Trustworthy and Effective Communication of Cybersecurity Risks: A Review

Jason R. C. Nurse, Sadie Creese, Michael Goldsmith, Koen Lamberts
University of Warwick
Coventry, CV4 7AL, UK
{j.nurse, s.creese, m.h.goldsmith, k.lamberts}@warwick.ac.uk

***Abstract*—Slowly but surely, academia and industry are fully accepting the importance of the human element as it pertains to achieving security and trust. Undoubtedly, one of the main motivations for this is the increase in attacks (e.g., social engineering and phishing) which exploit humans and exemplify why many authors regard them as the weakest link in the security chain. As research in the socio-technical security and trust fields gains momentum, it is crucial to intermittently pause and reflect on their progress while also considering related domains to determine whether there are any established principles which may be transferred. Comparison of the states-of-the-arts may assist in planning work going forward and identifying useful future directions for the less mature socio-technical field. This paper seeks to fulfil several of these goals, particularly as they relate to the emerging cybersecurity-risk communication domain. The literature reviews which we conduct here are beneficial and indeed noteworthy as they pull together a number of the key aspects which may affect the trustworthiness and effectiveness of communications on cybersecurity risks. In particular, we draw on information-trustworthiness research and the established field of risk communication. An appreciation of these aspects and precepts is imperative if systems are to be designed that play to individuals' strengths and assist them in maintaining security and protecting their applications and information.**

***Index Terms*—Cybersecurity risk; information trustworthiness; risk perception and communication; security communication recommendations**

I. Introduction

The importance of information and systems security can hardly be disputed. Year on year, numerous security surveys [1–3] have been published that highlight the significant losses incurred by businesses and individuals as they battle old and new attacks, particularly in the online space. As approaches to bolster cybersecurity have evolved, an increasing amount of consideration (e.g., [4–8]) has been devoted to human users and supporting their interactions with systems in relation to information security technologies. This effort has resulted in the Security Usability and Human-Computer Interaction and Security (HCISec/HCI-S) domains, and led to their establishment as crucial areas for research and application. Work within these fields typically studies the usability of mechanisms for authentication, encryption, PKI and device pairing, and generally security tools and secure systems [9].

Another domain which has also been gaining research emphasis is that of security-risk communications [10–13] and making security risks in cyberspace more tangible. The central question here is how best to communicate cybersecurity risks to individuals to facilitate understanding and promote good security judgement. The attention is therefore not solely on security software but also on most contexts where conveying security-risk information to individuals is necessary. It is this topic of cybersecurity-risk communication that forms the focus of our work. Specifically, this review paper aims to provide crucial insight into what aspects may be important in fostering trust and improving the effectiveness of a security-risk message. In earlier work [11], we have hinted towards this aim and identifying design principles to support accurate communication (perception, analysis and treatment) of online security risks. Camp [14] is one of many authors that highlights the significance of this type of research and seriously engaging security communications.

We begin this paper's literature survey in Section II with an assessment of how individuals perceive and make judgements on an information object's trustworthiness. A review of this mature field should provide useful guidance on how one might build trust in risk information displayed to individuals, and thereby increase its chance of being noticed and used in decision-making [15]. Next, we reflect on general risk-communications research and the significant work undertaken within the health and natural-disaster domains (Section III). Risk-perception and cognitive-science principles are important here as we seek to outline research progression in risk communications and later, attempt to adapt established recommendations to a security-risk context. With this foundation, in Section IV we then review the advances within the security-risk communications field specifically. This allows us to present the state-of-the-art in a topical socio-technical security field. Next, Section V draws from the general risk-communication and information-trustworthiness research to outline some potential cybersecurity-risk communication recommendations (to be evaluated in future work). The paper then concludes and discusses directions for further work in Section VI.

II. What Builds Trust in Information

To communicate risk is to communicate information. Central to the goal of trustworthy and effective risk communication therefore is understanding what aspects influence perceived information trustworthiness. In previous research [16] we have sought to address this question through a comprehensive review of the trust literature. That work spanned both offline and online domains and covered numerous key articles includ-

ing [15,17–19], which assess from trust in digital information and Web content to the new area of trust in social media information. The outcome of that survey was a list of over 35 factors which have an effect on whether an individual will decide to trust an information object. In Table I, we draw from that extensive list and categorise the trustworthiness factors based on whether they relate to (i) the information source (where the information originated or who sent it), (ii) the piece of information itself, or (iii) the end-user of the information (message receiver). Albeit simple, this breakdown can actually be quite beneficial in allowing system designers to see at what stages a user's trustworthiness in information might be built, or if design is poorly conceived where it might inadvertently be destroyed. From this factor listing, we note for example that the identity of a source, relevance and timeliness of information, and expertise of an end-user are all likely to influence the trustworthiness of an information message.

TABLE I
PERCEIVED INFORMATION TRUSTWORTHINESS FACTORS

Source:	Deception, Ethics, Identity (Source, Authority/Competence of source, Credentials, Digital signatures), Incentive, Location of source (Geographic location), Objectivity (Bias), Popularity, Positive intentions (Goodwill), Recommendation (Seals of approval, Rankings, Citations), Related resources, Reputation (Direct experience, Predication), Motivation, Similarity to receiver beliefs/context
Information:	Accuracy (Free-of-error, Reliability), Believability (Likelihood, Plausibility of arguments), Competence of information, Consistency/Internal validity, Corroboration (Agreement), Objectivity (Bias), Predictability, Presentation and format (Appearance, Appeals of a personal nature, Representational consistency, Concise representation), Provenance, Recommendation (Seals of approval, Rankings, Citations), Relevance, Specificity, Timeliness/Freshness (Age, Recency, Volatility, Response time, Speed of loading), Topic, Verifiability
End-user:	Bonding, Context and criticality, Beliefs, Disposition to trust, Expertise, Motivation, Propensity, Risk propensity, Trust in general technology, Familiarity, Limited resources/choices

In addition to their impact on information in general, by extension these factors may also have a significant influence on a security-risk message (e.g., a warning prompt) and whether it will be trusted by an individual. As such, a risk message that is accurate, specific, presented appropriately and is familiar, is more likely to be trusted and acted on than a message that is to the contrary. Many of the factors in Table I are already in use today to make decisions about trust on the Web. If we look at Web browsers for example, *identity* is a core factor in choosing a browser and linking names such as Chrome and Internet Explorer to their respective supporting companies i.e., Google Inc. and Microsoft Corp. Having identified the source, it is normal for individuals to then consider additional aspects such as the browser/company's *reputation*, *competence* and *popularity*. In terms of security, this example might also be extended to appreciate the fact that most browsers provide *relevant* and *specific* warning messages about Web sites infected with malware before letting individuals access them, thus emphasising *timeliness* of information as well. These factors, and particularly those related to the end-user and their cognitive state will be increasingly important as system designers seek to create an atmosphere of trust and facilitate effective risk communication leading to more secure and trusted systems.

III. COMMUNICATING RISKS: HURDLES AND PROGRESS

Risk communication is a relatively mature field and has been researched in detail for many years, especially within the health and natural-disaster domains. Risk communication can be defined as the interactive process of exchanging information about a risk (its nature, meaning, consequences, likelihood and response options) to individuals so that they can make informed judgements [20–23]. This activity can be split into three goals, advancing/changing knowledge and attitudes, modifying risk-relevant behaviour, and facilitating cooperative conflict resolution and decision-making [24]. All of these goals require individuals to initially consume risk information or in other words, perceive it. As a result, risk perception forms one of the critical initial stages within risk communication that considers the ways in which a person actually views a risk and the various factors that affect their perspective [23,25]. Some of the other processes core to risk communication which are implicit to the goals mentioned above are risk analysis, risk evaluation and risk treatment [23,24]. Together, these activities allow an individual to mentally understand a risk, weigh it and make an informed cognitive decision concerning how it should be treated.

Considering the broad focus of risk communication and its link with human perception and decision-making, it should be of no surprise that this is often regarded as a complex topic with numerous factors affecting and influencing its processes. Various articles [22, 26–28] support this reality and supply detailed studies that identify the multitude of factors that come into play. A large contributor to risk communication's complexity is the perceptual and subjective nature of a risk itself, with authors [27,29] in both the health and terrorism domains referring to it as a socially constructed and psychologically-oriented phenomenon. Literature has discussed the subjectivity of this topic in detail and listed key examples such as (i) the fact that actual risk and perceived risk (i.e., what a person perceives the risk to be) can be quite different [26,30] and (ii) the appetite and acceptability of a risk depends heavily on an individual's priorities and values [29].

Slovic [31] provides one of the most influential and significant reviews on the subject of risk perception, and draws from various fields to confirm the assertions above and stress how difficult, yet important, researching risk communication and perception is. In light of this complexity, it is crucial that any cybersecurity-risk communication approach has an adequate appreciation of possible difficulties and ways to avoid them, thereby communicating security information more effectively. Considering this, we review core aspects in the established health and natural-disaster fields that have been found to impact risk communication. It is hoped that progress in these domains might facilitate much quicker advancements in cybersecurity-risk communication research. To structure our review, we use the three areas identified by [32] in which hurdles in risk communication may arise, namely, the risk message itself, the message communicator/source and message

receiver (the individual). One should note the similarity here to the categories of information trustworthiness in Table I.

A. The risk message

As highlighted in the literature [32], the risk message itself presents a noteworthy challenge to risk communication. One of the first issues which arises is the innate complexity of the problems which risk information actually relates to [32]. In most situations, the problems of interest are likely to be quite complicated, and therefore the associated risk information is almost certainly not trivial. Another basic yet important consideration is deciding exactly what information to present in a risk message [26, 32]. This task should not be underestimated, because badly chosen information could have numerous adverse consequences and ultimately lead to individuals making ill-advised decisions [26]. The specificity of risk information has also emerged as a possible hurdle to communication. Jenkin [29] reports on this factor in the terrorism context and drawing from work in [33] highlights that communications that are not specific enough may increase anxiety without increasing an individual's actual awareness.

Once a risk message has been researched and the appropriate information selected for communication, the next crucial question is how should it be presented. The question of risk-information format is arguably one of the most heavily researched subtopics pertaining to a risk message. According to the literature, there are three broad formats of presentation: numeric (using percentages, frequencies and probabilities), verbal (which applies terms such as 'unlikely', 'possible' and 'definite') and visual (utilising graphics, graphs, charts and diagrams) [34–40]. Each of these types (and each specific format chosen) has its own unique strengths and weaknesses in facilitating productive risk communication. In some situations, there also may be the opportunity to combine formats.

To take the numeric format type as an example, it has the benefit of being precise, verifiable for accuracy and easily convertible from one metric to another [36]; all useful qualities for a risk message. The overarching weakness with this format, however, is that it assumes an understanding and ability regarding mathematical and probabilistic concepts (generally termed 'numeracy' [41]) that has proven to be misplaced, even amongst highly educated individuals (various studies [42–44] on numeracy in the medical field support this point). At the lower level of specific numeric formats, difficulties are also apparent both in terms of choosing the best format and secondly, ensuring it is properly applied. This is discussed in depth by several authors regarding probabilities, frequencies, relative risks and the importance of reference classes (i.e., stating to whom the risk information relates), and the need to appropriately frame a presented risk. Research in [45], for example, highlights that single event probabilities, conditional probabilities and relative risks tend to be confusing because it is difficult to understand what class of events a percentage or probability refers to. Other research on the numeric format (in [36,40,45]) presents more of these arguments and discussions, and are central articles that link to various additional works.

The verbal communication format is not used as widely as the other two but does have its benefits. Most notably, these include possibly being superior at representing an individual's intuitions and emotions, being natural and easy to use, and lastly, being good at expressing the source, level and imprecision of the uncertainty plaguing a typical risk message [36]. As summarised in [36], the core weakness in this technique is the high degree of variability in individuals' interpretations. This is further supported by [40, 46] from general and clinical consultation-specific perspectives. If one takes the term 'likely' to describe a risk for example, there is no real way to ensure that it will mean the same thing to all individuals. This problem is exacerbated especially in cases where it is the aim of communicators to portray precision in a risk message. Other factors found to influence interpretation are an individual's experience, knowledge and expectations [36]. Briefly comparing this format to the use of numerical information, some authors [40] conclude that numerical risk information is better understood and trusted than verbal information.

Visual mechanisms including graphs, charts and risk ladders have become popular formats for communicating risks as well. Summarising the literature on this topic, many researchers [34, 36] note that the advantage of visuals lies in their ability to attract and hold people's attention, to assist in visualising and portraying part-to-whole relationships, and to capture and summarise large amounts of data, thus allowing for easier identification of patterns. Visuals may also have the benefit of potentially being more apt communicators for individuals with low numeracy levels; in [47] for example, risk ladders are seen as particularly helpful in communication. The effectiveness of specific graphics rely on numerous aspects, including display characteristic (e.g., layout, use of cues and colours), data complexity, user characteristics (e.g., cognitive styles and demands) and the task at hand and cognitive load on individuals [34, 40].

The amount of information within a visual has been researched in [48], looking towards reducing complexity with a 'less is more' approach—of course the potential issue here is not displaying both frames of information and thus risking biasing judgement. Nonetheless, this paper does give a good example of the ongoing work in this area. Possible drawbacks of visual tools centre on poor design, significant complexity, patterns that may draw attention away from important details, not stating the reference class, and obscuring relevant comparisons [36,49]. If not addressed, these weaknesses may lead to misunderstandings and failures in cognitive processing that eventually result in poor risk communication.

B. The risk communicator/source

Returning to the broad areas defined earlier, our focus now shifts to the risk communicator/source and the challenges faced there. The main hurdle in this regard is the reality that communicators themselves have difficulties in processing and calculating risk [32]. Work on quantitative risk communication by [32] highlights various studies in which key information

sources (e.g., doctors, judges and experts) incorrectly or inconsistently calculated rather serious risks. Gigerenzer and Edwards [45] further support this based on their research on numeracy in risk communication in the health field. Influential sources such as the media have also been shown to cause problems through a lack of care in interpreting and reporting risk statistics properly [49].

C. The message receiver

The message receiver (or individual) introduces yet another dimension of hurdles. One of the most significant aspects is the question of how the individual perceives the risk at a personal level. Important influential factors include a person's culture, beliefs, needs, knowledge, awareness, familiarity with the risk, feeling of control, voluntariness, and the level of impact and dread of the risk [21,23,30,31,50–52]. Another key aspect is a person's literacy and (as stated before) numeracy levels; how literate or numerate an individual is will have serious repercussions on the effectiveness of risk communications (a variety of studies [32, 41, 42, 44, 47, 53] over numerous domains support this reality). Emotion (e.g., fear, anger and anxiety), attitude and affect (a good or bad feeling about something) may also influence an individual's perception and decisions regarding risks, as discussed in several articles [27,32,54,55]. The fact is that as humans, we often rely on affective responses, emotions, and even our current mood to motivate how we perceive things and make decisions.

Building on this discussion, a salient point made by [56] is that most risk analysis is handled quickly and automatically by the experiential mode of thinking (which is intuitive, automatic, alert to cues and fast). The affect heuristic is one notion which has shown itself to be important to this field and most importantly, risk communication [55, 56]. Research [28, 57] has also looked into other popular heuristics (e.g., availability, anchoring and adjustment, representativeness) and their positives and negatives when applied to making judgements and decisions regarding risks. The receiver's trust in the risk communicator is a pivotal aspect as well [21, 27, 29, 50]. If an individual does not trust the source, they are not likely to perceive the risk accurately, which may result in an overestimation or underestimation of the risk. In [58], the authors offer slightly opposing views noting that, contrary to popular belief, the influence of trust may actually be limited. Further research is needed to clarify these points. This general link to trust, however, does support the current paper's aim to draw on trustworthiness research for cybersecurity-risk communications. Before moving on, Table II provides a quick summary of the difficulties documented in risk-communication literature which were discussed above; for clarity we do not restate the references in the table.

From this brief review of the hurdles commonly faced in risk communication, it should be no surprise that one of the key proponents in the field (i.e., [27]) has referred to the process of risk assessment (a core part of communication which groups analysis and evaluation [23]) as a 'battlefield'. Combined, these challenges stress the fact that simply supplying accurate risk information is not enough to ensure that individuals will be able to process and comprehend the risk message [32], let alone act on it. There must be different strategies for different purposes and different target groups; one-size-fits-all is not a viable approach to effective risk communication in any domain. Following on from this, there have been recommendations and guidelines for effective risk communication within the literature. These span techniques for presentation, source association and optimising cognition, as well as ways to communicate to particular types of individuals, for example those with low numeracy levels. Some of the most noteworthy articles include [21, 32, 34–37, 44, 45]. We reflect on some of these recommendations as they relate to cybersecurity-risk communications later in this paper.

TABLE II
SUMMARISING HURDLES TO RISK COMMUNICATION

Risk message:	– Risk information and the situation it relates to are likely to be complex and therefore innately difficult to communicate – Deciding exactly what risk information to present is a critical task that ultimately affects decision-making – Vague and unspecific risk information may increase anxiety and not risk awareness – For numeric messages, key concerns relate to dealing with low-numerate individuals, choosing the best presentation format and ensuring that the chosen format is suitably applied – For verbal messages, the main hurdle is the high degree of variability in how individuals interpret messages – For visual messages, concerns relate to poor designs, complex diagrams, patterns drawing one's attention from crucial details, obscuring relevant comparisons and not stating reference classes. Effectiveness of visuals is also influenced by display characteristic, data complexity, user characteristics and the task at hand
Risk source / communicator:	– Understanding, calculating and conveying risk information can be a challenging task even to the message communicator/source
Message receiver:	– There are a variety of factors which influence how an individual perceives a risk and these must all be generally considered. These include personal (e.g., culture, emotion and familiarity with risk), skills-based (e.g., literacy and numeracy) and psychological (e.g., modes of thinking and heuristics) factors – The extent to which an individual trusts a risk communicator/source has a noteworthy effect on the success of risk communications

IV. The Communication of Cybersecurity Risks

To reiterate, risk communication in the cybersecurity context considers how best to communicate security-risk information to users of a system in order to facilitate understanding and promote informed judgement. In some cybersecurity situations, persuading users to adopt a particular course of action may also be the goal. Research in the security communications space is relatively new [14] and at this stage may be broken into work on perception of security risks and decision-making regarding these risks—this somewhat mirrors early work in the risk-communications field covered in Section III. As such, these two areas form the themes of our review below. We must emphasise that this review is on cybersecurity-risk communications and not general security usability research. At times these concepts overlap, but our focus at this time is on the former notion.

A. Perception of security risks

We start with work in [59] where the authors, accepting the significance of the human element of information security,

conduct a survey into factors influencing individuals information security perceptions. To guide their work, they draw on popular and established risk-perception literature. Their study concludes that a person's perception of information security can be defined by six core factors, namely knowledge, impact, severity, controllability, possibility and awareness. This finding links to influential factors of a message receiver identified in Section III-C. Another study [60] supports some of these factors as it researches what dimensions influence an individual's risk perception of online hazards (security risks/threats). From that review and analysis, the authors find that persons use four main dimensions in judging online risks, namely ability to control or avoid the risk, dread of consequences, unfamiliarity of risks and immediacy of consequences/impact. The authors note that researchers may be able, through an understanding of these aspects, to predict individual's reactions to online risks.

There has also been research on understanding and measuring security-risk perception. In [61] and later in [12], for example, the authors define a novel risk-perception measurement model. The model distils common perception factors and grounds itself in two security-risk characteristics—an individual's knowledge about a risk and the risk's consequences—with each characteristic having a scale of levels indicating different values and measures. For instance, for a given solution at a particular time, a person may have a low understanding (knowledge) of a risk (valued at 'Level 2: Understanding') and may view the possible consequences as quite serious (preferring, 'Level 2: Serious, ongoing and raises ethical concerns' on the scale). Using other defined parameters, total scores may be tallied and then either combined into group scores to define a group's perception of a risk, used at later time intervals to track changes in an individual's risk perceptions, or both.

In addition to investigating and measuring factors that affect security-risk perception, some authors have sought directly to influence risk perception in order to improve risk communication. Research in [10] exemplifies this, as the authors attempt to improve the process by embedding graphics and symbols in information security messages. Use of this technique was transferred from other fields (e.g., education) where it has proved effective for information presentation. Contrary to expectations however, the outcome of the authors' study did not identify any statistically significant differences between graphical and text-only test groups. This is particularly interesting for our context as it may suggest that not all aspects are easily transferable across fields, or simply, that as always, care must be taken in researching and presenting information. The authors also mention several important questions regarding how security-risk information is to be presented, querying wording, format, means, and even colour. These are all aspects which have links to literature in general risk-perception research and therefore once again show why recommendations in that field may be of great use to cybersecurity-risk communications.

B. Decision-making on security risks

In terms of decision-making as it relates to security risks, [62] provides a useful study into factors that impact security-risk decisions. Specifically, the authors examine how users/individuals make tradeoffs regarding security risks and rewards. From their empirical study, they conclude that an individual's risk perception, security skill and culture do influence decision-making on risks. Although from a general risk-communication perspective these findings are well understood (with most mentioned in Section III-C), it is encouraging to see them being applied and tested in the security field.

West et al. [63] focus on the direct question of why users make poor decisions and, through numerous case-study analyses, identify several human factors. These include humans' tendency to satisfice (choose quick and 'good enough' alternatives and not necessarily the best ones), to succumb to cognitive biases (e.g., representativeness heuristic and base rate and response bias), to be faced with time pressures, and to suffer from inattentional blindness, amongst other things. Salient points in that and other related work [64] are that users generally do not think that they are at risk, users tend to be unmotivated, safety is often considered an abstract concept and lastly, losses are usually perceived disproportionately to gains. The authors do offer some possible solutions and these range from improving security-risk awareness, to modifying risk messages/dialogues to attract attention, or if possible, removing users from the security decision completely (e.g. system scanning for viruses on an input pen drive without confirming this with a user). This work is very useful as it stresses the importance and application of various established risk-perception and communication tenets.

Another novel research article in security-risk decision making is found in [65]. Here, the authors examine how individuals evaluate online risks without all the necessary information; a reality they view as common in the online context. This gives rise to four levels of 'knowability' of risks, namely, known certainty, known uncertainty, unknown uncertainty and unknowable uncertainty. Their example of known certainty is when a supplier guarantees that because of its strong security mechanisms, none of its online transactions leads to identify theft. Whereas, unknowable uncertainty (the opposite end of the spectrum) is where no one knows and there is no way to determine exactly what amount/percentage of online transactions with the supplier leads to identity theft. Based on their study, the authors then show that these levels have varying effects on a person's decisions. This is even to the extent that expressing risks in a particular way (or level) may lead to particular choices.

Mental models have also been introduced to assist in security-risk communications. These models define internalised representations of external reality [13, 66]. Camp [14] proposes the application of these models, hoping to draw from their successes in improving risk communication in other domains. Five possible models were discussed in that work: physical security, medical infections, criminal behaviour, economics failure and warfare. Each of these models has different uses and benefits in application to security communication problems. As with any other method, however, the author stresses that these models are not perfect but may have

their unique uses in effective communication. Other work has sought to utilise mental models further in targeted security-risk communication. In [13], the authors combine mental models (embodied in videos) and activity recognition tools to display timely warnings linked to video stills. They note that video may lead to better comprehension than text. User tests will be necessary to determine how useful this innovative approach is. At a first glance, however, we identify that there might be practical organisational issues as it appears that users/individuals will need to watch a video prior to system use and secondly, activity recognition and the necessary logic need to be built into or on top of software or system schemas.

Similar to discussions in Section III, a core part of security-risk communications is the risk message itself. Two of the most recent and pertinent works on this topic are [67, 68]. In [67], the authors identify a key set of criteria for the design of security alerts. These include: creation of interface designs that match users' mental models, focus on aesthetic and simple design, establishing standard colours to capture users' attention, using icons as visual indicators, explicit words to classify risk levels, and consistent, meaningful terminology. Bravo-Lillo et al. [68] seek to advance the cybersecurity warnings field through the use of mental models. The main contributions of that research include, insight into how advanced and novice users make sense of warnings, and general notes about warning design and presentation (e.g., amount of content to display and considering all steps of how users process warnings).

The research reviewed above is aimed mainly at understanding from a non-security specialist perspective. There is also a body of work which seeks to improve communications between systems and security administrators of an organisation on the current risks from a cyberattack. From the academic perspective, Jaferian and colleagues [69, 70] have researched this topic in depth and defined design guidelines for IT security-management tools and heuristics for their evaluation. In terms of security communications in particular, the authors posit that designers should use a range of different presentation/interaction methods to display information, meaningful messages should be used, and interfaces and alerts should be appropriate for and customisable by users [69]. Other work targeted towards security managers/administrators such as [71,72] exists, but proposed guidelines do not readily extend to the security-risk communication aspects of interest.

Within industry, there have been developments in communicating cybersecurity risks as well. The domain of intrusion detection is particularly relevant to our research as it depends heavily upon technology-based tools to help communicate the potential risk of cyberattack. There are a wide range of graphical and textual-based methods available (e.g., [73–75]), however, most of these appear to be based upon a pragmatic application of well-known human cognition and perception principals as opposed to a deep consideration of how to optimise for communication of cyber risk for individuals whose job it is to monitor activities on information systems. One of the more profound exceptions to this observation can be seen in [76]. In that article, the authors investigate the use of novel visualisation and recommendation techniques to improve the performance of cybersecurity analysts working in real-time incident-management environments. They test their approach with user studies (involving professional cybersecurity analysts) and report a number of positive findings, one of which was increased accuracy in incident classification. In terms of our research, this article is particularly useful as it looks at key questions such as visual versus tabular displays and different information presentation sequences within cybersecurity notification systems. Undoubtedly we will reflect further on this article in the analysis of our recommendations in future work. Next, we consider the recommendations proposed for improving communication of cybersecurity risks.

V. Recommendations for Improving Communication of Cybersecurity Risks

From the review of security-risk communications research above, it is clear that there is an increasing amount of effort being placed on this new domain. Our aim is to complement these efforts and further the field by investigating whether principles and advances from associated domains might supply useful guidance to this new research area. In line with this aim, below we draw from the numerous guidelines and 'lessons learned' in the very related, general risk-communications field (original references are highlighted inline), to propose a list of recommendations that may pertain to cybersecurity-risk communications. Where applicable, we also incorporate information-trustworthiness factors into these recommendations to foster trust in security-risk information and thereby increase likelihood of subsequent usage.

- Planning how cybersecurity risks will be communicated is crucial. System designers should be clear on: (i) the goal of the communication (e.g., is it to educate users or draw them away from a security decision that may be too risky); (ii) what type of security messages and communication strategies would be most useful (in [13] for example, the authors emphasise strategies reliant on visuals and mental models); and lastly (iii) the characteristics (e.g., level of knowledge and education, literacy and numeracy, mental models, attitudes/beliefs about the security issue) of individuals targeted by risk messages (e.g., knowledgeable Web users might desire more specifics than novice users regarding a security risk posed by a potentially malicious Web site). It is also important to explain possibly unfamiliar terms or complex security aspects—if users are not able to properly understand a risk, it is unlikely they will appropriately treat it. References [51, 77] generally support the points above. We note that current tools arguably do not allow for much personalisation and thus, generally operate on a one-size-fits-all basis.
- It is well-understood that humans possess a limited processing capacity. As such, designers should focus on reducing the cognitive effort required by individuals in processing security-risk information and/or security-related interfaces [35–37]. This may be done by cutting

back on the initial amount of security details, and as much as is possible, keeping communications simple [51]. This suggestion will need to be tempered by the current context, as certain users (e.g., experts or security analysts/administrators whose job it is to monitor all levels of system security) may prefer to be presented with detailed information initially. The presentation and format, relevance and specificity of information also become key factors in increasing a user's trust in a security-risk message displayed [16]. Methods that appreciate all of this recommendation's aspects may be deployed in practice but we can find only one somewhat related study (i.e., [76]) on general performance and effectiveness.

- System designers should ensure that the meaning of information presented in security-risk messages is clear. Methods to achieve this include using appropriate message framing (including assessing whether positive (e.g., there is a 96% chance a Web site is legitimate) framing is more suitable than negative (e.g., there is a 4% chance the site is malicious) framing) and providing narratives of possible outcomes of making the decision to use or not use presented security-risk information (e.g., if the flagged file is indeed malware, installing it may result in disruption of normal system services and use, invasion of privacy, and so on). [11, 35]
- Users should be presented with clear and consistent directions for action, i.e., options to respond to a security risk faced [39, 53]. Also, designers should assist users in visualising what the actual experience (result) of a security-risk decision may be like. This is particularly pertinent in situations where users may be faced with unfamiliar choices. Narratives (descriptions with a resulting outcome, such as increased potential for system to be compromised) may also be helpful here. [35]
- When communicating cybersecurity risks numerically, note: (i) users with high-numeracy levels are likely to pay more attention to risk figures, while low-numerate users may draw more on emotions, mood states and expert guidance [41]; (ii) it may be beneficial to present security-risk numbers using frequencies (e.g., there is a 1 in 100 chance that code in a Web site has malicious intentions) instead of (or in addition to) percentages (e.g., there is a 1% chance that code in a Web site has malicious intentions) [32,44,45]; (iii) to avoid individuals dismissing small risks (e.g., 1% or less) or risks from familiar events (e.g., security information or warning messages from a particular source), an explicit message to this effect is in order [36]; (iv) some relative risk communication often results in an overestimation of perceived risk. This format is useful if aiming to influence users towards a specific (e.g., less risky) security decision but should be avoided otherwise (e.g., use absolute risk). [36, 44, 45]
- When communicating cybersecurity risks visually, note: (i) no single visual will work perfectly in all situations–bar charts, risk ladders, pie charts, icons, indicators, etc. all have slightly more useful application contexts [34,40], even security-specific research by [10] has alluded to this in terms of graphics and symbols; (ii) to promote educated judgements, displays should be representative of actual quantities/probabilities [36], this is particularly relevant if showing security-risk levels or virus infection statistics graphically; (iii) visuals showing security risks should avoid elements (e.g., extraneous pictures or distracting images) that divert attention away from the data [34]; (iv) if graphs are used (e.g., to show attack likelihood), these and any conclusions that might be drawn from the visuals should be explained clearly and not left up to an individual's sole interpretation [36].
- When communicating cybersecurity risks verbally: (i) it may be best to use multiple formats to present security-risk information as various authors have expressed that verbals are not to be completely relied on [40,44,78]. This is especially relevant for security as it is common to see messages quoting that attacks are 'likely' or 'probable'. The core issue therefore is, how does one ensure that these terms mean the same to all users; (ii) consider context as this may also influence perceptions [40]—context might span who the users are, where they are, what they are likely to be doing in the system and the gravity of the security decision they currently face.
- Some of the most important criteria for evaluating security-risk messages in pilot testing include: comprehension, agreement, hazard/dose-response consistency, uniformity, audience evaluation and types of failure in communication [79]. Thus, have users understood the security-risk message as expected? Is there uniformity in risk levels and responses? Do users view security communications as clear, helpful and accurate? Etc.
- In seeking to build trust in cybersecurity-risk communications, it is beneficial for messages to be given in a timely fashion (ideally as close to the risk situation/attack as possible), to be presented in a standard and predictable security message format which is generally familiar to users, clearly to highlight the reason for the communications and allow it to be verified and traced (e.g., to an information source), to identify the origin/system that generated the risk message, amongst other things. [16]

Grounded in existing risk-communications and information-trustworthiness research, these recommendations may be of great use in communicating cybersecurity-risk information. Existing work already cited has a few of these guidelines in terms of risk communication and generally usable security. For example, keeping communications simple and minimalistic [67], assisting users in seeing the consequences of decisions [7], and engaging in some level of customisation of security-risk information to specific target audiences [69]. Our future work therefore will seek to determine how useful these proposed recommendations may actually be.

VI. Conclusion and Future Work

In this paper, we critically reviewed the information trustworthiness, risk communication and cybersecurity-risk com-

munication research fields. From that survey, we identified several motivational factors and drew on recommendations that may be applicable to the largely uncharted field of cybersecurity-risk communications. An underlying goal was to determine the degree to which these aspects are already addressed in current security communications work and, going forward, how we might seek to increase the overall effectiveness of these security communications.

Having defined a number of these trustworthiness factors and communication recommendations, our future work will seek to thoroughly investigate their combined use and the extent to which they may prove beneficial in the cybersecurity context. We envisage that this investigation will involve several progressive steps. These include, the identification of a set of case scenarios where various facets of cybersecurity-risk communication could be tested, the development of a prototype system and/or add-on functionality in line with scenarios to provide a practical basis for assessment, and finally, user studies with suitable subjects to critically evaluate the effectiveness and trustworthiness of cybersecurity-risk communications with and without incorporation of the recommendations proposed. The online environment is a key target area for these application tests, as that environment may have unique characteristics to explore and/or exploit. Although some very relevant work has been done in these areas (most notably [67,68]), there are still many unanswered questions. Topics such as numeric and verbal communication of cybersecurity risks and personalisation for perceptual and individual factors are especially of interest as these have not been addressed in great detail as far as it relates to this research. Visual risk-communication has received some emphasis, particularly in using icons, indicators and graphics in browsers, firewalls and incident-management tools [67,76,80]. Undoubtedly the field of security usability will be drawn on in further work to enhance/supplement our recommendations.

ACKNOWLEDGMENT

This work was conducted as a part of the TEASE project, a collaboration between the University of Warwick, HW Communications Ltd and Thales UK Research and Technology. The project is supported by the UK Technology Strategy Board's Trusted Services Competition (www.innovateuk.org) and the Research Councils UK Digital Economy Programme (www.rcuk.ac.uk/digitaleconomy).

REFERENCES

[1] Computer Security Institute (CSI), "2008 CSI computer crime and security survey," 2008.

[2] PricewaterhouseCoopers LLP and Infosecurity Europe, "Information Security Breaches Survey 2010: Executive Summary," 2010.

[3] Detica and Office of Cyber Security and Information Assurance in the Cabinet Office UK, "The cost of cyber crime," 2011.

[4] M. Sasse, S. Brostoff, and D. Weirich, "Transforming the 'weakest link'–a human/computer interaction approach to usable and effective security," *BT Technology Journal*, vol. 19, no. 3, pp. 122–131, 2001.

[5] S. Smith, "Humans in the loop: Human-computer interaction and security," *IEEE Security & Privacy*, vol. 1, no. 3, pp. 75–79, 2003.

[6] D. Balfanz, G. Durfee, D. Smetters, and R. Grinter, "In search of usable security: Five lessons from the field," *Security & Privacy, IEEE*, vol. 2, no. 5, pp. 19–24, 2004.

[7] B. Payne and W. Edwards, "A brief introduction to usable security," *IEEE Internet Computing*, pp. 13–21, 2008.

[8] S. Parkin, A. van Moorsel, P. Inglesant, and M. Sasse, "A stealth approach to usable security: helping it security managers to identify workable security solutions," in *The New Security Paradigms Workshop (NSPW) 2010*. ACM, 2010, pp. 33–50.

[9] R. Kainda, I. Flechais, and A. Roscoe, "Security and usability: Analysis and evaluation," in *International Conference on Availability, Reliability and Security (ARES)*. IEEE, 2010, pp. 275–282.

[10] M. Pattinson and G. Anderson, "How well are information risks being communicated to your computer end-users?" *Information Management & Computer Security*, vol. 15, no. 5, pp. 362–371, 2007.

[11] S. Creese and K. Lamberts, "Can cognitive science help us make online risk more tangible?" *IEEE Intelligent Systems*, vol. 24, no. 6, pp. 32–36, 2009.

[12] F. Farahmand, M. Dark, S. Liles, and B. Sorge, "Risk perceptions of information security: A measurement study," in *International Conference on Computational Science and Engineering*. IEEE Computer Society, 2009, pp. 462–469.

[13] J. Blythe, J. Camp, and V. Garg, "Targeted risk communication for computer security," in *15th International Conference on Intelligent User Interfaces*, 2011, pp. 295–298.

[14] L. J. Camp, "Mental models of privacy and security," *IEEE Technology and Society Magazine*, vol. 28, no. 3, pp. 37–46, 2009.

[15] K. Kelton, K. R. Fleischmann, and W. A. Wallace, "Trust in digital information," *Journal of the American Society for Information Science and Technology*, vol. 59, no. 3, pp. 363–374, 2008.

[16] J. R. C. Nurse, S. S. Rahman, S. Creese, M. Goldsmith, and K. Lamberts, "Information quality and trustworthiness: A topical state-of-the-art review," in *International Conference on Computer Applications and Network Security (ICCANS)*. IEEE Press, 2011, pp. 492–500.

[17] Y. Gil and D. Artz, "Towards content trust of web resources," *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, no. 4, pp. 227–239, 2007.

[18] A. J. Pickard, P. Gannon-Leary, and L. Coventry, "Trust in 'E': Users' trust in information resources in the web environment," in *ENTERprise Information Systems*, ser. Communications in Computer and Information Science, J. E. Quintela Varajo, Ed., 2010, vol. 110, pp. 305–314.

[19] S. Moturu and H. Liu, "Quantifying the trustworthiness of social media content," *Distributed and Parallel Databases*, pp. 1–22, 2010.

[20] National Research Council (NRC) USA, *Improving Risk Communication*. National Academy of Sciences, 1989. [Online]. Available: http://www.nap.edu/openbook.php?record_id=1189

[21] K. Calman, "The language of risk: a question of trust," *Transfusion*, vol. 41, no. 11, pp. 1326–1328, 2001.

[22] ——, "The William Pickles Lecture. Issues of risk:'this unique opportunity'," *The British Journal of General Practice*, vol. 51, no. 462, pp. 47–51, 2001.

[23] International Organization for Standardization (ISO), "ISO/IEC guide 73 risk management – vocabulary – guidelines for use in standards," Tech. Rep., 2002.

[24] B. Rohrmann, "The evaluation of risk communication effectiveness," *Acta psychologica*, vol. 81, no. 2, pp. 169–192, 1992.

[25] ——, "Risk perception, risk attitude, risk communication, risk management: a conceptual appraisal," in *15th Internaional Emergency Management Society (TIEMS) Annual Conference*, 2008.

[26] B. Fischhoff, A. Bostrom, and M. Quadrel, "Risk perception and communication," *Annual Review of Public Health*, vol. 14, no. 1, pp. 183–203, 1993.

[27] P. Slovic, "Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield," *Risk analysis*, vol. 19, no. 4, pp. 689–701, 1999.

[28] E. Peters, K. McCaul, M. Stefanek, and W. Nelson, "A heuristics approach to understanding cancer risk perception: contributions from judgment and decision-making research," *Annals of Behavioral Medicine*, vol. 31, no. 1, pp. 45–52, 2006.

[29] C. Jenkin, "Risk perception and terrorism: Applying the psychometric paradigm," *Homeland Security Affairs*, vol. 2, no. 2, pp. 1–14, 2006.

[30] Y. Asnar and N. Zannone, "Perceived risk assessment," in *4th ACM workshop on Quality of protection*, 2008, pp. 59–64.

[31] P. Slovic, "Perception of risk," *Science*, vol. 236, no. 4799, pp. 280–285, 1987.

[32] C. Skubisz, T. Reimer, and U. Hoffrage, "Communicating quantitative risk information," in *Communication yearbook : Vol. 33*, C. S. Beck, Ed. Routledge, 2009, pp. 177–211.

[33] C. Stout, "Using Psychology to Counter Terrorism at the Personal and Community Level," in *Psychology of terrorism: Coping with the continuing threat*, C. Stout, Ed. Praeger Publishers/Greenwood Publishing Group, 2004, pp. 1–31.

[34] I. Lipkus and J. Hollands, "The visual communication of risk," *J. Natl Cancer Inst Monographs*, vol. 1999, no. 25, pp. 149–163, 1999.

[35] J. Hibbard and E. Peters, "Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice," *Annual Review of Public Health*, vol. 24, no. 1, pp. 413–433, 2003.

[36] I. Lipkus, "Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations," *Medical Decision Making*, vol. 27, no. 5, pp. 696–713, 2007.

[37] E. Peters, N. Dieckmann, A. Dixon, J. Hibbard, and C. Mertz, "Less is more in presenting quality information to consumers," *Medical Care Research and Review*, vol. 64, no. 2, pp. 169–190, 2007.

[38] D. Timmermans, C. Ockhuysen-Vermey, and L. Henneman, "Presenting health risk information in different formats: the effect on participants' cognitive and emotional evaluation and decisions," *Patient Education and Counseling*, vol. 73, no. 3, pp. 443–447, 2008.

[39] E. Waters, H. Sullivan, W. Nelson, and B. Hesse, "What is my cancer risk? how internet-based cancer risk assessment tools communicate individualized risk estimates to the public: content analysis," *Journal of Medical Internet Research*, vol. 11, no. 3, 2009.

[40] V. Visschers, R. Meertens, W. Passchier, and N. De Vries, "Probability information in risk communication: a review of the research literature," *Risk Analysis*, vol. 29, no. 2, pp. 267–287, 2009.

[41] E. Peters, "Numeracy and the perception and communication of risk," *Annals of the NY Academy of Sciences*, vol. 1128, no. 1, pp. 1–7, 2008.

[42] I. Lipkus, G. Samsa, and B. Rimer, "General performance on a numeracy scale among highly educated samples," *Medical Decision Making*, vol. 21, no. 1, pp. 37–44, 2001.

[43] S. Woloshin, L. Schwartz, M. Moncur, S. Gabriel, and A. Tosteson, "Assessing values for health: numeracy matters," *Medical Decision Making*, vol. 21, no. 5, pp. 380–388, 2001.

[44] A. Fagerlin, P. Ubel, D. Smith, and B. Zikmund-Fisher, "Making numbers matter: present and future research in risk communication," *American Journal of Health Behavior*, vol. 31, no. 1, pp. 47–56, 2007.

[45] G. Gigerenzer and A. Edwards, "Simple tools for understanding risks: from innumeracy to insight," *BMJ: British Medical Journal*, vol. 327, no. 7417, pp. 741–744, 2003.

[46] R. Thomson, A. Edwards, and J. Grey, "Risk communication in the clinical consultation," *Clinical Medicine, Journal of the Royal College of Physicians*, vol. 5, no. 5, pp. 465–469, 2005.

[47] C. Keller, M. Siegrist, and V. Visschers, "Effect of Risk Ladder Format on Risk Perception in High-and Low-Numerate Individuals," *Risk analysis*, vol. 29, no. 9, pp. 1255–1264, 2009.

[48] B. Zikmund-Fisher, A. Fagerlin, and P. Ubel, "A Demonstration of Less Can Be Morein Risk Graphics," *Medical Decision Making*, vol. 30, no. 6, pp. 661–671, 2010.

[49] E. Kurz-Milcke, G. Gigerenzer, and L. Martignon, "Transparency in risk communication," *Annals of the New York Academy of Sciences*, vol. 1128, no. 1, pp. 18–28, 2008.

[50] A. Alaszewski and T. Horlick-Jones, "How can doctors communicate information about risk more effectively?" *British Medical Journal*, vol. 327, no. 7417, pp. 728–731, 2003.

[51] P. Wiedemann, H. Schutz, and M. Clauberg, "Lessons learned: Avoiding pitfalls in risk communication," in *International Conference and COST 281 Workshop on Emerging EMF Technologies, Potential Sensitive Groups and Health*, 2006.

[52] I. Lin and D. D. Petersen, "Risk Communication in Action: The tools of message mapping," U.S. Environmental Protection Agency (EPA), Tech. Rep. EPA/625/R-06/012, 2007.

[53] R. Rudd, J. Comings, and J. Hyde, "Leave no one behind: improving health and risk communication through attention to literacy," *Journal of health communication*, vol. 8, pp. 104–115, 2003.

[54] L. Sjöberg, "Factors in risk perception," *Risk analysis*, vol. 20, no. 1, pp. 1–12, 2000.

[55] C. Keller, M. Siegrist, and H. Gutscher, "The role of the affect and availability heuristics in risk communication," *Risk Analysis*, vol. 26, no. 3, pp. 631–639, 2006.

[56] P. Slovic, M. Finucane, E. Peters, and D. MacGregor, "Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality," *Risk analysis*, vol. 24, no. 2, pp. 311–322, 2004.

[57] J. Jackson, N. Allum, and G. Gaskell, "Perceptions of risk in cyberspace," in *Trust and Crime in Information Societies*. Edward Elgar, 2005, pp. 245–281.

[58] T. Earle and M. Siegrist, "Trust, Confidence and Cooperation model: a framework for understanding the relation between trust and Risk Perception," *International Journal of Global Environmental Issues*, vol. 8, no. 1, pp. 17–29, 2008.

[59] D.-L. Huang, P.-L. Rau, and G. Salvendy, "A survey of factors influencing peoples perception of information security," in *Human-Computer Interaction. HCI Applications and Services*, ser. Lecture Notes in Computer Science, J. Jacko, Ed. Springer, 2007, vol. 4553, pp. 906–915.

[60] I. Gabriel and E. Nyshadham, "A cognitive map of people's online risk perceptions and attitudes: An empirical study," in *41st Hawaii International Conference on System Sciences*, 2008, pp. 274–283.

[61] F. Farahmand, M. Atallah, and B. Konsynski, "Incentives and perceptions of information security risks," *29th International Conference on Information Systems (ICIS)*, pp. 25–41, 2008.

[62] L. Chen and D. Farkas, "An investigation of decision-making and the tradeoffs involving computer security risk," *15th Americas Conference on Information Systems (AMCIS)*, pp. 610–618, 2009.

[63] R. West, C. Mayhorn, J. Hardee, and J. Mendel, "The weakest link: A psychological perspective on why users make poor security decisions," in *Social and Human Elements of Information Security: Emerging Trends and Countermeasures*. IGI Global, 2009, pp. 43–60.

[64] R. West, "The psychology of security," *Communications of the ACM*, vol. 51, no. 4, pp. 34–40, 2008.

[65] P. Wang and E. Nyshadham, "Knowledge of online security risks and consumer decision making: An experimental study," in *44th Hawaii International Conference on Systems Sciences (HICSS)*. IEEE Computer Society, 2011, pp. 1–10.

[66] P. Johnson-Laird, *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard Univ Press, 1983.

[67] T. Ibrahim, S. Furnell, M. Papadaki, and N. Clarke, "Assessing the usability of end-user security software," in *Trust, Privacy and Security in Digital Business*, ser. Lecture Notes in Computer Science, S. Katsikas, J. Lopez, and M. Soriano, Eds. Springer, 2010, vol. 6264, pp. 177–189.

[68] C. Bravo-Lillo, L. Cranor, J. Downs, and S. Komanduri, "Bridging the gap in computer security warnings: A mental model approach," *IEEE Security & Privacy*, vol. 9, no. 2, pp. 18–26, 2011.

[69] P. Jaferian, D. Botta, F. Raja, K. Hawkey, and K. Beznosov, "Guidelines for designing IT security management tools," in *2nd ACM Symposium on Computer Human interaction For Management of information Technology*, 2008, pp. 1–10.

[70] P. Jaferian, K. Hawkey, A. Sotirakopoulos, and K. Beznosov, "Heuristics for evaluating it security management tools," in *2011 Annual Conference Extended Abstracts on Human factors in Computing Systems*. ACM, 2011, pp. 1633–1638.

[71] S. Chiasson, R. Biddle, and A. Somayaji, "Even experts deserve usable security: Design guidelines for security management systems," in *Symposium on Usable Security and Privacy Workshop at Usable IT Security Management (USM)*, 2007, pp. 1–4.

[72] E. Kandogan and E. Haber, "Security administration tools and practices," in *Security and Usability: Designing Secure Systems that People Can Use*. O'Reilly, 2005, pp. 357–376.

[73] i2 Limited, "i2 Clarity Platform and Analysis Product Line." [Online]. Available: http://www.i2group.com/uk/products–services

[74] Future Point Systems, "Starlight Visual Information System™(VIS)." [Online]. Available: http://www.futurepointsystems.com/

[75] Lookingglass Cyber Solutions, "ScoutVision™." [Online]. Available: http://www.lgscout.com/products/scoutvision

[76] J. Rasmussen, K. Ehrlich, S. Ross, S. Kirk, D. Gruen, and J. Patterson, "Nimble cybersecurity incident management through visualization and defensible recommendations," in *7th International Symposium on Visualization for Cyber Security (VizSec)*. ACM, 2010, pp. 102–113.

[77] V. Bier, "On the state of the art: risk communication to the public," *Reliability Engineering & System Safety*, vol. 71, no. 2, pp. 139–150, 2001.

[78] K. O'Doherty and G. Suthers, "Risky communication: pitfalls in counseling about risk, and how to avoid them," *Journal of Genetic Counseling*, vol. 16, no. 4, pp. 409–417, 2007.

[79] N. D. Weinstein and P. M. Sandman, "Some Criteria for Evaluating Risk Messages," *Risk Analysis*, vol. 13, no. 1, pp. 103–114, 1993.

[80] P. Shi, H. Xu, and X. Zhang, "Informing security indicator design in web browsers," in *2011 iConference*. ACM, 2011, pp. 569–575.