Can Trust be Trusted in Cybersecurity?

Daniel Pienta
Baylor University, dan_pienta@baylor.edu

Stefan Tams
HEC Montreal, stefan.tams@hec.ca

Jason Thatcher
The University of Alabama, jason.b.thatcher@gmail.com

Can Trust be Trusted in Cybersecurity?

Daniel Pienta
Baylor University
dan_pienta@baylor.edu

Stefan Tams
HEC Montreal
stefan.tams@hec.ca

Jason Bennet Thatcher
University of Alabama
jbthatcher1@cba.ua.edu

Research-in-progress

Abstract

Human compliance in cybersecurity continues to be a persistent problem for organizations. This research-in-progress advances theoretical understanding of the negative effects of trust formed between individuals and the cybersecurity function (i.e., those responsible for protection), cybersecurity system (i.e., the protective technologies), and organization (i.e., those verifying (e.g., hiring, championing, vouching.) the cybersecurity department) that leads to suboptimal compliance behaviors. In contrast to the current information security literature that focuses on how organizations can induce compliance, this study begins to provide understanding into the degradation of compliance through organizational actions. Additionally, understanding is provided on how to combat the negative effects of trust. An integrated model is conceptualized using the theories of trust and attention. This model provides the theoretical foundation to study the role of dark side trust in the context of cybersecurity and provides initial mechanisms to reduce it. By developing this conceptualization of dark side trust and model, this study contributes to the general study of trust in information systems research outside of the domain of cybersecurity.

1. Introduction

Organizations are increasingly developing more complex cybersecurity ecosystems that rely on people, technology, and processes to function effectively [34]. Information security research (ISec) has identified the criticality of studying human compliance behavior in cybersecurity ecosystems [8]. ISec research studied human compliance using a rich theoretical base such as deterrence [10, 39], motivation [7, 23], fear [22], accountability [44], and mindfulness [21]. Interestingly, there has been little research attention directed at the role of trust in human compliance in the context of cybersecurity.

Trust in information systems (IS) research has largely been shown to be positive in different contexts such as electronic markets [2], e-commerce [40], website design [9], and online social networks [3]. Trust though has also been conceptualized and shown to result in negative consequences [12, 34, 37, 50] as individuals may maintain trust unconditionally and over and above evidence to the contrary [12]. Trust, therefore, has been shown to have a "bright" and "dark" side stimulating both positive and negative consequences for individuals and organizations. Similar to the context of project management, trust like commitment can be beneficial and detrimental. Escalating commitment can lead to detrimental consequences by continually absorbing resources without delivering benefits [24, 25]. Trust like commitment can potentially lead to detrimental consequences in contexts rife with risk, uncertainty, and vulnerability like cybersecurity.

The dark side of trust, which we refer to as when an individual maintains trust unconditionally despite contradictory relevant stimuli, is important in the context of cybersecurity as it can be detrimental to the defense of cyber-attacks. For example, in cybersecurity, one with dark side trust may trust that the cybersecurity system will filter all phishing emails and therefore not correctly identify malicious emails even though stimuli (e.g., an unsecure domain, incorrect domain address, request for urgent confidential) should alert them otherwise. Conversely, bright side trust, when an individual maintains trust until presented with contradictory relevant stimuli, can be beneficial. Individuals with bright side trust, contrarily, may identify a malicious email that bypasses cybersecurity countermeasures when presented with stimuli that alert them the email is unsafe.

Trust, be it dark or bright, is a highly relevant mechanism in understanding mitigation of uncertainty between people, organizations, and technology [33]. When individuals are in situations of uncertainty, risk, and vulnerability, like during cyber-attacks they tend to

URI: https://hdl.handle.net/10125/64264
978-0-9981331-3-3

rely on agents (e.g., cybersecurity department, cybersecurity system) they trust to guide behaviors. When relying on trusting agents it is conceivable that it may result in negative outcomes thus trust's dark side manifests. In this research-in-progress, we provide understanding of the role of dark side trust in high-risk contexts like cybersecurity as the consequences can be detrimental to the individual, organization, and society as a whole.

ISec research has a rich theoretical and practical understanding of how to facilitate compliance through mechanisms such as motivation [22, 23], accountability [44, 45], and deterrence[10, 16, 39], amongst others. Interestingly, existing ISec research has yet to fully understand how organizations can inhibit human compliance through mechanisms like dark side trust as implied above. In this research-in-progress study, we investigate the dark side of trust in cybersecurity and its inhibiting effect on human compliance. In doing so, we develop a conceptual model on cybersecurity trust, attention to cybersecurity, and compliance behaviors. We examine the following research question:

How does dark side trust inhibit compliance in the context of cybersecurity?

This research-in-progress contributes to ISec literature and practice in several ways. This research builds upon dark side trust in the context of cybersecurity [34] by studying its effect on attention paid to cybersecurity as a mechanism to reduce compliance. This begins to provide understanding of how organizations can inadvertently reduce compliance. Insight is also provided into the moderating roles of cybersecurity mindfulness, suspicion, and intention to protect between trust and attention and how they can combat dark side trust. The study also contributes to the IS literature in general, by building upon the pervious conceptualizations of the nuances of trust, that trust is not always beneficial. From a practical perspective, this study will enable Chief Information Security Officers (CISOs), cybersecurity managers, and system designers to understand how to protect against dark side trust in cybersecurity. By providing understanding into the moderating role of cybersecurity mindfulness, suspicion, and intention to protect this study allows cybersecurity stakeholders to incorporate these three levers into training and system design to increase attention for compliance and alleviate the dark side of trust detrimental consequences.

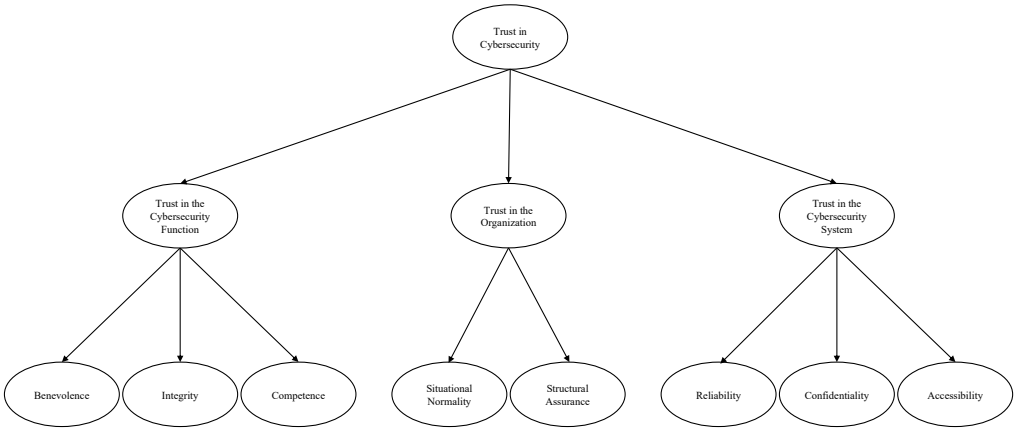
2. Trust In Cybersecurity

Trusting relationships have been studied from four primary perspectives between 1) people and groups 2) people and organizations 3) organizations, and 4) people and technology [38]. Cybersecurity and proper defense though relies on the individual trusting the technology (i.e., cybersecurity system), people (i.e., cybersecurity department), and the organization. Each of these trusting agents can deploy cybersecurity countermeasures that an individual can rely on to prevent and defend against cyber-attacks.

In the context of cybersecurity, an organization can develop legal strategies to prosecute, terminate, or discipline individuals to use fear to motivate identification and protective behaviors. The cybersecurity function can provide training for individuals to call upon to identify and protect against a cyber-attack. The cybersecurity system can filter and identify malicious files or emails to warn and notify the user of potential harm. Collectively, these different trusting agents can shape the individual's trusting beliefs about cybersecurity. We, therefore, posit that trust in cybersecurity is a multidimensional superordinate construct comprised of trust in the cybersecurity function, cybersecurity system, and organization (Figure 1).

In conceptualizing trust in cybersecurity as a multidimensional superordinate construct we adapt and

Figure 1. Trust in Cybersecurity as Multidimensional Superordinate Construct



contextualize items for trust in the cybersecurity function, trust in the organization, as well as develop new measures and dimensions for trust in a cybersecurity system. Trust in the cybersecurity function relates to those responsible for protecting the individual with the dimensions of benevolence, competence, and reliability [15, 30, 38, 40]. Institutional trust in cybersecurity relates to the beliefs the individual has about cybersecurity in the organizational protective context [29] with the dimensions of situational normality and structural assurances. Trust in the cybersecurity system relates to how the individual believes he or she is safeguarded by protective technologies from a cyber-attack [29].

Trust in the cybersecurity function was adapted from trust in a specific technology. Trust in a specific technology was conceptualized as users' perceptions of technology attributes. McKnight et al. [29] included system-like constructs of reliability, functionality, and helpfulness, which were derived from integrity, competence, and predictability [29]. Similar to the conceptualization of trust in a specific technology we argue that trust in cybersecurity system elicits different system-like attributes since it primarily serves as a protective technology. We adapt new dimensions that are derived from reliability, functionality, and helpfulness and incorporate the CIA triad (confidentiality, integrity, availability) of security into trust in cybersecurity system dimensions of reliability, confidentiality, and availability (Table 1). Reliability is the assurance that confidential information and computing resources are protected. Confidentiality is that access to confidential information and computing resources is restricted to only those who need it. Availability is the guarantee of access to confidential information and computing resources by authorized persons when necessary. Overall we posit that this new conceptualization of trust in cybersecurity reflects the holistic socio-technical nature of trust in various cybersecurity countermeasures elicited by different facets of the organization.

2.1 Dark Side Trust in Cybersecurity

When an individual extends trust they are effectively assuming risk, uncertainty, and making himself or herself vulnerable to the trustee [11, 18, 33]. Establishing trust allows an individual to close informational gaps in situations of risk, uncertainty, and vulnerability by abdicating responsibility to the trustee [15]. This indicates that issues of concern like uncertainty and risk can be quelled when trust is placed in a trustor.

In mitigating uncertainty, risk, and vulnerability IS research has found that trust is a key factor in reducing concerns of information privacy, information security, and opportunism concerns in an e-commerce context [33]. Taking an agency perspective, in the context of e-commerce, Pavlou et al. [33] found that sellers send signals by posting information security policies, explaining information security technologies, and third party-verification of information security practices. These pre-contract signals, in turn, lead to trusting beliefs by buyers and mitigate information security concerns. Pavlou et al. [33] also found post-contract incentives play a role as high-quality sellers have reason to safeguard information security due to reputational concerns. Similarly, in the context of cybersecurity, the cybersecurity function sends trusting signals and has incentives to mitigate cyber-attacks in the organization. For example, the cybersecurity function implements information security and safe computing policies, explains the purpose of and provides protective technologies via the cybersecurity system, and the organization verifies the cybersecurity function and system via support. Additionally, the cybersecurity function has incentives to mitigate cyber-attacks due to reputational concerns.

Frequently, individuals are tasked with the making decisions fraught with uncertainty, risk, and vulnerability during the course of cyber-attacks. Trust allows individuals to reduce feelings of uncertainty, risk, and vulnerability by placing trust in the cybersecurity function, system, and/or the organization. For example, social engineered cyber-

Table 1. Trust in cybersecurity system dimensions

Trust Between People (McKnight et al. 2011)	Trust in a Specific Technology (McKnight et al. 2011)	Trust in a Cybersecurity System (This paper)
Competence – The efficacy of the trustee to fulfill a promise in terms of their ability or power to do something for us (McKnight et al. 2011, p. 12:5)	Functionality - whether one expects a technology to have the capacity or capability to complete a required task (McKnight et al. 2011, p. 12:5)	Confidentiality - access to confidential information and computing resources is restricted to only those who need it
Benevolence – When the trustee cares enough to offer help when needed (McKnight et al. 2011, p. 12:5)	Helpfulness – if the help function of the technology is adequate and responsive (McKnight et al. 2011, p. 12:5)	Availability - the guarantee of access to confidential information and computing resources by authorized persons when necessary
Integrity – The hope that trustees are consistent, predictable, and reliable (McKnight et al. 2011, p. 12:5)	Reliability – the technology works consistently and predictably (McKnight et al. 2011, p. 12:5)	Reliability - the assurance that confidential information and computing resources are protected

attack like phishing attempt to psychological manipulate individuals by acting as a trusted entity [19]. Individuals, when faced with a socially engineered phishing email, may defer the decision regarding if the email is safe to a trusting agent. They may trust that 1) the cybersecurity system did not filter the email therefore the email is safe, 2) the cybersecurity function did not block the sender or notify of phishing threats therefore the email is safe, and 3) the intimate details of the operation of the organization or branding are in the context of the email therefore it is safe. While these may raise red flags and elicit the bright side of trust in enacting proper defense (e.g., checking the domain for safety, reviewing the sender email address rather than relying on the email header, etc.) there may also be negative consequences due to the dark side of trust. The dark side of trust could manifest resulting in the failure of the individual to identify the phishing email as they abdicate responsibility to trusting agents. This could result in the individual providing confidential information (e.g., username and passwords), downloading a malicious file, or even responding and wiring funds to a cybercriminal.

Overall, the dark side of trust in management literature has been found to have negative consequences for individuals and organizations such as allowing unethical behavior [50], overreliance on automation [27], poor judgment [13], management complacency [26], and underperformance [32]. In light of this understanding in management literature, little attention has been paid to the dark side of trust in ISec and IS research. The dark side of trust therefore, could have negative consequences in high-risk situations like cybersecurity as it may alleviate individuals' cybersecurity concerns in an organization and reduce compliance behaviors.

3. Attention to Cybersecurity

Cybersecurity is riddled with failures of the human element with almost 90 percent of attacks relying on deceiving humans [52]. For example, the Equifax data breach of 2017 could have been avoided had humans updated security patches when released in March 2017. The cyber-attack was launched from May to June 2017, yet the breach was not discovered until the end of July 2017 and by that time it is estimated 143 million records had been breached [53].

One of the primary purposes of cybersecurity ecosystems is the detection and identification of malicious cyber-attacks [31]. These systems rely on socio-technical countermeasures to detect, identify, and respond in mitigating an attack. Cybersecurity systems use technical means like information security warnings to engage users and elicit pro-cybersecurity behaviors [43]. Recently, ISec research has turned to collective intelligence via knowledge management systems to leverage the human element by creating human firewalls in identifying malicious email [20]. Leveraging humans and engaging them as an effective countermeasure has proven elusive since it requires them to pay attention to information security warnings, phishing emails, information security training, and notifications [21, 36, 45, 48, 49].

When humans, a key component of cybersecurity defense, fail to detect cyber-attacks that bypass technical countermeasures the individual and organization are subject to detrimental consequences. ISec research has found that security is not a priority for individuals [17]. Additionally, there is little understanding of basic cybersecurity and associated countermeasures [51]. Inherently, this shows that individuals are not concerned about cybersecurity or do not understand it.

To combat the tendency to ignore cybersecurity, ISec research has turned its attention to how individuals tune out information security warnings [1, 43, 45]. In a series of studies, it was found that over time individuals become habituated to repeated security warnings and thus fail to engage with them. Polymorphic design of information security warnings was found to decrease habituation and increase adherence to said warnings [1, 43]. This research shows that individuals over time are vulnerable to paying less attention to information security warnings. Thus garnering the attention of individuals in cybersecurity is critical in the effective detection, response, and mitigation of cyber-attacks.

4. Compliance

Human compliance in cybersecurity has received substantial attention from a diverse array of theoretical lenses, yet is still an elusive problem for academics and practice. As stated previously, ISec research has noted the criticality of behavioral security and the importance of inducing compliance behaviors in the effective defense of cyber-attacks. [8].

Security compliance policies and associated security education, training, and awareness programs are designed to provide guidance for individuals to protect the organization [22]. A commonality amongst these policies and programs is that it allows the cybersecurity function to articulate and demonstrate expectations for individual pro-cybersecurity behaviors. When individuals fail to engage in these compliance behaviors the organization is at risk. Although, we have a rich understanding of what drives compliance behaviors theoretically such as fear [22],

accountability [44], and deterrence [10], ISec research regarding what lessens these behaviors is still relatively new. Therefore, understanding what diminishes individuals to engage in cybersecurity compliance, like dark side trust is important for future ISec research.

5. Cybersecurity Mindfulness, Cybersecurity Suspicion, and Intention to Protect

Mindfulness and suspicion are concepts that are similar in that they engage users in cognitive activation [5, 6, 20, 41]. Although similar, there are distinct individual differences between the constructs in that mindfulness does not require the element of malicious intent, while suspicion does. For instance, an individual can engage in mindful behaviors even if they do not sense they will be harmed, while suspicion does not rise unless the individual believes they will be harmed. Mindfulness invokes engagement with users in one dimension through awareness of multiple perspectives [41]. Conversely, suspicion invokes engagement through the dimension of cognitive activation, where the individual mentally conjectures multiple explanations for harm [6]. For instance, in identifying phishing emails an individual that is mindful might engage in the multiple perspectives related to past compliance training, while an individual that is suspicious may inherently question all emails looking for alternative explanations of why the email can be harming above and beyond what he or she has been trained for. Regardless, both mindfulness and suspicion in cybersecurity further engage individuals in deeper thinking.

Individuals with an intention to protect may be more likely to engage in compliance behaviors. Intention has been shown to be a key factor in driving behavior [46, 47]. Past ISec and IS research has shown the strength of this relationship and those that have a positive behavioral intention are more likely to engage in subsequent behaviors [7, 23, 36, 46, 47]. Therefore, those that have an intention to protect may be more likely to engage in more concentrated efforts for cybersecurity.

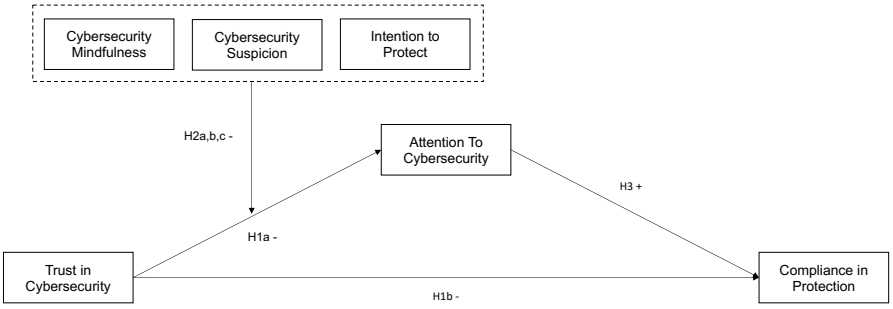
6. Research Model

We draw on the theoretical bases of trust, attention, and compliance behaviors to build an integrative model of the consequences of dark side trust in cybersecurity (Figure 2). In this view, we propose that dark side trust in cybersecurity lessens attention to cybersecurity and subsequent compliance behaviors. Bright trust, in general, has been shown as a mechanism that lessens concern in situations of uncertainty, vulnerability, and risk for beneficial consequences [33]. This model, in turn, posits that dark side trust can have the same effects in that it can alleviate concerns for cybersecurity causing individuals attention to cybersecurity to decrease and thus making them more susceptible to cyber-attacks since security compliance behaviors are not followed. Table 2 presents the model constructs and definitions.

Trust and its beneficial side is shown to have positive outcomes in the contexts of e-commerce [2, 33, 40], social networks [3], and recommendation agents [4]. In these contexts, trust has been shown to mitigate uncertainty, risk, and vulnerability leading to improved exchanges between the trustee and the trustor. Specifically, Pavlou et al. [33] found that trust mitigated uncertainty related to information security in online shopping. Customers' perceptions of risk in information security were reduced the more they trusted the organization thus increasing purchasing intentions [33]. Similar to this dark side trust can mitigate individuals concerns for cybersecurity in an organization due to trust in cybersecurity.

We suspect that individuals may not be concerned about cybersecurity in the organization since they trust the cybersecurity function, system, and organization prevent cyber-attacks. Therefore, individuals may not pay attention to cybersecurity since they trust

Figure 2. Dark side trust research model



cybersecurity agents in the organization will have taken care of mitigating cyber-attacks. Additionally, trust in cybersecurity may cause individuals not to engage in cybersecurity compliance behaviors again since it is not of their concern.

H1: The dark side of trust in cybersecurity has a negative (i.e., de-intensifies) relationship with a) attention to cybersecurity and b) compliance in protection.

Cybersecurity mindfulness, suspicion, and an individual's intention to protect all may increase the level of attention an individual directs at cybersecurity. Cybersecurity mindfulness and suspicion have both been shown to increase the identification of cyber-attacks. Jensen et al. [21] showed that training individuals in phishing mindfulness increases their ability to avoid falling victim to phishing attacks. Gay et al. [14] found increasing suspicion in military drone vehicles lead to an improvement in identifying and mitigating cyber-attacks on the systems running the drones. As noted previously, each of these constructs includes cognitive activation that engages the user, mindfulness through multiple perspectives in training compliance and suspicion through cognitively assessing a harmful situation for alternative explanations for how, when, and why a cyber-attack is happening thus eliciting vulnerability. Intention to protect has also been shown in IS research to be a positive predictor of behavior. We therefore posit that cybersecurity mindfulness, suspicion, and intention to protect will negatively moderate the relationship between dark side trust and attention to cybersecurity.

H2a: Cybersecurity mindfulness negatively moderates (i.e., de-intensifies) the relationship between the dark side of trust in cybersecurity and attention to cybersecurity.

H2b: Cybersecurity suspicion negatively moderates (i.e., de-intensifies) the relationship between dark side of trust in cybersecurity and attention to cybersecurity.

H2c: Intention to protect negatively moderates (i.e, de-intensifies) the relationship between the dark side of trust in cybersecurity and attention to cybersecurity.

Getting members in an organization to pay attention to cybersecurity is a concern for ISec researchers and practice. Vance et al. [43] showed that individuals over time stop paying attention over time to information security warnings as they become habituated to them. By changing design principles in the information security warning they showed that individuals were more attentive to the warning. Practice has also shown that although cybersecurity is a concern and widely prevalent in organizations a large number of people do not understand basic concepts [51].

Garnering individuals' attention toward cybersecurity is therefore important in eliciting compliance in protection. For instance, those that pay attention to phishing emails may identify them and subsequently forward them to the cybersecurity function or flag them within the system. Although, those that do not pay attention may not engage in requested compliance behaviors such as appropriately reporting a phishing email. This is important to ISec research as newer streams have shown the importance of collective human efforts in defending and diffusing cyber-attacks [20]. We therefore, propose that when individuals direct attention to cybersecurity they are more likely to engage in requested compliance behaviors.

H3: Attention to cybersecurity has a positive (i.e., intensifies) relationship with compliance in protection.

Table 2. Construct definitions

Construct	Definition
Trust in Cybersecurity	The degree to which an individual believes he or she is safeguarded from a cyber-attack [30]
Cybersecurity Mindfulness	A state of alertness and lively awareness in assessing threats to sensitive and confidential information [41]
Cybersecurity Suspicion	A person's simultaneous state of cognitive activity, uncertainty, and perceived malintent about cybersecurity [6]
Intention to Protect	The user's intention to continue protecting confidential information [46]
Attention to Cybersecurity	The process of directing our awareness to relevant stimuli while ignoring irrelevant stimuli in cybersecurity
Compliance in Protection	Behaviors that follow cybersecurity compliance policies for protecting confidential information and computing resources

7. Proposed Method

As this is research-in-progress, data will be collected in the near future in partnership with an organization in the southeastern United States. The research design for the full study will use a multi-method approach comprised of a survey, eye-tracking

study and a field experiment to triangulate data in the context of a phishing e-mail attack.

In partnership with our research site, training on phishing was provided to individuals in Fall of 2019, which included six heuristics or rules for individual to follow in identifying phishing emails: 1) ensuring the email header is legitimate 2) reviewing the domain name for legitimacy 3) reviewing for generic salutations 4) reviewing for spelling errors 5) looking for urgent requests for sensitive information, and 6) hovering and ensuring an embedded link or attached file is legitimate [35]. The organization digitally tracks all users that complete and do not complete the online phishing training.

The survey is being developed through a literature review of the trust, mindfulness, suspicion, attention, and compliance literatures. When available, scales will be adapted to the context of trust in cybersecurity [29], mindfulness [41], suspicion [6], and attention [42]. Compliance behaviors will be operationalized with a scale with items such as "I report suspected phishing emails to the cybersecurity department" and "I do not click on links unless they are deemed safe (i.e., the link is noted as safe after hovering over it.)".

To test the model, we will survey a sample of those individuals requested to complete the training (i.e., those that did and did not). We will first measure trust in cybersecurity, cybersecurity mindfulness, cybersecurity suspicion, and intention to protect. Items for attention and compliance will be operationalized to the above noted generalized phishing rules and subsequent compliance behaviors. Validity and reliability will be verified following established IS guidelines [28].

We will then recruit subjects for a follow up study from those that completed the survey and those that did not. Eye-tracking technology will be used to corroborate survey responses that individuals are in fact paying attention to phishing emails during the course of an experiment. The experiment will request individuals to identify potential phishing emails. Eye movements will be recorded to verify what an individual looks at when attempting to identify an email as malicious or legitimate.

Finally, we will conduct a simulated phishing exercise to enact a breach at an organization with simulated phishing software. This software provides tracking mechanisms to obtain objective data on end user compliance behaviors. Data will then be triangulated from each phase of the research design.

7.1 Preliminary Pilot Test Findings

Since this is a research-in-progress study the full study has yet to be finalized and started. However, to test the preliminary instrument and research model, an initial pilot test was conducted with 87 respondents from a university in the South Eastern United States. The research design for the pilot test consisted of administering the pilot instrument to participants. Participants were first presented with items measuring trust in cybersecurity, cybersecurity mindfulness, cybersecurity suspicion, and intention to protect. Participants were then presented with the above noted generalized phishing rules and shown the application of the rules to example phishing emails. Subjects were then presented with items regarding attention and compliance that were operationalized to the six general phishing heuristics. The results of the initial pilot test were promising and provided some initial support for the factor structure and model. At the time of submission, the research team was working on finalizing the instrument based on the pilot test results.

8. Analysis

Analysis will be done in multiple steps to assess the validity of our research design. We will use structural equation modeling (SEM), EQS 6.4 a covariance-based SEM, for data analysis as the research model has latent variables. We will examine the measurement model to verify the instrument yields reliable and valid measurements. We will then assess the structural model to test the effects of dark side trust on attention to cybersecurity and compliance for protection. Additionally, we will assess the moderating effects of cybersecurity mindfulness, suspicion, and intention to protect between dark side trust in cybersecurity and attention to cybersecurity. Objective data obtained from eye-tracking and the field experiment will also be analyzed.

9. Discussion

9.1 Theoretical Implications

Upon completion of this research-in-progress, this study contributes to cybersecurity and the ISec literature in several ways. While trust may be a key enabler in the use of cybersecurity systems, research has shown that dark side trust may result in negative consequences [12, 13]. By studying this concept in the context of cybersecurity we extend understanding of trust and its adverse impact on individuals and organizations to the ISec literature. Specifically, this study sheds light on how organizations can reduce mindless compliance due to dark side trust. The integrated model of dark side trust provides an explanatory framework for ISec literature to study the

degradation of compliance and combatting of dark side trust.

We introduce mindfulness and suspicion into our model as a means to provide actionable countermeasures for dark side trust. Of particular interest is gaining insight into the construct of suspicion [6, 14], which in this study begins to understand its role in cyber-defense and compliance. We provide understanding into the moderating role of mindfulness, suspicion, and intention to protect and seek to understand how they affect attention to cybersecurity and subsequent compliance behaviors.

Although, the prior empirical findings show the importance of attention, little work has extended our understanding of its role in compliance beyond information security warnings [1, 43, 45]. This research will provide further understanding of the interplay of trust and attention in cybersecurity as mechanisms to control compliance behavior.

We also extend current research in cybersecurity regarding trust, as we reconceptualize it as a superordinate multi-dimensional construct comprised of trust in people, technology, and the organization. This new conceptualization will capture the different nature of trust in cybersecurity influenced by technology, people, and processes. New dimensions of trust in a specific technology are also conceptualized due to the unique context of cybersecurity.

Finally, IS research has been almost exclusively about the benefits of trust for organizations, individuals, online transactions and technology [2, 4, 15, 30, 33, 40]. The existing body of trust and information systems (IS) research has not extensively explored the dark side of trust and its negative consequences [34]. The dark side trust model in this study may provide researchers outside of the ISec community to study it in other contexts such as e-commerce, virtual teams, and outsourcing relationships.

9.2 Practical Implications

This research-in-progress paper has several potential implications for practices ranging from training, design, and management. Training currently has been emphasized as a solution to increasing information security awareness. This paper conceptually proposes the idea that the effects of training, SETA, and notifications can be decreased since it may encourage trust. Designs of cybersecurity systems typically take a consistent iterative approach in issuing warnings, notifications, and calls to action and do not focus on failures of the function, system, or organization. This paper may provide the impetus to implement more malleable designs that engage the user to process decisions more mindfully through the levers of mindfulness, suspicion, and intention to protect. It may also encourage CISOs and cybersecurity managers to leverage the role of failure in training and communications (e.g., showing that cybersecurity function, system, and organization cannot prevent all cyber-attacks) as a cybersecurity countermeasure to combat dark side trust. Lastly, this paper shows that trust must be protected by the cybersecurity department and ensuring that individuals do not rely too heavily on the function, system, and organization for protection.

10. References

[1] Anderson, B., A. Vance, C.B. Kirwan, D. Eargle, and J.L. Jenkins, "How users perceive and respond to security messages: a NeuroIS research agenda and empirical study", *European Journal of Information Systems 25*(4), 2016, pp. 364–390.

[2] Ba, S., and P.A. Pavlou, "Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior", *MIS quarterly*, 2002, pp. 243–268.

[3] Bapna, R., A. Gupta, S. Rice, and A. Sundararajan, "Trust and the Strength of Ties in Online Social Networks: An Exploratory Field Experiment.", *MIS Quarterly 41*(1), 2017, pp. 115–130.

[4] Benbasat, I., and W. Wang, "Trust in and adoption of online recommendation agents", *Journal of the association for information systems 6*(3), 2005, pp. 4.

[5] Bobko, P., A. Barelka, and L. Hirshfield, *The Construct of State-Level Suspicion A Model and Research Agenda for Automated and Information Technology (IT) Contexts:*, Defense Technical Information Center, Fort Belvoir, VA, 2013.

[6] Bobko, P., A.J. Barelka, L.M. Hirshfield, and J.B. Lyons, "Invited Article: The Construct of Suspicion and How It Can Benefit Theories and Models in Organizational Science", *Journal of Business and Psychology 29*(3), 2014, pp. 335–342.

[7] Boss, S., D. Galletta, P.B. Lowry, G.D. Moody, and P. Polak, "What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors", *MIS Quarterly (MISQ) 39*(4), 2015, pp. 837–864.

[8] Crossler, R.E., A.C. Johnston, P.B. Lowry, Q. Hu, M. Warkentin, and R. Baskerville, "Future directions for behavioral information security research", *computers & security 32*, 2013, pp. 90–101.

[9] Cyr, D., M. Head, H. Larios, and B. Pan, "Exploring human images in website design: a multi-method approach", *MIS quarterly*, 2009, pp. 539–566.

[10] D'Arcy, J., A. Hovav, and D. Galletta, "User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach", *Information Systems Research 20*(1), 2009, pp. 79–98.
[11] Dasgupta, P., "Trust as a commodity. Trust: Making and Breaking cooperative relations. D. Gambetta", 1988.
[12] Gambetta, D., "Can we trust trust", *Trust: Making and breaking cooperative relations 13*, 2000, pp. 213–237.
[13] Gargiulo, M., and G. Ertug, "The dark side of trust1", *Handbook of trust research 165*, 2006.
[14] Gay, C., B. Horowitz, J. Elshaw, P. Bobko, and I. Kim, "Operator Suspicion and Decision Responses to Cyber-Attacks on Unmanned Ground Vehicle Systems", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 61*(1), 2017, pp. 226–230.
[15] Gefen, D., E. Karahanna, and D.W. Straub, "Trust and TAM in online shopping: an integrated model", *MIS quarterly 27*(1), 2003, pp. 51–90.
[16] Herath, T., and H.R. Rao, "Protection motivation and deterrence: a framework for security policy compliance in organisations", *European Journal of Information Systems 18*(2), 2009, pp. 106–125.
[17] Herath, T., and H.R. Rao, "Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness", *Decision Support Systems 47*(2), 2009, pp. 154–165.
[18] Hosmer, L.T., "Trust: The connecting link between organizational theory and philosophical ethics", *Academy of management Review 20*(2), 1995, pp. 379–403.
[19] Jakobsson, M., and S. Myers, *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*, John Wiley & Sons, 2006.
[20] Jensen, M., A. Durcikova, and R. Wright, "Combating Phishing Attacks: A Knowledge Management Approach", (2017).
[21] Jensen, M.L., M. Dinger, R.T. Wright, and J.B. Thatcher, "Training to mitigate phishing attacks using mindfulness techniques", *Journal of Management Information Systems 34*(2), 2017, pp. 597–626.
[22] Johnston, A.C., M. Warkentin, M. Siponen, Mississippi State University, M. Siponen, and University of Jyväskylä, "An Enhanced Fear Appeal Rhetorical Framework: Leveraging Threats to the Human Asset Through Sanctioning Rhetoric", *MIS Quarterly 39*(1), 2015, pp. 113–134.
[23] Johnston, and Warkentin, "Fear Appeals and Information Security Behaviors: An Empirical Study", *MIS Quarterly 34*(3), 2010, pp. 549.
[24] Keil, M., "Escalation of commitment in information systems development: A comparison of three theories.", Academy of Management Briarcliff Manor, NY 10510 (1995), 348–352.
[25] Keil, M., B.C. Tan, K.-K. Wei, T. Saarinen, V. Tuunainen, and A. Wassenaar, "A cross-cultural study on escalation of commitment behavior in software projects", *MIS quarterly*, 2000, pp. 299–325.
[26] Langfred, C.W., "Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams", *Academy of management journal 47*(3), 2004, pp. 385–399.
[27] Lee, J.D., and K.A. See, "Trust in Automation: Designing for Appropriate Reliance", *Human Factors*, 2004, pp. 31.
[28] MacKenzie, Podsakoff, and Podsakoff, "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques", *MIS Quarterly 35*(2), 2011, pp. 293.
[29] Mcknight, D.H., M. Carter, J.B. Thatcher, and P.F. Clay, "Trust in a specific technology: An investigation of its components and measures", *ACM Transactions on Management Information Systems 2*(2), 2011, pp. 1–25.
[30] McKnight, D.H., V. Choudhury, and C. Kacmar, "Developing and validating trust measures for e-commerce: An integrative typology", *Information systems research 13*(3), 2002, pp. 334–359.
[31] Pasqualetti, F., F. Dorfler, and F. Bullo, "Attack Detection and Identification in Cyber-Physical Systems", *IEEE Transactions on Automatic Control 58*(11), 2013, pp. 2715–2729.
[32] Patzelt, H., and D.A. Shepherd, "The decision to persist with underperforming alliances: The role of trust and control", *Journal of management studies 45*(7), 2008, pp. 1217–1243.
[33] Pavlou, P.A., H. Liang, and Y. Xue, "Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective", *MIS quarterly*, 2007, pp. 105–136.
[34] Pienta, D., H. Sun, and J. Thatcher, "Habitual and Misplaced Trust: The Role of the Dark Side of Trust Between Individual Users and Cybersecurity Systems", 2016.
[35] Pienta, D., J.B. Thatcher, and A.C. Johnston, "A Taxonomy of Phishing: Attack Types Spanning Economic, Temporal, Breadth, and Target Boundaries", 2018, pp. 18.
[36] Schuetz, S.W., P.B. Lowry, and J.B. Thatcher, "DEFENDING AGAINST SPEAR PHISHING: MOTIVATING USERS THROUGH FEAR APPEAL MANIPULATIONS", (2016), 12.
[37] Skinner, D., G. Dietz, and A. Weibel, "The dark side of trust: When trust becomes a 'poisoned chalice'", *Organization 21*(2), 2014, pp. 206–224.

[38] Sollner, M., I. Benbasat, D. Gefen, J.M. Leimeister, and P.A. Pavlou, "'Trust', in MIS Quarterly Research Curations, Ashley Bush and Arun Rai, Eds," *MIS Quarterly*, 2016.
[39] Straub, D.W., "Effective IS Security: An Empirical Study", *Information Systems Research 1*(3), 1990, pp. 255–276.
[40] Sun, H., "Sellers' trust and continued use of online marketplaces", *Journal of the Association for Information Systems 11*(4), 2010, pp. 2.
[41] Thatcher, J.B., R.T. Wright, H. Sun, T.J. Zagenczyk, and R. Klein, "Mindfulness in information technology use: definitions, distinctions, and a new measure", *MIS Quarterly 42*(3), 2018, pp. 831–847.
[42] Thøgersen, J., "Psychological Determinants of Paying Attention to Eco-Labels in Purchase Decisions: Model Development and Multinational Validation", *Journal of Consumer Policy 23*(3), 2000, pp. 285–313.
[43] Vance, A., J.L. Jenkins, B.B. Anderson, et al., "Tuning Out Security Warnings: A Longitudinal Examination of Habituation Through fMRI, Eye Tracking, and Field Experiments", *MIS Quarterly 42*(2), 2018, pp. 355–380.
[44] Vance, A., P.B. Lowry, and D. Eggett, "Using accountability to reduce access policy violations in information systems", *Journal of Management Information Systems 29*(4), 2013, pp. 263–290.
[45] Vance, A., P.B. Lowry, and D.L. Eggett, "Increasing accountability through the user interface design artifacts: A new approach to addressing the problem of access-policy violations", *Mis Quarterly 39*(2), 2015, pp. 345–366.
[46] Venkatesh, Morris, Davis, and Davis, "User Acceptance of Information Technology: Toward a Unified View", *MIS Quarterly 27*(3), 2003, pp. 425.
[47] Venkatesh, V., J.Y. Thong, F.K. Chan, P.J. Hu, and S.A. Brown, "Extending the two-stage information systems continuance model: Incorporating UTAUT predictors and the role of context", *Information Systems Journal 21*(6), 2011, pp. 527–555.
[48] Wright, R.T., M.L. Jensen, J.B. Thatcher, M. Dinger, and K. Marett, "Research Note—Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance", *Information Systems Research 25*(2), 2014, pp. 385–400.
[49] Wright, R.T., and K. Marett, "The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived", *Journal of Management Information Systems 27*(1), 2010, pp. 273–303.
[50] Yip, J.A., and M.E. Schweitzer, "Trust promotes unethical behavior: Excessive trust, opportunistic exploitation, and strategic exploitation", *Current Opinion in Psychology 6*, 2015, pp. 216–220.
[51] "What Americans Knows About Cybersecurity | Pew Research Center", 2017. https://www.pewinternet.org/2017/03/22/what-the-public-knows-about-cybersecurity/
[52] "2019 Data Breach Investigations Report", *Verizon Enterprise*. https://enterprise.verizon.com/resources/reports/dbir/
[53] "Equifax Data Breach, 1 Year Later: Obvious Errors, No Fixes | Fortune", http://fortune.com/2018/09/07/equifax-data-breach-one-year-anniversary/