towards the technology (e.g. as in the technology acceptance models [VMGF03]).

In case of implicit elicitation, either as addition or as alternative to explicit elicitation, the applicable mental model needs to be extracted from the user's behavior. One potential approach is to use machine-learning techniques to learn from standard reactions to shown warnings ("Remind me later") or more sophisticated queries ("I know this risk"). Implicit elicitation would also help to address the necessity of reacting to changes over time – for example, from increasing expertise of the user. Further challenges include privacy concerns from the sensitive data collected in the process of elicitation, and the handling of multi-user devices and individuals' multiple devices.

5.2 Supporting the developer and other stakeholders

While users may benefit from better risk communication, several other stakeholders in the information security ecosystem may face additional challenges. For the software developer, it is already complex to create one user interface for risk communication. For individual risk communication, the complexity significantly increases without appropriate development support. We thus need clear methodologies to develop individualized risk communication artifacts and appropriate knowledge bases as libraries and APIs to reduce the effort of its implementation.

For IT administration, complications will arise in the area of training and helpdesk procedures, since different warnings and behaviors of applications may be reported due to the individualization. Accordingly, individualized risk communication will only be viable, if it proves significantly superior. Considering the track record of current warnings, for example in Web browsing, the base line might not be overly high, though.

Acknowledgments

The work presented in this paper is supported by funds of the Federal Ministry of Food, Agriculture and Consumer Protection (BMELV) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

We thank Michaela Kauer and Christoph Seikel for their support on designing and conducting the study.

References

[ALC07] Farzaneh Asgharpour, Debin Liu, and L. Jean Camp. Mental Models of Computer Security Risks. In *WEIS '07: Workshop on the Economics of Information Security*, 2007.

[ALC08] Anne Adams, Peter Lunt, and Paul Cairns. A qualitative approach to HCI research. Cambridge Univ. Press, Cambridge, 2008.

[AS99] Anne Adams and M. Angela Sasse. Users are not the enemy. *Commun. ACM*, 42:40–46, December 1999.

[BLCDK11] Cristian Bravo-Lillo, Lorrie F. Cranor, Julie S. Downs, and Saranga Komanduri. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *IEEE Security & Privacy*, 9(2):18 – 26, Mar–Apr 2011.

[Bry88] Alan Bryman. *Quantity and quality in social research*. Unwin Hyman, London, UK, 1988.

[BvOP+09] Robert Biddle, P. C. van Oorschot, Andrew S. Patrick, Jennifer Sobey, and Tara Whalen. Browser interfaces and extended validation SSL certificates: an empirical study. In *Proceedings of the 2009 ACM workshop on Cloud computing security*, CCSW '09, pages 19–30, New York, NY, USA, 2009. ACM.

[Cam09] L. Jean Camp. Mental models of privacy and security. *IEEE Technology and Society Magazine*, 28(3), Fall 2009.

[DHC06] Julie S. Downs, Mandy B. Holbrook, and Lorrie Faith Cranor. Decision strategies and susceptibility to phishing. In *SOUPS '06: Proceedings of the second symposium on Usable privacy and security*, pages 79–90, New York, NY, USA, 2006. ACM.

[DTH06] Rachna Dhamija, J. D. Tygar, and Marti Hearst. Why phishing works. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590, New York, NY, USA, 2006. ACM.

[DYT+09] Frederik De Keukelaere, Sachiko Yoshihama, Scott Trent, Yu Zhang, Lin Luo, and Mary Zurko. Adaptive Security Dialogs for Improved Security Behavior of Users. In Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, Raquel Prates, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2009*, volume 5726 of *Lecture Notes in Computer Science*, pages 510–523. Springer Berlin / Heidelberg, 2009.

[ECH08] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 2008.

[FHH+02] Batya Friedman, David Hurley, Daniel C. Howe, Helen Nissenbaum, and Edward Felten. Users' conceptions of risks and harms on the web: a comparative study. In *CHI '02 extended abstracts on Human factors in computing systems*, pages 614–615, New York, NY, USA, 2002. ACM.

[GC12] Vaibhav Garg and Jean Camp. End User Perception of Online Risk under Uncertainty. In *2012 45th Hawaii International Conference on System Sciences*, pages 3278–3287. IEEE, 2012.

[GS67] Barney G. Glaser and Anselm L. Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Transaction, 1967.

[Her10] Cormac Herley. So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. In *NSPW '09: Proceedings of the 2009 workshop on New security paradigms*, 2010.

[JL80] P.N. Johnson-Laird. Mental models in Cognitive science. *Cognitive Science*, 4(1):71–115, Jan 1980.

[JST88] Helmut Jungermann, Holger Schütz, and Manfred Thüring. Mental Models in Risk Assessment: Informing People About Drugs. *Risk Analysis*, 8(1):147–155, 1988.

[KBM+11] Patrick Gage Kelley, Robin Brewer, Yael Mayer, Lorrie Faith Cranor, and Norman Sadeh. An Investigation into Facebook Friend Grouping. In *INTERACT 2011*. Springer, 2011.

[KPV+12] Michaela Kauer, Thomas Pfeiffer, Melanie Volkamer, Heike Theuerling, and Ralph Bruder. It is not about the design – it is about the content! Making warnings more efficient by communicating risks appropriately. In *GI SICHERHEIT 2012 Sicherheit – Schutz und Zuverlässigkeit*, 2012.

[KT82] Daniel Kahneman and Amos Tversky. The simulation heuristic. Cambridge University Press, Cambridge, MA, USA, 1982.

[MAR74] NANCY J. MARSHALL. Dimensions of Privacy Preferences. *Multivariate Behavioral Research*, 9(3):255–271, 1974.

[MFBA02] Millet Granger Morgan, Baruch Fischhoff, Ann Bostrom, and Cynthia J. Altman. *Risk Communication: A mental models approach*. Cambridge Univ. Press, 2002.

[OYKB12] Kaan Onarlioglu, Utku Ozan Yilmaz, Engin Kirda, and Davide Balzarotti. Insights into User Behavior in Dealing with Internet Attacks. In *Proceedings of the NDSS Symposium 2012*. Internet Society, 2012.

[RK99] Alexander J. Rothman and Marc T. Kiviniemi. Treating People With Information: an Analysis and Review of Approaches to Communicating Health Risk Information. *J Natl Cancer Inst Monogr*, (25), 1999.

[RM97] Gordon Rugg and Peter McGeorge. The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems*, 14(2):80–93, 1997.

[SEA+09] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *USENIX Security 2009*, 2009.

[Spe09] D. Spencer. *Card Sorting: Designing Usable Categories*. Rosenfeld Media, 2009.

[VMGF03] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3):425–478, 2003.

[Was10] Rick Wash. Folk models of home computer security. In *SOUPS '10: Proceedings of the 6th Symposium on Usable Privacy and Security*, New York, NY, USA, 2010. ACM.

[WR11] Rick Wash and Emilee 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.