This is a peer-reviewed, accepted author manuscript of the following research article: Schaik, P. V., Renaud, K., Wilson, C., Jansen, J., & Onibokun, J. (2020). Risk as affect: the affect heuristic in cybersecurity. *Computers and Security*, *90*, [101651]. https://doi.org/10.1016/j.cose.2019.101651

Risk as affect: the affect heuristic in cybersecurity

Table of Contents

This is a peer-reviewed, accepted author manuscript of the following research article: Schaik, P. V., Renaud, K., Wilson, C., Jansen, J., & Onibokun, J. (2020). Risk as affect: the affect heuristic in cybersecurity. *Computers and Security*, *90*, [101651]. https://doi.org/10.1016/j.cose.2019.101651

Abstract

Risk perception is an important driver of netizens' (Internet users') cybersecurity behaviours, with a number of factors influencing its formation. It has been argued that the affect heuristic can be a source of variation in generic risk perception. However, a major shortcoming of the supporting research evidence for this assertion is that the central construct, affect, has not been measured or analysed. Moreover, its influence in the cybersecurity domain has not yet been tested. The contribution of the research reported in this paper is thus, firstly, to test the affect heuristic while measuring its three constructs: affect, perceived risk and perceived benefit and, secondly, to test its impact in the cybersecurity domain. By means of two carefully designed studies (N = 63 and N = 233), we provide evidence for the influence of the affect heuristic on risk perception in the cybersecurity domain. We conclude by identifying directions for future research into the role of affect and its impact on cybersecurity risk perception.

Keywords: affect, risk perception, cybersecurity, benefit perception, affect heuristic

Highlights

- Previous work has proposed that affect influences risk perception
- We directly tested the affect heuristic in cybersecurity
- We found evidence for the first affect heuristic model
- Evidence for the second model was variable
- Only affect valence (but not arousal) had an impact on risk perception

This is a peer-reviewed, accepted author manuscript of the following research article: Schaik, P. V., Renaud, K., Wilson, C., Jansen, J., & Onibokun, J. (2020). Risk as affect: the affect heuristic in cybersecurity. *Computers and Security*, *90*, [101651]. https://doi.org/10.1016/j.cose.2019.101651

1. Introduction

The huge advantages global citizens gain from being online are somewhat clouded by the significant risks they are exposed to while accessing online services (de Bruijn and Janssen, 2017). The World Economic Forum ranked cyber-attacks third in worldwide threats in 2018. In the UK, the Office for National Statistics[1] reported that, during the year ending in March 2018, around 4.5 million cybercrimes were committed in England and Wales. In 2017, around 17 million UK residents fell victim to cybercrime, with losses of approximately £130 billion (Hern, 2018). Two in five UK businesses, too, were subjected to a cyber-attack in 2018 (HM Government, 2018a). These statistics are for a single country, but serve to demonstrate the scale of the problem.

Being aware of this, the UK government (HM Government, 2018b) considers increasing cybersecurity skills to be one of their top priorities. This priority is reflected in the other Five Eyes governments' strategy documents too (Public Safety Canada, 2018; New Zealand Government, 2016; US Government, 2018; Australian Government, 2016). The government strategy documents refer specifically to the need for individuals to *understand* the risks and to *know how* to protect themselves online.

It is indeed important for netizens (Internet citizens) to know the facts about online risks, because this has an influence on their ability to develop an informed perception of a particular risk (Hansson, 2010; Bodemer & Gaissmaier, 2015). In particular, people need to have an accurate perception of the risk. Risk perceptions predict uptake of precautionary cybersecurity behaviours (van Schaik et al., 2018), so efforts to inform citizens are indeed advisable. Yet facts and knowledge, on their own, do not reliably lead to accurate risk perceptions (Pidgeon et al., 1992; Cross, 1998). Hansson (2010) explains that risk perception is both objective (fact-based) and subjective (socially constructed and emotional). It is important to understand the influences that inform risk perceptions because people will only take precautions if they have a genuine perception of the risks related to online activities (Slovic et al., 1980).

The uptake of precautionary behaviours can be modelled by protection motivation theory (PMT) (Rogers, 1975). PMT suggests that both threat and coping appraisals will influence people's motivations to take precautionary measures against threats. According to the theory, the riskier a behaviour is perceived to be, the more likely it is that precautionary measures will be taken. The activity might also be avoided altogether if the risk is perceived to be too high (Lienard, 2011). PMT can help us to clarify individual variations in protective cybersecurity behaviours too, because we know that action-related decisions build on individual risk perceptions (Sjöberg et al., 2004; Warkentin et al., 2012; Jansen & van Schaik, 2017a, 2017b, 2018a, 2018b, 2018c).

It is clearly important to understand netizens' formation of cybersecurity risk perception, because this plays such a major role in prompting protective action.

[1] https://www.ons.gov.uk/search?q=cyber

Accurate and objective risk perceptions will prompt deliberate precautionary actions by individuals (Renaud & Warkentin, 2017), organisations (Allodi & Massacci, 2017) and governments (Renaud et al., 2018).

To engender risk perception accuracy, it is necessary to understand exactly how risk perceptions are formed and what factors influence them (Slovic et al., 1980; Hansson, 2010). Kühberger and Schulte-Mecklenbeck (2017) explain that risk perception is not purely a cognitive process, but that it is also informed by affective influences. Finucane et al. (2003) refer to this as the "dance of affect and reason". In fact, Slovic et al. (2002) argue that affect is essential to rational action.

Affect is a potentially powerful yet poorly investigated influence on risk perception (Slovic & Peters, 2006). Core affect is the central concept in Russell's (2003) influential dimensional approach to studying affect and emotion (cited 2261/4560 times, according to Scopus/Scholar, 9/9/2019). The author explains affect as a simple feeling, *core affect*, to which at all times people have conscious access to; this is a mix of two dimensions: valence (pleasure-displeasure) and arousal (activation-deactivation). Here, we use this approach to investigate the role of affect on cybersecurity-related risk perception.

Although the Finucane et al.'s (2000) publication on the affect heuristic has been frequently cited (1313/2737 times, according to Scopus/Scholar, 9/9/2019) and has been invoked to explain risk perception and human behaviour, it remains a conjecture in the domain cybersecurity. Therefore, the aim of the research reported here was directly to test the affect heuristic and its applicability to cybersecurity, with implications for our understanding of the role of affect in risk perception to inform cybersecurity practice. We conducted two studies using Finucane et al.'s (2000) methods. In our first study, we tested Finucane et al.'s (2000) first affect heuristic model; in our second study, we tested the second.

Cybersecurity is a particularly pertinent issue in the UK, because 82% of the UK population shops online (CBS, 2018), and 94.78% of their population carries out online activities, one of the highest percentages in the world. Therefore, we carried out our study on UK citizens.

2. The affect heuristic in cybersecurity risk perception

In the risk literature and PMT literature, risk perception has often been studied from a cognitive perspective (Loewenstein et al., 2001). Studies consider the impact of threat-related knowledge, voluntariness and control over exposure, newness, catastrophic potential and severity of consequences (van Schaik et al., 2017). Yet cognitive evaluations only partially explain variations in risk perceptions and risk-related decision-making (Loewenstein et al., 2001). Therefore, the need for a different approach, in other words 'risk as feelings', has been proposed (Loewenstein et al., 2001). According to Loewenstein et al., risk judgment involves not only cognitive evaluations, but also affect as an essential influence. They argue that risk is not calculated, but based on affect. If risk perception is indeed based on affect, risk perception will not be subject to cognitive evaluations to the extent that cognitive accounts of judgement would lead us to believe. Loewenstein et al. (2001) use the

terms "affect" and "feelings" interchangeably. However, we focus on affect, the central concept in the affect heuristic, which is the subject of the current research. Here we adopt Russell's (2003) influential conceptualisation of affect as discussed in the previous section.

The affect heuristic has been proposed in domains other than cybersecurity, such as health and natural disasters, to explain variation in risk perception (Bowen et al., 2004; Siegrist et al., 2006; Terpstra et al., 2009). According to the affect heuristic, people's perceptions of the risk of an activity or technology are influenced by their affect related to the activity. In particular, Finucane et al. (2000) claim that the affect heuristic can explain the relationship between perceived risk and perceived benefit. According to the authors' objective analysis, risk and benefit are positively correlated in the external world. However, in people's risk perceptions of technologies and activities, risk and benefit are negatively correlated (Finucane et al., 2000). The affect heuristic suggests that this is due to people's affect triggered by these technologies and activities. For example, if people feel that browsing the dark net is a good thing to do, they are likely to consider it beneficial and having low risk.

Finucane et al. (2000) present two affect heuristic models and empirically evaluate these models with two experiments. Perceptions of the risk and benefit of various technologies (e.g., nuclear power) outside of the cybersecurity domain were analysed. According to their first affect heuristic model (Figure 1), the reason for the negative correlation between perceived risk and perceived benefit is that affect is *both* a negative determinant of perceived risk *and* a positive determinant of perceived benefit. According to the second affect heuristic model, framing a message in terms of the risks of an activity (e.g., information about the risks posed by the use of a technology) leads to two consequences (Figure 2, Panel [a]). First, the level of perceived risk acts as a negative determinant of affect. Second, in turn, the level of affect acts as a positive determinant of perceived benefit. Therefore, affect is a mediator of the effect of perceived risk on perceived benefit. Similarly, when the message emphasises the beneficial nature of an activity, affect is a mediator of the negative effect of perceived benefit on perceived risk (Figure 2, Panel [b]).

Although Finucane et al. (2000) present their results as evidence for the affect heuristic, they only provide indirect evidence. They did not directly test their two affect heuristic models. In particular, they did not measure or analyse affect, even though this is the main factor in both models. Moreover, to our knowledge, research reporting direct testing of the affect heuristic remains scarce outside of the domain of cybersecurity and has not been tested directly in cybersecurity.

3. The current research

Despite the limitations of affect-heuristic studies, cybersecurity research has invoked this heuristic as an explanation for variations in risk perceptions (Garg & Camp, 2013). For example, in cybersecurity, the malicious activities of insiders have been attributed to the affect heuristic (Farahmand & Spafford, 2013). To our knowledge, the affect heuristic has not been tested in the domain of cybersecurity, neither indirectly (without measuring affect, as in Finucane et al. [2000]) nor directly (by

measuring and analysing the impact of affect on perceived risk and perceived benefit). The current research is original by directly testing both of Finucane et al.'s affect heuristic models in cybersecurity. The research is theoretically significant by testing these model-based explanations of risk perception and practically significant by presenting potential applications of the results to cybersecurity.

4. Study 1: testing the affect heuristic with affect

The aim of Study 1 is to test Finucane et al.'s (2000) first affect heuristic model (Figure 1), where affect is a negative determinant of perceived risk and a positive determinant of perceived benefit.

4.1. Method

4.1.1. Research design

The design is based on that of Finucane et al.'s (2000) Study 1. However, the design is novel. First, its focus is specifically on cybersecurity threats, in addition to other threats. Second, in order to allow us to fully test the first affect heuristic model, affect was measured. This is essential in extending Finucane et al.'s (2000) research, to more completely validate their model. The measured variables are (1) perceived risk, (2) perceived benefit and (3) affect. The order of risk rating and benefit rating was counterbalanced (risk followed by benefit or vice versa).

4.1.2. Participants

Participants were recruited through a large UK online survey panel (June-July 2018). All panel members aged 18 or over were eligible and these were automatically invited by the panel organisers. There were 63 participants (exceeding Finucane et al.'s [2000] sample size of 54). The sample was balanced in terms of gender (female/male split: 51%/49%). Mean age was 51.19 (*SD* = 13.08), with a range of 22 to 79. Most participants' highest level of completed education was high-school diploma or equivalent (43%) or degree (33%). Most participants were employed (57%) or retired (25%). A majority used the Internet three times or more per day (60%) and most spent several hours (46%) or one hour (24%) per day using the Internet.

4.1.3. Stimuli and measurement

Stimuli (Appendix 1) were 14 Facebook privacy- and security settings (S1-14) studied by Van Schaik et al. (2018), 2 non-cybersecurity-related Internet activities (S15-16), 3 further cybersecurity-related Internet activities analysed in Study 2 (S17-19) and a subset of 3 (S20-S22) out of 23 activities/technologies also studied by Finucane et al. (2000). Another 3, out of 23 activities/technologies from Finucane et al. (2000), were used as practice stimuli (Appendix 2).

Perceived risk and perceived benefit were measured using 7-point Likert scales (as in Finucane et al., 2000), with end-points 'not at all risky (beneficial)' and 'very risky (beneficial)' (see Appendix 2). Although Finucane et al. (2000) did not measure or analyse affect, this is essential in testing affect heuristic models. More fundamentally, the authors did not theoretically specify the affect concept. We adopt Russell's (1980) influential circumplex model of human affect to conceptualise affect (5915/11728 citations according to Scopus/Google Scholar, 9/9/2019). The model

has two dimensions: *valence*, ranging from miserable to happy, and *arousal*, ranging from sleepy to aroused. We used a 5-point self-assessment manikin for measuring valence and arousal (Bradley & Lang, 1994). We understand Finucane et al.'s (2000) conceptualisation of affect as valence, but we analysed both valence and arousal to demonstrate specificity of the affect heuristic to valence.

4.1.4. Procedure

Affect was measured in all affect-rating trials *before* either risk or benefit ratings for all stimuli. This is because, according to the first affect heuristic model, the determinant (affect) needs to precede the consequents (risk and benefit). For each participant, the order of stimulus presentation was randomised, but (once randomised) the same order was used for affect rating, risk rating and benefit rating. This was to ensure a constant gap (of 22 - 1 =) 21 stimuli between the affect rating and the risk/benefit rating, and between the risk/benefit rating and the benefit/risk rating). Therefore, in the affect-rating trials S1-S22 were presented in random order. Then, in the risk rating trials, the same stimuli were presented in the same order, followed by the benefit rating trials (or *vice versa* when benefit rating trials came first).

4.2. Results and discussion

To test the first affect heuristic model, Pearson's correlation *r* was analysed between valence and, (1) risk (negative correlation hypothesised) and (2) benefit (positive correlation hypothesised). In support of the hypothesis for risk, the correlation with valence was negative, with an average correlation over stimuli of -0.40 (Figure 3), bootstrapped BCa *CI*(0.95) = [-0.46; -0.35]. In support of the hypothesis for benefit, the correlation with valence was positive, with an average correlation of 0.52 (Figure 4), bootstrapped BCa *CI*(0.95) = [0.44; 0.60]. Nonetheless, when benefit was partialled out, the correlation between risk and valence reduced to a small effect size, *pr* = -0.22, on average (Figure 3). Similarly, when risk was partialled out, the correlation between benefit and valence reduced, but to a lesser extent, to medium-sized effect, *pr* = 0.43, on average (Figure 4). In the light of these results, support for the hypothesis for risk was clearly reduced and limited, but support for the hypothesis for benefit remains relatively strong.

For comparison with Finucane et al.'s (2000) results, we also calculated the correlation between risk and benefit per stimulus. They presented these correlations as evidence for the model, but did not measure or analyse the main factor affect. Therefore, their results can, at best, provide only indirect evidence for the hypothesis. In support of their results, we found negative correlations, with an average of -0.46 (Figure 5), bootstrapped BCa *CI*(0.95) = [-0.51; -0.40]. Nevertheless, in our analysis, 13 out of the 19 correlations between risk and benefit remained medium-sized or large and statistically significant when affect was partialled out, with an average of -0.29 (Figure 5). Consequently, although affect may be a determinant of both perceived risk and perceived benefit (even though the evidence for risk is limited, according to the results presented above), this does not imply that affect fully explains the shared variance between risk and benefit. However, this was precisely Finucane et al.'s (2000) implicit assumption and their

basis for interpreting their results as support for the affect heuristic. By contrast, with our analysis of the correlations between affect and risk and between affect and benefit, we directly test and provide appropriate evidence for the first affect heuristic model.

An analysis by stimulus category (Figure 6) shows that for most categories the correlations of valence with risk and valence with benefit were substantial (> 0.32), providing evidence for the affect heuristic. The exceptions were security settings for access to shared information in Facebook (e.g., keep a restricted list of friends) and hacking. We notice large discrepancies in the pattern of correlations between risk and benefit, on the one hand, and either valence and risk, or valence and benefit (e.g., for hacking), on the other. Therefore, if we were to purely use the correlation between risk and benefit (as Finucane et al. [2000] did) rather than the correlations between valence and risk and between valence and benefit, we would incorrectly conclude that there is substantial evidence for the affect heuristic for some of the stimuli. We would also incorrectly conclude the opposite for some other stimuli. In particular, support for the affect heuristic would be substantially overestimated for hacking and keeping a restricted list of friends (large correlation between risk and benefit, but small correlations between valence and risk and between valence and benefit). However, support would be substantially underestimated for allowing others to post on timeline and allowing all old posts to be shared (large correlations between valence and risk and between valence and benefit, but small correlation between risk and benefit).

To demonstrate the specificity of valence as the main affect dimension, we conducted an analysis of correlations and partial correlations again, but now with valence replaced by arousal (the other affect dimension in Russell's circumplex model of affect). The pattern of results for arousal (presented in Appendix 3) was clearly different from that for valence. In particular, correlations and partial correlations between arousal and risk were negligible (average absolute values < 0.05) rather than medium-size and small, respectively, when valence was analysed. Correlations and partial correlations between arousal and benefit were small (average values 0.12) rather than large and medium-sized, respectively, when valence was analysed. The average correlation between risk and benefit remained unchanged when arousal was partialled out rather than reduced, when valence was partialled out. Therefore, the evidence for the first affect heuristic model is specific to the affect dimension of *valence*.

5. Study 2: testing the affect heuristic with affect and framing

The aim of Study 2 is to test Finucane et al.'s (2000) second affect heuristic model (Figure 2). In this model, the effect of perceived risk (benefit) on perceived benefit (risk) is mediated by affect.

5.1. Method

5.1.1. Research design

The experimental design is based on that of Finucane et al.'s (2000) Study 2. However, the design is novel. First, the focus is on cybersecurity-related activities.

Second, in order to allow us to fully test the second affect heuristic model, affect was measured and analysed in relation to risk and benefit. This is the main factor in the model, but was not studied by Finucane et al. (2000); therefore, they were not able fully to validate their model.

Specifically, a mixed, 2 (domain: risk/benefit information)-by-2 (extent: high risk/benefit, low risk/benefit)-by-(4) (technology/activity: social network/social media, online banking, online gaming, nuclear power), design was used. The first two factors were used *between* subjects and the last factor *within* subjects. There were four experiment versions: (1) high risk, (2) low risk, (3) high benefit and (4) low benefit. The effect of risk or benefit framing information was analysed. In the analysis of the design per activity/technology, the dependent variables were perceived risk and perceived benefit.

In the analysis of the second affect heuristic model, the predictor was perceived risk and the outcome variable was perceived benefit in the high-risk- and low-risk conditions. The predictor was perceived benefit and the outcome variable was perceived risk in the high-benefit and low-benefit conditions. In both analyses, the mediator was affect.

5.1.2. Participants

Participants were recruited through a large UK online survey panel (July-August 2018). All panel members aged 18 or over were eligible and these were automatically invited by the panel organisers. There were 233 participants (n = 59, 62, 56 and 56 in the four experimental conditions, respectively, and exceeding Finucane et al.'s total sample size of 213). The sample was balanced in terms of gender (female/male split: 50%/50%). Mean age was 52.13 (SD = 13.13), with a range of 22 to 79. Most participants' highest level of completed education was high-school diploma or equivalent (41%) or degree (31%). Most participants were employed (52%) or retired (20%). A majority used the Internet three times or more per day (54%) and most spent several hours (42%) or one hour (23%) per day using the Internet.

5.1.3. Stimuli and measurement

The stimuli were activities/technologies. Three of these were cybersecurity-related (S1: online social network/social media; S2: online banking; S3: online gaming); a fourth, non-cybersecurity, technology (S4: nuclear power) was previously studied by Finucane et al. (2000). The following types of framing were used: high risk, low risk, high benefit and low benefit (see Appendix 4). Perceived risk, perceived benefit and affect were measured as in Study 1.

5.1.4. Procedure

During each trial, a stimulus (technology/activity) was presented with a description of the technology framed either in terms of risk (low or high) or benefit (low or high).

In high- and low-risk conditions, the risk rating for each stimulus preceded the affect rating (not included in Finucane et al., 2000), followed by the benefit rating. This is because, according to the model, the determinant risk needs to precede the

consequent affect and, in turn, the mediator affect needs to precede the consequent benefit.

In the high- and low-benefit conditions, the benefit rating for each stimulus preceded affect rating (not included in Finucane et al., 2000), followed by the risk rating. This is because, according to the model, the determinant benefit needs to precede the consequent affect and, in turn, the mediator affect needs to precede the consequent risk. In all conditions, the order of stimulus presentation was randomised per participant.

5.2. Results and discussion

5.2.1. Analysis of message effect

The effect of the framing of risk and benefit was tested with 2-way analysis of variance (ANOVA) for each stimulus.

5.2.1.1 Perceived risk

For *nuclear power*, the main effect of extent (F [1, 229] = 12.22, $p < 0.001$, *pes* = 0.05) and the interaction effect between extent and domain (risk or benefit) (F [1, 229] = 14.19, $p < 0.001$, *pes* = 0.06; see also Figure 7) were significant. Perceived risk was higher under the high-risk condition than under the low-risk condition, but the difference was exceedingly small between the two benefit conditions.

For *social media*, the main effect of domain (risk or benefit) was significant (F [1, 229] = 5.02, $p = 0.03$, *pes* = 0.02; see also Figure 7). Perceived risk was higher under the risk conditions and the difference between the high-risk and the low-risk condition was exceedingly small.

For *online banking*, the interaction effect between extent and domain (risk or benefit) (F [1, 229] = 7.41, $p = 0.01$, *pes* = 0.03; see also Figure 7) was significant. Perceived risk was higher under the high-risk condition than under the low-risk condition, but the difference was between the two benefit conditions was in the opposite direction.

For *online gaming*, neither the main effects of domain ($F < 1$) and extent (F [1, 229] = 2.58, $p = 0.11$, *pes* = 0.01) nor the interaction effect ($F < 1$) were significant.

5.2.1.2 Perceived benefit

For *nuclear power*, the interaction effect between domain (risk or benefit) and extent was significant (F [1, 229] = 6.62, $p = 0.01$, *pes* = 0.03; see also Figure 8). Perceived benefit was higher under the high-benefit condition than under the low-benefit condition, but the difference was small under the two risk conditions.

For *social media*, neither the main effects of domain (F [1, 229] = 3.50, $p = 0.06$, *pes* = 0.02) and extent ($F < 1$) nor the interaction effect ($F < 1$) were significant.

For *online banking*, the interaction effect between domain (risk or benefit) and extent was significant (F [1, 229] = 4.89, $p = 0.02$, *pes* = 0.03; see also Figure 8). Perceived benefit was higher under the high-benefit condition than under the low-benefit condition.

For *online gaming*, the main effect of domain (risk or benefit) on perceived benefit was significant (F [1, 229] = 6.74, p = 0.01, *pes* = 0.03; see also Figure 8). Perceived benefit was higher under the benefit conditions and the difference between the high-benefit and the low-benefit condition was exceedingly small.

In sum, framing was fully effective for nuclear power and online banking, as the interaction effect of domain and extent were significant on perceived risk and perceived benefit. Framing was partially effective for social media and online gaming (significant effect of domain for both). Another important observation is that, as expected, the pattern of results for risk and benefit differed. For example, for online banking, the effect of risk extent (low or high) increased perceived risk, but the effect of extent of benefit (low or high) did not increase perceived risk. At the same time, the effect of benefit extent (low or high) increased perceived benefit, but the effect of extent of risk (low or high) did not increase perceived benefit.

5.2.2. Analysis of the affect heuristic

According to Finucane et al.'s (2000) second affect heuristic model, perceived risk (benefit) is a determinant of perceived benefit (risk), and this effect is mediated by affect. Mediation analysis was conducted to test this conjecture.

5.2.2.1. Mediation of the effect of risk on benefit

Mediation of the effect of perceived risk on perceived benefit by affect was analysed with mediation analysis for high- and low-risk conditions separately. Six out of the eight analyses showed a significant negative indirect effect of perceived risk on perceived benefit (Figures 9-12). Over the eight analyses, the average size of the indirect effect was -0.25, CI(0.95) = [-0.49; -0.01]. Therefore, the results provide evidence for the second research model when perceived risk drives perceived benefit.

5.2.2.2. Mediation of the effect of benefit on risk

Mediation of the effect of perceived benefit on perceived risk by affect was analysed with mediation analysis for high- and low-benefit conditions separately. Four out of the eight analyses showed a significant negative indirect effect of perceived risk on perceived benefit (Figures 9-12). Over the eight analyses, the average size of the indirect effect was -0.20, CI(0.95) = [-0.58; 0.18]. Therefore, results provide mixed evidence for the second affect research model when perceived benefit drives perceived risk.

In sum, the mediation of the negative effect of risk on benefit was more consistent than the mediation of the negative effect of benefit on risk. The significant mediated effects were indirect-only (Zhao et al., 2010), in other words the effect of risk was fully mediated by affect, consistent with the affect heuristic. Here, the exception was the competitive mediation (Zhao et al., 2010) in the low-benefit frame for nuclear power. This indirect negative effect is consistent with the affect heuristic, but the direct positive effect is consistent with the idea that, in the real world, risk and benefit are positively correlated (Finucane et al., 2000). Even in the non-significant mediation effects of risk and those of benefit, the pattern was consistent with the affect heuristic: the indirect effect was negative and the confidence interval over the

eight analyses was skewed towards negative values.
To show the specificity of valence as the main affect dimension, we conducted the mediation analyses again, but now with valence replaced by arousal (the other affect dimension in Russell's circumplex model of affect). The pattern of results (Appendix 5) for arousal was clearly different from that for valence. In particular, there was no consistent evidence for mediation, as only 1 of the 15 indirect (mediated) effects was significant. Therefore, the evidence for the second affect heuristic model is specific to the valence dimension of affect.

6. General discussion

We carried out two studies to test the affect heuristic as proposed by Finucane et al. (2000). In particular, our studies sought to include measurements of affect, both in terms of valence and arousal, and to assess their impact on risk and benefit perceptions. We carried out these studies in the cybersecurity domain because it is crucial for us to understand how best to ensure that people have an accurate perception of both the risks and benefits of engaging with the online world. PMT suggests that accurate risk perceptions are likely to prompt people to take appropriate precautions while they enjoy the benefits of online activities. However, according to the affect heuristic, these perceptions are influenced by people's affect in response to the technology that is involved in the activity. We tested the impact of both of affect's constituent parts in our studies: valence and arousal, to determine whether either or both were influential in informing risk and benefit perceptions in the cybersecurity domain, within both of Finucane et al.'s (2000) proposed models.

The results of our first study did indeed support the first affect heuristic model with direct evidence in terms of impact of affect valence on perceived benefit, but its impact on risk was small enough for us to question the role of valence in informing risk perceptions. The impact of arousal on perceived risk and benefit, however, ranged from negligible to small. We found small to medium correlations between perceived risk and perceived benefit, which provides some indirect evidence for Finucane et al.'s (2000) first model. However, there were some inconsistencies between the two, empirically demonstrating the fallacy of testing the model by analysing the correlation between perceived risk and perceived benefit.

Our second study tested Finucane et al.'s (2000) second affect heuristic model within the cybersecurity domain. Here, our findings were not as convincing. We did find evidence for risk perception's influence on benefit perception, mediated by affect valence. Yet there was less compelling evidence of impact of benefit perception on risk perception mediated by affect valence, at least for the scenarios we used. Once again, we did not find evidence for mediation by arousal. In sum, our findings for affect mediating between perceived risk and perceived benefit apply only variably to affect valence (i.e. not for all scenarios).

6.1. The applicability of the affect heuristic to cybersecurity

Now that we understand the impact of affect valence on perceived risk and perceived benefit, we can consider how these findings can inform cybersecurity risk communication.

It might seem obvious that making people feel aroused by delivering a fear appeal message would lead to their taking precautions to reduce this experienced arousal. A number of authors have indeed suggested using fear appeals to encourage cybersecurity behaviours (Herath & Rao, 2009; Johnson & Warkentin, 2010; Liang & Xue, 2010; Ifinedo, 2012; Posey et al., 2011; Boss et al., 2015). Yet, our finding that affect arousal does not reliably inform perceived risk suggests that such arousal-based appeals might not be particularly effective in reality. Marett et al. (2019) tested the impact of fear appeals regarding ransomware and discovered that too-vivid messages could lead to maladaptive rather than protective responses, which seems to confirm our finding.

On the other hand, the fact that affect valence impacts perceived benefit, and that perceived benefit and perceived risk are negatively correlated (Study 1) is concerning in the cybersecurity domain. In effect, the lure of a particular risky activity could lead to a heightened perceived benefit, which could, in turn, dampen down perceived risk so that the person will not take sensible precautions. Indeed, Rhodes and Pivik (2011) found exactly this effect for risky driving. A liking for the activity led to reduced risk perceptions, and a discounting of the true risks of the activity. It seems that the perceived benefit was so heightened by liking that the perceived risk was dampened down so that it no longer mattered.

The fact that affect's mediation of the negative effect of perceived risk on perceived benefit was more consistent than the mediation of the positive effect of perceived benefit on perceived risk also has implications for cybersecurity messaging. Indeed, De Bruijn and Janssen (2017) make this very point in their insightful paper. This finding could skew the affect heuristic's predicted inverse relationship between perceived benefit and perceived risk. For example, a risky online service could offer someone an enticement to persuade them to use it (high perceived benefit leading to low perceived risk). A subsequent security warning message (inducing negative affect) could well lead them to question the advantages offered by the service (reduced benefit). The intervention might serve to amplify perceived risk disproportionally so that they discontinue usage, due to this skewed relationship, instead of engaging in a thoughtful and rational trade-off of risk and benefit. Essentially, the negative affect induced by heightened perceived risk is more influential than the positive affect triggered by perceived benefit.

6.2. The applicability of the "risk-as-feelings" hypothesis

Loewenstein et al. (2001) coined the term "risk as feelings". Their paper has made a huge academic impact (2780/5667 citations, according to Scopus/Scholar 9/9/2019). Yet, in other recent risk-related reviews, affect is not mentioned (Renn & Benighaus, 2013). Still, a number of researchers have carried out studies to measure the role of affect in risk perception in non-cybersecurity domains (Mathur & Levy, 2013; Lerner et al., 2015; Cottingham & Fisher, 2016), which prompted our interest in its impact in cybersecurity.

There is a potentially confounding factor in affect heuristic studies. Sometimes studies state that they are measuring affect's impact on risk, but the instruments they use seem to test emotions rather than affect (Keller et al., 2006; Taylor & Snyder,

2017). Russell's influential conceptualisation of affect (2003) provides clarity by distinguishing affect and emotion within a unified framework. First the author explains affect as a simple feeling, *core affect*, to which at all times people have conscious access to; this is a mix of two dimensions: valence (pleasure-displeasure) and arousal (activation-deactivation). Second, according to Russell, an emotional episode consists of several components, including core affect, antecedent event, affective quality, attribution, appraisal, instrumental action, physiological and expressive changes, subjective conscious experiences and emotional meta-experience. In Russell's dimensional approach to conceptualizing affect, core affect can be measured without necessarily being confounded by emotions (Bradley & Lang, 1994). This is the approach that we have followed in the current research by measuring affect with validated scales to capture both valence and arousal.

We know that people tend to judge hazards based on how they feel about them, and not based on a cognitive and objective assessment of the actual risks (Slovic and Peters, 2006; Bearth and Siegrist, 2016). Our study provided some confirmation of this in the cybersecurity domain.

Researchers have investigated the impact of the affect heuristic in other areas. For example, in their study of product innovations, King and Slovic (2014) found that where people were uncertain about risks or benefits, they relied on affect to make judgements. The triggering of the affect heuristic is also indicated when decision making is complex or when psychological resources are limited (Wu et al., 2018). In cybersecurity, the risks are often uncertain and the context complex, which could lead people to rely on affect instead of objectively evaluating actual probabilities.

In the field of gene technology, Siegrist and Sütterlin (2016) showed that a pre-existing attitude towards a particular technology would bias people's assessment of the risks and benefits of the technology, because of the heuristic affect. They demonstrated that even if people were given accurate information about risk, they interpreted it differently. Wu et al. (2018) refer to "external emotions" having an impact on the assessment. Hine et al. (2007) found supporting evidence for the affect heuristic related to the health risks of wood-burning fires. They discovered that people who owned wood-burning heaters had more positive associations (pre-existing positive affect), and this led to their downplaying the risks.

We also have to acknowledge differences in perspectives. Denscombe (2010) found that teenagers were reluctant to condemn their peers who smoked. They explain that the priorities of the authorities cannot be assumed to match the concerns of people being targeted with a risk-related message. Västfjäll et al. (2014) explain that "affect, accessible thoughts and motivational states influence perceptions of risks and benefits" (p. 527). Denscombe's findings reflect this: the motivations of the authorities are to reduce smoking; the young people are more focused on their group membership and their desire not to demonise one of their own. Assuming that everyone will respond in the same way to a particular risk is naïve: the same message will invoke different affects, and hence different judgements.

Extrapolating from this, consider that a cybersecurity expert presents people with accurate information about the benefits of using a particular technology in the cyber domain. A pre-existing affect invoked by negative encounters with the technology might lead to the message failing to adjust perceptions as anticipated. For example, a message focused on the benefits of encryption might trigger a negative affect. The person might not necessarily accept any proclaimed benefits, due to negative affect related to previous experiences of complexity and usage difficulties related to encryption technologies.

Watson et al. (2017) carried out an investigation into unlawful file sharing behaviours. They discovered that the affect heuristic also came into play in this domain. Perceived benefit reduced perceptions of risk so that information about the legal risks did not have much of an impact. They recommend that this kind of illegal activity be addressed by focusing attention on benefits of lawful alternatives rather than by focusing on trying to increase risk perceptions.

A study into the impact of communicating the risks and benefits of cancer tests to patients is reported by Scherer et al. (2018). Their findings, like ours, were that affect played a role in how people responded differently to receiving information about the risks or benefits of medical tests. In particular, they report that risk-related information reduced perceived benefit, as predicted by the affect heuristic. On the other hand, information about uncertainty of benefit did not impact perceived risk to the same extent.

6.3. Results in relation to existing models of human behaviour

Researchers have modelled the impact of other factors on risk and benefit perception. For example, Ganzach (2000) validates a model based on familiarity. He proposes the models shown in Figure 13. The first model depicts the impact of global preference (affect) on perceived risk and return (benefit), with unfamiliar stimuli; this corresponds to the first affect heuristic model. The second model represents the effect of perceived risk and return on global preference with familiar stimuli. This is a variation on the second affect heuristic model, under conditions when objective information is available about both risk and benefit.

Other researchers have investigated the role of *experience* in risk perception (De Dominicis et al., 2015; Golman et al., 2015; Raue et al., 2018). Researchers have studied the impact of experience from the inference perspective (Ganzach, 2000; Hassenzahl & Monk, 2010; van Schaik et al., 2012), anchoring and adjustment (Kahneman, 2011) and the theory of process memory (Vedadi & Warkentin, 2018).

Siegrist et al. (2000) investigated the role of social trust and found that *social trust* is a major predictive factor of perceived risks and benefits of a technology. Kim et al. (2015) also investigated risk perception in the context of social networking. Their finding was that the act of establishing a new link would lead to heightened risk perceptions, whereas those who acted to cement existing relationships would not experience this effect. This might go some way towards explaining our findings regarding access to shared information in Facebook, which were unlike those of other risky behaviours.

Slovic (2004) compares and contrasts experiential and analytic systems of thinking, and his analysis suggests that affect comes to play in the former, but not in the latter. These two systems of thinking act in parallel to inform our decisions. Pachur and Hertwig (2012) found evidence for the affect heuristic as well as the availability heuristic working in combination in informing decision-making. Finucane and Holup (2006) suggest that "risk as feelings" and "risk as analysis" combine to lead to a "risk-as-value" judgement. Other factors can also play a role in informing risk perception, such as age (Finucane, 2008), gender (Gustafsod, 1998), attachment (Renaud et al., 2019) and liking (Rhodes & Pivik, 2011). The affect heuristic is thus merely one factor influencing risk perception and feeding, together with other factors, into the complexity of the human brain to generate an eventual perception of both risk and benefit. Therefore, combining affect and other factors such as experience and social trust may be fruitful in future cybersecurity research on risk perception.

Pham and Avnet (2009) found that affect is relied on far more when the focus is *promotion* rather than *prevention*. Higgins (1987) explains that regulatory-focus theory specifies the existence of two distinct systems: promotion and prevention. The former is involved with regulating growth and cultivation. The latter, prevention, is concerned with protection and security. In particular, as Pham and Avnet (2009) explain, how goals are pursued differs between these two regulatory systems. Promotion is approach-oriented whereas prevention is avoidance-oriented. Promotion suggests a strategy of exploration and seizing of opportunities whereas prevention is essentially vigilance-focused and cautionary. Because the usual cybersecurity messages focus on prevention (Ayala, 2016; Holland and Shey, 2015; Williams et al., 2018), messages prompting negative affect, as Sunstein (2003) advises, might not be the best way of discouraging risky behaviours and promoting precautionary behaviour in the cybersecurity context.

All of these studies highlight the importance of *decision context* when considering risk perception, as emphasised by Bateman et al. (2007) and Leiserowitz (2006). This underlines the difficulty of studying affect using surveys – context cannot be as rich and informative as it is when real life decisions are made. Future studies may therefore benefit from studying affect in (simulated) real-world situations.

7. Conclusion and Future Work

In this paper, we present original research that directly tests both of Finucane et al.'s affect heuristic models in cybersecurity. Our research is theoretically significant by testing these model-based explanations of risk perception and practically significant by presenting potential applications of the results to cybersecurity. Specifically, we present two controlled studies that we carried out to test the validity of the two affect heuristic models proposed by Finucane et al. (2000) in the cybersecurity domain. Our results provide support for the first model, in particular, first, through direct evidence (the impact of affect valence on risk- and benefit perception, as specified by the model) and, second, indirect evidence (the correlation between risk- and benefit perception, not specified by the model). However, there were some inconsistencies between the two, empirically demonstrating the flaw of testing the model by analysing the correlation between perceived risk and perceived benefit.

The evidence for the second model was variable, with evidence for risk perception's influence on benefit perception mediated by affect valence. Furthermore, only affect valence (but not arousal) had an impact on risk perception in both studies.

A number of future research directions are suggested by this research. First, we need to design experiments to test cybersecurity message designs that do not aim to arouse, but rather aim to encourage precaution uptake by focusing on the benefits thereof. This is because affect is more impactful in a promotion-based mindset (Pham and Avnet, 2009). Second, Sunstein (2003) argues that if the affect heuristic has an impact on risk perception, we should be able to alter affect and thereby change the perception. Indeed, Slovic (2004) explains that this is exactly what marketers and advertisers do. Sunstein cites LeDoux (2003) and Nussbaum (2003) to make the point about the way cognition can be changed by alterations in emotions. This, too, might be a fruitful avenue for future investigations in the cybersecurity domain.

Third, it would be beneficial to repeat these studies with a new set of risky cybersecurity behaviours so as to establish the generality of the findings. More generally, the extent to which the affect heuristic is generally applicable to the cybersecurity domain can be tested by establishing boundary conditions. For example, according to the person-artefact task model (Finneran and Zhang, 2003), it is worth exploring whether the affect heuristic holds across different persons (familiarity of users with Internet activities and technologies), artefacts (Internet activities and technologies [a number of which were included in our studies]; the reliability and effectiveness of technologies in goal achievement), and task characteristics (type of use [e.g. hedonic or goal-oriented, Van Schaik & Ling, 2009]; the frequency and duration of using the technologies). Furthermore, in organisational computer use, does the power of the affect heuristic vary across different organisational contexts?

This research contributes to the body of knowledge related to the impact of the affect heuristic. Future decisions and choices in cybersecurity need to be based on relevant theories and models and with this research we support the cumulative knowledge building that can facilitate and inform this process.

This is a peer-reviewed, accepted author manuscript of the following research article: Schaik, P. V., Renaud, K., Wilson, C., Jansen, J., & Onibokun, J. (2020). Risk as affect: the affect heuristic in cybersecurity. *Computers and Security*, *90*, [101651]. https://doi.org/10.1016/j.cose.2019.101651

Figure captions

Figure 1. First affect heuristic model.

Figure 2. Second affect heuristic model.

Figure 3. Correlations and partial correlations of valence with perceived risk (Study 1).

Figure 4. Correlations and partial correlations of valence with perceived benefit (Study 1).

Figure 5. Correlations and partial correlations of perceived risk with perceived benefit (Study 1).

Figure 6. Correlations and partial correlations of risk and benefit with valence (Study 1). rVR: correlation between valence and risk. rRB: correlation between risk and benefit. rVB: correlation between valence and benefit.

Figure 7. Perceived risk as a function of risk- and benefit information (Study 2).

Figure 8. Perceived benefit as a function of risk- and benefit information (Study 2).

Figure 9. Mediation analysis, nuclear power – mediator: valence (Study 2).

Figure 10. Mediation analysis, social media – mediator: valence (Study 2).

Figure 11. Mediation analysis, online banking – mediator: valence (Study 2).

Figure 12. Mediation analysis, online gaming – mediator: valence (Study 2).

Figure A3.1. Correlations and partial correlations of risk and benefit with arousal (Study 1). rAR: correlation between arousal and risk. rRB: correlation between risk and benefit. rAB: correlation between arousal and benefit.

Figure A5.1. Mediation analysis, nuclear power – mediator: arousal (Study 2).

Figure A5.2. Mediation analysis, social media – mediator: arousal (Study 2).

Figure A5.3. Mediation analysis, online banking – mediator: arousal (Study 2).

Figure A5.4. Mediation analysis, online gaming – mediator: arousal (Study 2).

Note: colour should not be used for any figures in print.