Do SETA Interventions Change Security Behavior? – A Literature Review

Uchechukwu Nwachukwu
University of Turku
ucjunw@utu.fi

Jiri Vidgren
University of Jyvaskyla
jiri.e.vidgren@student.jyu.fi

Marko Niemimaa
University of Agder
marko.niemimaa@uia.no

Jonna Järveläinen
University of Turku
jonna.jarvelainen@utu.fi

Abstract

Information security education, training, and awareness (SETA) are approaches to changing end-users' security behavior. Research into SETA has conducted interventions to study the effects of SETA on security behavior. However, we lack aggregated knowledge on 'how do SETA interventions influence security behavior?'. This study reviews 21 empirical SETA intervention studies published across the top IS journals. The theoretical findings show that the research has extended Protection Motivation Theory by (1) enhancements to fear appeals; (2) drawing attention to relevance; (3) incorporating temporality; (4) and shifting from intentions to behavior. In terms of behavior, the SETA interventions have targeted (1) information security policy compliance behavior; and (2) information protection behavior. We argue that while these studies have provided insights into security intentions and behavior, knowledge on designing effective SETA training has remained primarily anecdotal. We contribute (1) by pointing out gaps in the knowledge; and (2) by proposing tentative design recommendations.

Keywords: Awareness, training, SETA, security management, security behavior

1. Introduction

Security education, training, and awareness (SETA) is part of the portfolio of security management approaches to changing users' Information Systems Security (ISS) behavior (Kretzer & Mädche, 2015). These management approaches seek to impose change on users' intentions or behavior through different methods, such as classroom teaching, mobile applications, or posters, that train and educate users on how to behave securely when processing organizations' information, i.e., approaches aimed to change ISS behavior. These approaches seek to influence the behavior persistently and collectively, typically exemplified and conceptualized as the security culture. It is thus no wonder that collaborative group trainings have been suggested as the most recommendable (Karjalainen & Siponen, 2011). Literature documents various and diverse methods and theories to implement SETA, such as experiments from games (Dincelli & Chengalur-Smith, 2020) and scenarios (Tsohou et al., 2015) to mindfulness interventions (Jensen et al., 2017). In addition, literature reviews have studied the aptness of serious games for cybersecurity training (Hendrix et al., 2016), theories explaining information security behavior (Lebek, Uffen, Neumann, Hohler, & Breitner, 2014), cybersecurity training in the critical infrastructure area (Chowdhury & Gkioulos, 2021), and the effectiveness of social engineering training (Aldawood & Skinner, 2019). SETA is an interventionist approach, as it seeks to actively change users' intentions or behavior using methods to achieve long-term effects in the form of a security culture. We define SETA intervention as *any experimental or naturalistic manipulation aiming to change an individual's security behavior.*

While other studies have reviewed related aspects (Lebek, Uffen, Neumann, Hohler, & H. Breitner, 2014), regardless of the significance of SETA approaches for the management of security and organizational culture, it is surprising that we lack reviews on SETA *interventions*. A review of used theories, training delivery methods and aimed behavioral changes would facilitate planning of SETA interventions. To address this gap, we analyze past literature to study how SETA interventions influence security behavior. We explore this question through a systematic literature review to find papers focusing on security education, training, or awareness of end-users, both employees and private persons. In particular, we focus on the interventions by reviewing what theories are used to explain the interventions, what methods are used to induce the behavioral change, and what behaviors these interventions seek to change.

2. Prior Research and Research Approach

To ensure rigor in the systematic literature review process, we followed the vom Brocke et al. (2015) guidelines for searching and reviewing the literature.

URI: https://hdl.handle.net/10125/103396
978-0-9981331-6-4

First, we reviewed central IS journals for this scope, AIS conferences and HICSS, which resulted in 1086 papers. Thus, we decided to limit the screening of papers to a representative sample of journals. We focused on the Senior Scholars' Basket of Eight with the addition of articles from the Computers & Security journal to ensure relevance, impact, and theoretical and practical rigor (Levy & Ellis, 2006). Further, we expected these journals to contain the most mature knowledge on SETA to be able to draw conclusions on their effectiveness. The literature that fits the scope and requirements of this review covers 21 articles published between 2010 and 2021 (available upon request).

We uncovered several other literature reviews that have focused on security compliance and awareness theories (Lebek, Uffen, Neumann, Hohler, & Breitner, 2014), success factors of SETA (Kirova & Baumöl, 2018), inter-organizational information security (Karlsson et al., 2016), human factors (Glaspie & Karwowski, 2018), behavioral influencers (Alohali et al., 2018), and personal information awareness (Ögütçü et al., 2016).

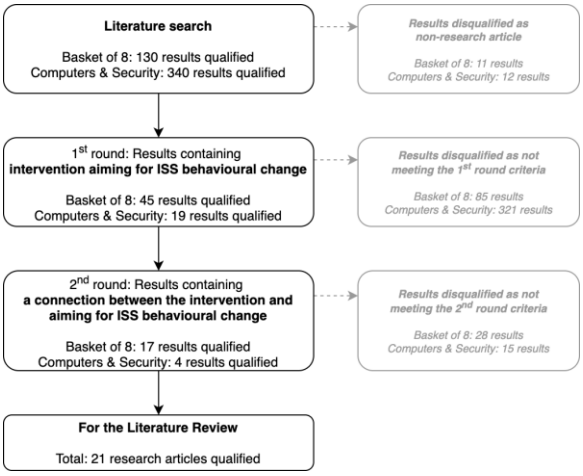

Figure 1. Literature search process.

After selecting the scope, we targeted a specific search to all information about the articles with the search word combination ("security awareness "OR "security training "OR "security education "). An initial search of the publications returned 470 results from the journals. We discarded twenty-three articles that were editorial issues, prefaces, endnotes, and executive overviews.

The research articles from the search result were then qualified by reading the abstract of each article and looking for the intervention that aims to change the ISS behavior of users. After this first round of qualifications, we selected 64 research articles for further examination and analysis. During the second round, we read the research articles entirely and qualified the ones which included the connection between the SETA intervention and the ISS behavior change.

The selected papers were read through, and from their contents, a matrix of concepts was created containing details about the interventions used. The matrix of concepts encompassed intervention goals, context, type of intervention, behavior targeted, target population, outcomes, intervention mediums, theories, and research methods used. Initial buckets or themes of interventions were identified and refined by the second and third read-through. The synthesis of these commonalities is presented in the subsequent sections of this article.

Quantitative research was the prevalent research method used in these articles, in a total of 15 papers. The rest were based on Mixed Methods (3), Design Science Research (2), and Action Research (1). Most (16) of the delivery mediums of the SETA training had an experimental approach. Thirteen were conducted in the field and three in the laboratory environment. Other delivery mediums include scenarios (3) and games (2).

We structure the findings according to a framework of the trifecta of SETA interventions with three components: theory, training methods, and behavior. SETA interventions research should be 1) founded on the theory that can either inform the design of the training methods and/or provide an explanation/prediction on behavior change, 2) make use of some training method to deliver the intervention, and 3) aim to change a specific behavior. That is, as SETA *interventions*, the research should provide researchers and practitioners prescriptions in the form of design and action theories, rather than explanations and/or predictions (Gregor, 2006).

The trifecta and the key components of SETA intervention research are illustrated in Figure 2.

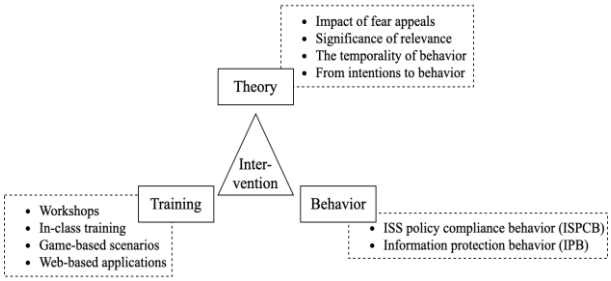

Figure 2. Trifecta of SETA intervention studies.

3. Theories in SETA interventions

The literature review revealed that a large and diverse body of theories had been used in the SETA intervention studies. Despite the diversity, the protection motivation theory (PMT) is most prominent (used in almost half of the studies). This chapter will first provide an overview of the diverse theory-base and

then elaborate on the theoretical learnings for PMT gained through the intervention studies.

3.1 Theories of SETA interventions

The theory base for explaining behavioral change through SETA is very diverse. For example, while Protection Motivation Theory (PMT) is the most widely used theory in our sample, we found a total of 23 different theories used. In addition, some of the research builds on several theories. Table 1 summarizes the most common theories in SETA research, i.e. those used more than once in our sample.

Table 1. Theories in SETA interventions.

Theory	Brief description
Protection Motivation Theory (PMT) (7 articles)	Explains an individual's protection behavior through threat and coping appraisals.
Deterrence Theory (DT) (3 articles)	Explains the effects of deterrents to an individual's behavior to deter harmful behavior
Elaboration Likelihood Model (ELM) (2 articles)	Explains how individuals process persuasive information through either central or peripheral routes to change attitudes.

Next, we will focus on the theoretical insights developed through the PMT-based interventions but omit discussion on other theories due to the limited prior research and space constraints.

3.2 Protection Motivation Theory

The origins of PMT relate to how people react to stressful situations but have since become widely applied across different fields to explain various intentions and behavior. In short, PMT "theorizes that when an individual is confronted with a threat, he or she cognitively assesses the threat and a possible associated remedy" (Menard et al., 2017).

The popularity of PMT in SETA intervention studies is intuitive as it is also the most common theory in non-interventionist security behavior studies (Kirova & Baumöl, 2018; Lebek, Uffen, Neumann, Hohler, & Breitner, 2014). PMT enables studying differences or changes in the participants' attitudes or actual behavior in the SETA intervention studies. PMT cannot directly explain why a particular SETA approach works, nor can it provide prescriptions on how a SETA should be delivered through training methods. PMT enables researchers to study the differences either between respondents that are manipulated differently (e.g., experiment group(s) and control group(s)), differences between pre and post-intervention intentions or behavior of the respondents (Puhakainen & Siponen, 2010) or intentions, or behavioral changes over different points in time.

Liang and Xue (2009) initially proposed the technology Threat Avoidance Theory (TTAT) to adapt the PMT to IS security context. Due to its origins in PMT, we categorized TTAT studies in the same category as PMT studies. TTAT presents a process model in which secure behavior represents a coping response to the recognized information security threat.

In this review, the body of literature is too limited, and applications of the theory are too diverse to make any far-reaching conclusions or generalizations. However, the reviewed studies can still indicate what has been learned about effective SETA interventions through the theories. Next, we review these learnings.

Appealing to fear. Fear appeals (i.e., messages intended to evoke fear in the recipients to engage in protection behavior) have been widely used as a theoretical construct in PMT intervention studies. The fear appeals provide a way to manipulate participants' behavior, especially in experimental settings where researchers can control the effects of differently constructed fear messages. Boss et al. (2015) argued that while fear appeals are a core construct of PMT, most studies have omitted them. The authors used strong and weak fear appeals to address the lack of fear appeal manipulations. They argue that strong fear appeals produce more fear and supporting threat, inspiring protection motivation. Johnston and Warkentin (2010) extended the PMT with the fear appeals model (perceived threat severity, perceived threat susceptibility). They found that self-efficacy, response efficacy, threat severity, and social influence, in part, determine behavioral intentions.

Similarly, Johnston et al. (2015) argue that the fear appeals rhetoric has been misspecified in the extant IS research and propose a comprehensive fear appeal rhetorical framework. Their research extends the PMT by introducing sanctions to fear appeals, thereby integrating personal relevance to the theory, which correlates with positive security behavior. On a more general note, progressively providing multiple explicit fear appeal messages, emphasizing the frequency of occurrence and potential harms with concrete examples, results in increased fear and behavioral intention (Boss et al., 2015; Johnston & Warkentin, 2010). While fear appeals can motivate users to secure behavior, Silic and Lowry (2020) note that most users prefer working in a fun and supportive environment rather than being chained to rules and fearing punishment. This is also significant given that highly frustrated users are less

likely to comply with security policies (Ormond et al., 2019).

Significance of relevance. Studies have found the message's relevance essential when manipulating intentions or behavior. In practice, relevance can be achieved by incorporating personal experiences or documents into the training and modeling the training tasks to produce a relatable cause-and-effect mental model (Puhakainen & Siponen, 2010). Schuetz et al. (2020) introduced contextual relevance and abstractness of fear appeal messages. According to the authors, while context has often been called out to explain inconsistent findings across studies on message relevance, the influence of context on PMT appraisal variables has not been articulated before.

In contrast to the extant research, the authors showed that in the organizational context, users reacted more strongly to fear appeals than in the personal context. They state: "our findings suggest that organizational users respond with lower perceptions of self-efficacy than personal users but higher perceptions of response costs, fear, and protection motivation." (p. 746). In addition, the authors found that message abstractness can explain the variance across studies and consistently influences intentions and behavior.

Incorporating relevance-related variables to PMT has been found significant for SETA and not merely on fear appeals. Jaeger and Eckhardt (2021) studied the influence of situational information security awareness on protection motivation. Their findings contribute to PMT by showing how situation awareness "serves as an initiator of the cognitive mediating processes that are the focus of PMT, meaning that situational information security awareness can act as a source of information needed for threat and coping appraisal "(p. 448). Further, Abraham et al. (2019) extend PMT "by proposing learner-controlled ISec [SETA] training as an antecedent of self-efficacy and threat severity and susceptibility perceptions "(p. 8). By learner-controlled, the authors refer to training tailored for the learner's learning style, abilities, and knowledge, thereby increasing the personal relevance of the provided training. The learner-controlled training was found to have a "positive effect on training satisfaction, training performance, self-efficacy, and threat severity and a marginally negative impact on threat susceptibility." (p. 8). Further, their study shows how learner-controlled training increases training retention, thereby drawing attention to the temporal aspects of security behavior.

The temporal considerations. In comparison, Abraham et al. (2019) studied the cognitive responses of the trainees over time (immediately after and two weeks later) as measures of training retention, while Steinbart et al. (2016) studied the security continuance behavior. The authors build on TTAT rather than PMT and focus on the cybernetic loop (i.e., input, process, output with a feedback loop, which they argue has been an understudied aspect of the theory. By incorporating the cybernetic loop into PMT studies, the authors can study how people respond to a threat and how the IT artifact affects their behavior. The authors argue that two independent goals influence the continuance of secure behaviors after manipulations. These are the user's desire for security (as predicted by TTAT) and the user's desirability for usability (as explained by TTAT and enhanced PMT models).

Behavior rather than intentions. Despite the initial focus of the theory on intentions, PMT has also been extended to study actual behavior. Furthermore, while the focus on intentions has been a trait of PMT studies, Jenkins et al. (2021) argued that the intention/behavior gap is, in fact, prevalent across security behavior studies. By studying the intentions and actual behavior simultaneously, the authors showed how the required effort to follow security policies did not significantly influence behavior directly but negatively influenced intentions on behavior.

4. Training Delivery Methods

The interventions use various training delivery mediums, primarily workshops, game-based scenarios, and web-based applications. Most of the interventions were initiated with some form of in-class or video training and most combined several delivery methods. Next, we discuss each of these training delivery methods.

4.1 In-class trainings

In-class, instructor-led trainings are among the most typical training delivery methods. While it is typical that these trainings include what Karjalainen and Siponen (2011) refer to as transmission-oriented training, they can consist of more participatory and transformational approaches. Puhakainen and Siponen (2010) divided the training into several parts. The first part was a generic lecture on security's significance, followed by hands-on training on encryption software.

In-class trainings can be combined with other methods. Wright and Marett (2010) sought to improve students' phishing awareness by giving each participant a code they purposefully referred to as "super-secure code" (SSC) to emphasize the importance of the piece of information they were given. The students were then given generic information security training during their lectures. During these lectures, "the instructor taught the concepts of phishing, hacking, and other relevant security/privacy topics" (p. 284). But in addition, the authors used other means to emphasize the importance

of the SSC, such as handing the information in a sealed envelope. Rather strikingly, despite all the efforts, 32% of the participants fell for a phishing email the authors had crafted and disclosed the SSC, which testifies to the difficulty of effective SETA.

Researchers have also found it helpful to separate participants into different groups during training based on their knowledge levels (Puhakainen & Siponen, 2010; Tsohou et al., 2015). Grouping allows tailoring the training to each user group in a contextually relevant manner. For example, Tsohou et al. (2015) showed that non-technical end-users found different topics helpful. A crucial aspect of in-class training success is incorporating active participation and interaction with reflections into the training sessions (Hart et al., 2020; Jensen et al., 2017; Puhakainen & Siponen, 2010).

4.2 Workshops

In contrast to in-class transmission-oriented methods, scholars have specifically approached SETA with interactive workshops. Albrechtsen and Hovden (2010) emphasized the interactive nature of the SETA workshops, which were organized around seven plausible scenarios the employees could face in their daily work and free time. The authors conclude that "the intervention was powerful enough to significantly change awareness and behavior among the participants in the intervention group" (p. 442). Tsohou et al. (2015) studied not only the implementation of individual SETA sessions but sought to embed awareness programs into case organization through workshops. While the exact details are not disclosed, the authors divided the users into three interactive workshops: two for non-technical users and one for IT personnel. The content in all three sessions focused on generic security topics such as an introduction to privacy and security for non-technical users and privacy-enhancing technologies for IT personnel. Indeed, while 40% of the non-technical users and 90% of the IT personnel found the workshops necessary, only 25% of the end-users found the workshops practical. Interestingly, despite the relatively low practicality, none saw the workshops as impractical.

4.3 Game-based methods

Game-based methods have several benefits over non-games-based delivery methods. The game-based methods include game elements, like stories and interactivity, that make the training more immersive and engaging (Silic & Lowry, 2020). Dincelli and Chengalur-Smith (2020) used visual/image-based and text-based games to compare the effects of the approaches. The authors found that visual game-based interventions incorporating stories and reflections were easier to learn and aided in improving recall and the ability to reconstruct threats from partial cues (redintegration). In contrast, text-based game interventions had a longer-lasting effect behaviorally. The games should be built on IT artifacts that users are already familiar with to reduce the cognitive burden such that the users can focus on the content of the SETA rather than on the IT artifact itself (Steinbart et al., 2016). Further, gamified interventions are more effective in motivating users than email awareness messages due to the more immersive experience (Silic and Lowry, (2020). Additionally, game-based methods can bring hedonistic benefits to the users (Ibid.).

4.4 Web-based applications

Several studies advocate the use of websites or web-based applications to deliver SETA. Abraham and Chengalur-Smith (2019) used a web application to develop a method for learner-controlled SETA. By implementing the SETA as an online portal, the users were given the possibility to control how they wish to learn the material (e.g., the pace and sequence of going through the material), which would not be possible in an in-class setting. Learner control is beneficial for delivering the training, but users' ability to skip material may have adverse outcomes as users may omit essential parts of the training.

Web-based applications can also function merely as a medium to deliver specific content rather than as a specific approach for training. For instance, Johnston et al. (2015) used a web application, but their focus was on how rhetoric in the content of the delivered message influences the users' behavioral intentions. It is unknown how critical web-based applications were for behavioral change or if the content had a more significant role. Nevertheless, what was learned from the study is that sanctioning rhetoric can enhance the effectiveness of fear appeals.

5. Behavioral change of SETA training

The literature analysis suggested that SETA interventions can be broadly categorized to aim for changes in two types of ISS behavior: ISS policy compliance behavior (ISPCB); and information protection behavior (IPB). Next, we provide a review of both of these categories with illustrations drawn from the literature.

5.1 Information Security Policy Compliance Behavior (ISPCB)

The ISPCB category contains eight papers aimed at changing the users' behavior to enforce policy compliance for protecting organizational resources. For example, this category included studies by Siponen et al. (2020), Siponen and Vance (2010), and Johnston et al. (2015). In these studies, password security trainings and USB-drive usage scenarios were used as vehicles to change users' compliance with information security policies. Password security, for example, was only used as an example of ISPCB. Some of the articles conducted post-tests and measured the actual behavioral change (Silic & Lowry, 2020; Siponen et al., 2020) while some articles performed an intervention but measured only intention to comply as an indicator of behavioral change (Siponen & Vance, 2010). On the other hand, some articles had no report attempting to measure post-intervention behavioral change (Tsohou et al., 2015).

In a thematically related manner, Siponen and Vance (2010) and Johnston et al. (2015) emphasized how informal sanctions such as guilt and self-blame predicted security policy compliance without neutralization techniques. Users become more motivated to comply with policies if they fear facing colleagues' disapproval, being teased, or being ostracized by colleagues than by formal sanctions. Thus, SETA approaches can leverage the potency of such severe informal sanctions by focusing on how policy non-compliance can lead to letting down colleagues or result in peer embarrassment (Johnston et al., 2015). Siponen and Vance (2010) suggested that neutralization techniques nullified the effect of formal and informal sanctions and recommended considering how to counter these techniques when designing information security training. One suggested way was creating dissonance between actual behavior and perceived recommended/best practices. Puhakainen and Siponen (2010) emphasized the importance of management in participating in information security policy drive and engaging in a visible manner that can motivate users.

5.2 Information Protection Behavior (IPB)

The IPB category consisted of 13 articles aiming to change users' susceptibility to phishing attacks and attacks to obtain users' private information. In addition, this category included the articles which considered the users' susceptibility to disclosing sensitive organization information, covering interventions aiming to improve users' password security, authentication methods, encryption for sensitive data, spyware, and general information security awareness.

The articles in the IPB category focused primarily on phishing exercises. Five of the thirteen articles were concerned mainly with changing the users' susceptibility to hand out information to phishing attacks (Goel et al., 2017; Jaeger & Eckhardt, 2021; Jensen et al., 2017; Schuetz et al., 2020; Wright & Marett, 2010). Furthermore, Jensen et al. (2019) attempted to change users' susceptibility to phishing and their behavior of using weak passwords. Target groups of the SETA interventions were primarily organizational employees and sometimes used universities as organizations. However, in a few cases, students acted as imaginary new hires (Abraham & Chengalur-Smith, 2019) or everyday citizens (Boss et al., 2015).

One reoccurring theme was the importance of contextual relevance. Some studies indicated that the alignment and personalization of phishing email content to the users' working context potentially raises the user's susceptibility (Jaeger & Eckhardt, 2021). This is mainly due to the need to focus on information needed to solve a work task and push aside other competing message cues creating cognitive overloads (Goel et al., 2017; Jaeger & Eckhardt, 2021). This susceptibility to contextually relevant phishing content can be mitigated by providing appropriate training content to related audience portions and supplementing such rule-based training with mindfulness approaches (Jaeger & Eckhardt, 2021; Jensen et al., 2017). Jensen et al. (2017) indicated that incorporating mindful techniques aids the transition from awareness to behavioral change where the mindful reflections of 'stop,' 'think,' and 'check' are taken before user actions.

The influence of self-efficacy was also discussed in papers. Jaeger & Eckhardt (2021) described the positive influence improving self-efficacy has on protection motivation and reducing susceptibility to attack vectors. When the awareness level of users is raised, they develop the confidence to act in response to a security situation and believe their actions will be effective. As a result, their protection motivation is increased. In contrast, Jensen et al. (2017, 2019) demonstrated that self-efficacy could negatively impact susceptibility when overestimated self-efficacy. Further, seeing a concrete and detailed description of threats with relatable examples facilitates a better understanding of the threats for users, which results in increased fear and protection motivation, more so for users in organizational settings (Schuetz et al., 2020).

Seven articles focused on end-user training to adopt information protection practices or tools. The interventions in these articles included an experiment on password security (Steinbart et al., 2016), backing up essential data (Boss et al., 2015), use of anti-spyware (Boss et al., 2015; Johnston & Warkentin, 2010), self-

disclosure of information in social media (Dincelli & Chengalur-Smith, 2020), and generic information security awareness behaviors (Abraham & Chengalur-Smith, 2019; Albrechtsen & Hovden, 2010; Hart et al., 2020).

Our analysis showed differences between ISS behavior categories and used training delivery methods. SETA interventions aimed at changing ISPCB often consisted of task-based training methods like scenario tasks (e.g. Jenkins et al., 2021; Siponen & Vance, 2010), instructor-led trainings (e.g. Puhakainen & Siponen, 2010; Siponen et al., 2020; Tsohou et al., 2015), and web-based activities (Johnston et al., 2015; Silic & Lowry, 2020). In contrast, SETA interventions with the aiming to change IPB, training delivery mediums range from email and text-based communications (e.g. Jensen et al., 2017; Johnston & Warkentin, 2010), workshops and in-class trainings (e.g. Albrechtsen & Hovden, 2010; Wright & Marett, 2010), games (e.g. Hart et al., 2020; Steinbart et al., 2016), web-based tutorials (e.g. Abraham & Chengalur-Smith, 2019; Goel et al., 2017) to visual trainings (e.g. Jensen et al., 2019; Schuetz et al., 2020) and comics (Dincelli & Chengalur-Smith, 2020).

6. Discussion

In this literature review, we studied 'how do SETA interventions influence behavior?' by focusing on the trifecta of SETA approaches: 1) the theories used, 2) the delivery methods, and 3) the behavior. Next, we elaborate on how the study's findings shed light on the research question.

First, the review uncovered numerous theories authors have used to explain and predict changes in ISS behavior. While PMT was the most prevalent, many other theories were also tested and developed. In particular, the review indicates that effective SETA trainings include fear appeals (Johnston et al., 2015), are relevant (Puhakainen & Siponen, 2010), account for temporality (Abraham & Chengalur-Smith, 2019), and focus on actual behavior rather than intentions (Jenkins et al., 2021).

Second, we found SETA can influence ISS behavior through several different methods. While in-class trainings, web applications, and workshops have all been used to influence the users' behavior successfully, our analysis suggests that game-based approaches can be particularly suitable for SETA due to their ability to engage and motivate users (Silic & Lowry, 2020). Workshops provide opportunities for interaction (Albrechtsen & Hovden, 2010; Puhakainen & Siponen, 2010) that might be difficult to implement in a game-based environment.

Third, we found SETA interventions to influence information security policy behavior (ISPCB) and information protection behavior (IPB). While we found a few differences in the training delivery methods used to influence behavior in these categories, the studies are too limited (in number) for conclusions. Further, the category of IPB is vast and includes research on influencing widely different ISS behavior (e.g., general awareness contra phishing detection). As such, SETA interventions seem to be used to influence a broad range of ISS behavior, but these boundaries and limitations of SETA need to be studied further.

Next, we will present the research gaps in the literature with a framework of a trifecta of SETA interventions and offer tentative design recommendations for SETA.

6.1 Literature Gaps

As interventions, we posit that SETA approaches should consist of three interrelated elements: theory to explain and predict ISS behavior; theory to guide the design and development of training delivery methods (e.g., by instantiating a theory); and training methods (objects created for or during training) to provide a rigorous way to instantiate changes in ISS behavior. By doing so, SETA research can develop theories of design and action that can provide prescriptions for interventions (Gregor, 2006).The review showed that PMT had been used to establish correlations between intervention and users' intentions/behavior.

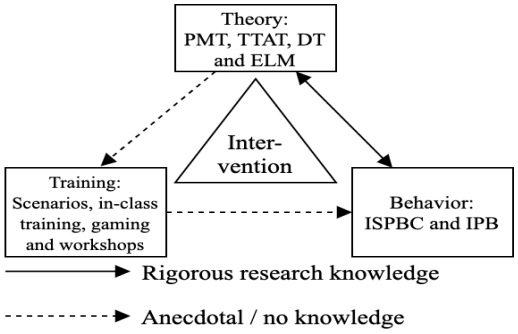

Figure 3 SETA Intervention Trifecta and research gaps

Our research indicates that less research has sought to establish a theory/artifact relationship (Karjalainen & Siponen, 2011) or artifact behavior relationship (Albrechtsen & Hovden, 2010; Puhakainen & Siponen, 2010). We illustrate this in Figure 3 and elaborate on it next.

Theory ←→ Security Behavior: The existing research has demonstrated changes in security behavior through interventions. Despite that PMT prevails, a

wide range of theories have been offered to explain the measured differences. These studies have established strong, empirically based evidence on factors that indicate intentions and secure behavior (e.g., fear appeals strongly to antecede protection-motivated behavior). Further, the studies have shown how theories can explain (in)secure behavior (e.g., what factors contribute to users' rationalization of failures to protect organizational information). While more studies are needed to provide further evidence and confirm earlier results, the review indicates that more significant gaps and room for more novel contributions exist in the other parts of the trifecta. Based on the review, the theories explain a broad range of ISS behavior. A more systematic analysis would be needed to study which theories can explain/predict more specific behavior. This could be achieved with the help of standardized taxonomies for ISS behavior. We have identified two types of behavior (ISPCB and IPB), but more detailed taxonomies of ISS behavior could be used to understand which theories explain different types of behavior.

Theory ←→ Training delivery methods: While the literature review uncovered one study that used design science research (Dincelli & Chengalur-Smith, 2020), there is a shortage of studies addressing the theory-based design and implementation of SETA approaches. Our analysis shows that the literature documents several different training delivery methods and artifacts to instantiate behavioral change. However, these methods and the training artifacts (e.g., a mobile application) are only loosely connected to theory. The focus of research is on behavioral changes rather than on what design features or material aspects of the methods and artifacts engender those behavioral changes. As such, the training methods, and artifacts, as part of the explanation/prediction of specific ISS behavior, tend to fade into the background. Here, design science research-based approaches are likely fruitful as they can be used both to develop theory and instantiate theory as artifacts. We expect that building the methods and artifacts on theories can contribute to novel and rigorous SETA approaches and enable the development of an expanding body of design theories (Karjalainen & Siponen, 2011). This body of knowledge could result in design principles for SETA interventions and design and action theories (Gregor, 2006), of SETA interventions to be used and applied in practice.

Training methods and artifacts ←→ Security Behavior: While the literature review shows that various training methods have been used in the interventions ranging from comic strips to mobile applications to different in-class methods, the actual relationship between the training methods (and the related artifacts) and the changes in security behavior have been less studied. Puhakainen and Siponen (2010) have applied action research to conduct intervention through physical in-class training. Action research provides sound foundations for interventions that aim to improve a problematic situation and offer a rigorous and systematic basis for interventions (Davison et al., 2012). However, while the study provides novel insights and their use of action research makes a valuable methodological contribution, the actual relation between the used methods and artifacts and the behavioral changes has remained blurry. For instance, the authors found that the training approach resonated with the participants when it was relevant to their work. However, the relevancy itself tells very little about how the actual methods and artifacts used in the intervention can be used to make the training relevant. A promising and novel approach to this issue of "invisible training artifact" is provided by Steinbart et al. (2016). They introduced the cybernetic loop to account for the reciprocity between the training and the security behavior. Given the centrality of training methods and artifacts in SETA, we propose further research to explore the possibilities of extending the current methodological approaches to SETA with Action Design Research (ADR) that enable researchers to study trainings as means to naturalistic interventions (Sein et al., 2011) in contrast to artificial experiments. the literature review shows that researchers have applied broadly different and even innovative training approaches in the intervention. This broadness raises concerns about whether it is possible to generate systematic knowledge on the effectiveness of training methods and artifacts unless some form of standardization on the means of intervention exists. Siponen and Baskerville (2018) have argued that IS security studies should focus on intervention rates to generate knowledge on the effectiveness of different IS security approaches. However, unless some form of standardized SETA intervention approach exists, such intervention rates will not be meaningful, and the findings will remain anecdotal. To elaborate, if each SETA intervention uses idiographic methods and artifacts for intervention, generalizations on these interventions are likely to stay idiographic as well. We propose that establishing and defining genres or typologies for SETA methods and artifacts can serve as a relevant basis for establishing intervention rates and allow flexibility for novel design.

6.2 Tentative Design Recommendations for SETA approaches

We derived six tentative design recommendations from the literature to support the further development of SETA approaches. These recommendations are meant as general guidelines, but they require further research

and verification for their scope of applicability in various contexts.

Table 2. Tentative design recommendations for SETA interventions.

Design recommendation	*Description*
Engage participants through interaction	Facilitate interaction, active participation, and discussion (see, e.g., Karjalainen and Siponen (2011), which can be facilitated through workshops (Albrechtsen & Hovden, 2010), gamification (e.g. Dincelli & Chengalur-Smith, 2020) and learner control (Abraham & Chengalur-Smith, 2019).
Ensure contextual relevance	Contextualize SETA by focusing on a specific and relevant task, e.g., password usage (Steinbart et al., 2016).
Tailor the training for specific user(s) tasks or threats	Include personal relevance for the users (Wright & Marett, 2010) and should take into account their particular susceptibility to threats (Goel et al., 2017).
Use concrete and strong fear appeal messages	Base communication on persuasive messages with specific negative consequences (Boss et al., 2015; Johnston & Warkentin, 2010; Siponen & Vance, 2010)
Periodical training	Run periodically to ensure behavioral permanence and cultural embeddedness (Steinbart et al., 2016).
Develop skills	Develop skills required for correct behavior rather than facilitate knowledge on how end-users *should* behave (Jenkins et al., 2021).

7. Conclusions

Based on the review, we conclude that SETA interventions have potential to influence security behavior. While prior research provides evidence of several successful interventions, we lack prescriptions on how to design effective methods, e.g., what design choices make specific training methods effective has remained rather anecdotal. Our research has contributed to narrowing this gap by proposing six tentative design recommendations for effective SETA interventions. These recommendations should be tested empirically to verify their completeness and accuracy.

As a general limitation of this literature review, we emphasize that the reviewed articles are gathered only from a limited set of journals.

References

Abraham, S., & Chengalur-Smith, I. S. (2019). Evaluating the effectiveness of learner controlled information security training. *Computers and Security*, *87*, 101586.

Albrechtsen, E., & Hovden, J. (2010). Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. *Computers and Security*, *29*(4), 432–445.

Aldawood, H., & Skinner, G. (2019). Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review. *Proceedings of 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2018*, 62–68.

Alohali, M., Clarke, N., Furnell, S., & Albakri, S. (2018). Information security behavior: Recognizing the influencers. *Proceedings of Computing Conference 2017*, *2018-January*, 844–853.

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., Polak, P., & Lowry, P. B. (2015). What Do Systems Users Have to Fear ? Using Fear Appeals to Engender Threats and Fear That Motivate Protective Security Behaviors. 39(4), 837–864.

Chowdhury, N., & Gkioulos, V. (2021). Cyber security training for critical infrastructure protection: A literature review. *Computer Science Review*, *40*, 100361.

Davison, R. M., Martinsons, M. G., & Ou, C. X. J. (2012). The roles of theory in canonical action research. *MIS Quarterly: Management Information Systems*, *36*(3), 763–786.

Dincelli, E., & Chengalur-Smith, I. (2020). Choose your own training adventure: designing a gamified SETA artefact for improving information security and privacy through interactive storytelling. *European Journal of Information Systems*, *29*(6), 669–687.

Glaspie, H. W., & Karwowski, W. (2018). Human factors in information security culture: A literature review. *Advances in Intelligent Systems and Computing*, *593*, 267–280.

Goel, S., Williams, K., & Dincelli, E. (2017). Got phished? Internet security and human vulnerability. *Journal of the Association for Information Systems*, *18*(1), 22–44.

Gregor, S. (2006). The Nature Of Theory In Information Systems. *MIS Quarterly*, *30*(3), 611–642.

Hart, S., Margheri, A., Paci, F., & Sassone, V. (2020). Riskio: A Serious Game for Cyber Security Awareness and Education. *Computers and Security*, *95*, 101827.

Hendrix, M., Al-Sherbaz, A., & Bloom, V. (2016). Game Based Cyber Security Training: are Serious Games suitable for cyber security training? *International Journal of Serious Games*, *3*(1), 53–61.

Jaeger, L., & Eckhardt, A. (2021). Eyes wide open: The role of situational information security awareness for security-related behaviour. *Information Systems Journal*, *31*(3), 429–472.

Jenkins, J. L., Durcikova, A., & Nunamaker, J. F. (2021). Mitigating the security intention-behavior gap: The moderating role of required effort on the intention-behavior relationship. *Journal of the Association for Information Systems*, *22*(1), 246–272.

Jensen, M. L., Dinger, M., Wright, R. T., & Thatcher, J. B. (2017). Training to Mitigate Phishing Attacks Using Mindfulness Techniques. *Journal of Management Information Systems*, *34*(2), 597–626.

Jensen, M. L., Durcikova, A., & Wright, R. T. (2019). Using susceptibility claims to motivate behaviour change in IT security. *European Journal of Information Systems*, *30*(1), 1–19.

Johnston, A. C., & Warkentin, M. (2010). Fear Appeals and Information Security Behaviors: An Empirical Study. *MIS Quarterly*, *34*(3), 549–566.

Johnston, A. C., Warkentin, M., & Siponen, M. (2015). An Enhanced Fear Appeal Rhetorical Framework: Leveraging Threats to the Human Asset Through Sanctioning Rhetoric. *MIS Quarterly*, *39*(1), 113-A7.

Karjalainen, M., & Siponen, M. (2011). Toward a New Meta-Theory for Designing Information Systems (IS) Security Training Approaches. *Journal of the Association for Information Systems*, *12*(8), 518–555.

Karlsson, F., Kolkowska, E., & Prenkert, F. (2016). Inter-organisational information security: A systematic literature review. *Information and Computer Security*, *24*(5), 418–451.

Kirova, D., & Baumöl, U. (2018). Factors Affecting the Success of Security Education, Training and Awareness Programs – A Literature Review. *Journal of Information Technology Theory and Application (JITTA)*, *19*(2).

Kretzer, M., & Mädche, A. (2015). Which are the Most Effective Measures for Improving Employees ' Security Compliance ? *ICIS 2015 Proceedings*, 1–17.

Lebek, B., Uffen, J., Neumann, M., Hohler, B., & Breitner, M. H. (2014). Information security awareness and behavior: A theory-based literature review. *Management Research Review*, *37*(12), 1049–1092.

Lebek, B., Uffen, J., Neumann, M., Hohler, B., & H. Breitner, M. (2014). Information security awareness and behavior: a theory-based literature review. *Management Research Review*, *37*(12), 1049–1092.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: The International Journal of an Emerging Transdiscipline*, *9*, 181–213.

Liang, H., & Xue, Y. (2009). Avoidance of information technology threats: A theoretical perspective. *MIS Quarterly: Management Information Systems*, *33*(1), 71–90.

Menard, P., Bott, G. J., & Crossler, R. E. (2017). User Motivations in Protecting Information Security: Protection Motivation Theory Versus Self-Determination Theory. *Journal of Management Information Systems*, *34*(4), 1203–1230.

Ögütçü, G., Testik, Ö. M., & Chouseinoglou, O. (2016). Analysis of personal information security behavior and awareness. *Computers and Security*, *56*, 83–93.

Ormond, D., Warkentin, M., & Crossler, R. E. (2019). Integrating cognition with an affective lens to better understand information security policy compliance. *Journal of the Association for Information Systems*, *20*(12), 1794–1843.

Puhakainen, P., & Siponen, M. (2010). Improving Employees' Compliance Through Information Systems Security Training: An Action Research Study. *MIS Quarterly*, *34*(4), 757–778.

Schuetz, S. W., Benjamin Lowry, P., Pienta, D. A., & Bennett Thatcher, J. (2020). The Effectiveness of Abstract Versus Concrete Fear Appeals in Information Security. *Journal of Management Information Systems*, *37*(3), 723–757.

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *Management Information Systems Quarterly*, *35*(1), 37–56.

Silic, M., & Lowry, P. B. (2020). Using Design-Science Based Gamification to Improve Organizational Security Training and Compliance. *Journal of Management Information Systems*, *37*(1), 129–161.

Siponen, M., & Baskerville, R. (2018). Intervention effect rates as a path to research relevance: Information systems security example. *Journal of the Association for Information Systems*, *19*(4), 247–265.

Siponen, M., Puhakainen, P., & Vance, A. (2020). Can individuals' neutralization techniques be overcome? A field experiment on password policy. *Computers and Security*, *88*.

Siponen, M., & Vance, A. (2010). Neutralization: New Insights into the Problem of Employee Information Systems Security Policy Violations. *MIS Quarterly*, *34*(3), 487-A12.

Steinbart, P. J., Keith, M. J., & Babb, J. (2016). Examining the continuance of secure behavior: A longitudinal field study of mobile device authentication. *Information Systems Research*, *27*(2), 219–239.

Tsohou, A., Karyda, M., Kokolakis, S., & Kiountouzis, E. (2015). Managing the introduction of information security awareness programmes in organisations. *European Journal of Information Systems*, *24*(1), 38–58.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, *37*(1), 205–224.

Wright, R. T., & Marett, K. (2010). The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems*, *27*(1), 273–303.