June 2024

RECONCEPTUALIZING THE HUMAN IN THE LOOP: A PROBLEMATIZATION OF TAKEN-FOR-GRANTED METAPHORS IN CYBERSECURITY RESEARCH

Wael Soliman
University of Agder, wael.soliman@uia.no

Jonna Järveläinen
University of Jyväskylä, jonna.k.jarvelainen@jyu.fi

RECONCEPTUALIZING THE HUMAN IN THE LOOP: A PROBLEMATIZATION OF TAKEN-FOR-GRANTED METAPHORS IN CYBERSECURITY RESEARCH

Completed Research Paper

Wael Soliman, University of Agder, Kristiansand, Norway, wael.soliman@uia.no

Jonna Järveläinen, University of Jyväskylä, Jyväskylä, Finland, jonna.k.jarvelainen@jyu.fi

Abstract

The use of metaphors is crucial to advancing not only scientific fields but also in facilitating the development of knowledge and understanding in general. While metaphors facilitate the exchange of novel concepts and ideas, they can also be a hindrance if we do not critically question the root metaphors and the extent to which they apply and do not apply to the subjects we study in cybersecurity research. We find two metaphors to be seemingly popular in characterizing the human actor: (1) the "frightened animal" metaphor with traceable roots to fear-inducing theories, such as Deterrence Theory, and (2) the "weakest link" metaphor with traceable roots to Taylorism and scientific management. We revisit the roots of these two metaphors, critically examine their applicability to the subject matter in cybersecurity, and provide suggestions for improving the status quo.

Keywords: Cybersecurity, Information Security, Information Systems Security, Metaphor, Frightened Animal, Weakest Link.

1 Introduction

"Invoking a metaphor means opening the door for a listener to enter a subject in a different way."
(Kendall & Kendall, 1993, p. 149)

Readers of the cybersecurity[1] literature may notice that the role of "humans" is often portrayed with a negative undertone. Many may be familiar with the opening statement of academic articles and industry reports that make assertions like, "humans are the weakest link in the cybersecurity chain", "employees are the biggest cybersecurity threat", or more candidly that "the core cybersecurity problem lies between the keyboard and the chair". In a parallel development, versed readers may observe that a popular solution to the weakest link problem is through fear-induced behavioral correction, as professed by some of the most popular theories in behavioral cybersecurity research, such as Deterrence Theory (DT) and Protection Motivation Theory (PMT). Whereas DT emphasizes the fear of (severe, certain, and swift) punishment as a mechanism to deter bad behaviors, PMT emphasizes fear appeals (i.e., messages) as a mechanism to steer people away from unhealthy behaviors. We suspect that these ideas have become so prevalent in the cybersecurity discourse that they could be classified as root metaphors (Morgan, 1980). As Kendall and Kendall (1993) in the above quote suggest, metaphors allow us to observe an issue from a certain perspective, often from a more common perspective as in the "students as consumers" metaphor

[1] We acknowledge that there are nuances that distinguish between the terms 'cybersecurity', 'information security' and 'information systems security', however, in this paper we use the term 'cybersecurity' as an umbrella term covering this broad areas of literature.

(McMillan & Cheney, 1996). From a scientific point of view, a root metaphor is reflective of different schools of thought that accept and perhaps take for granted certain perspectives "as foundation for inquiry" (Morgan, 1980, p. 607).

Metaphors can be so deeply rooted in our discourse that it becomes difficult to note them and their influence on our thinking (Berente, 2020). As such, espousing a certain image manifested in root metaphors could inadvertently steer our research in biased directions. For instance, perpetuating the rhetoric that "humans are the weakest link" can be damaging to problem formulation and the solution search possibilities. This is not to say that the "humans are the weakest link" argument has not been challenged before (e.g., Sasse et al., 2001), but rather to highlight that accepting this rhetoric unquestionably could limit cybersecurity researchers' and practitioners' attention to focus their blaming and controlling efforts on the humans. In the same vein, adopting a criminological lens, such as Deterrence Theory, might lead us to consider employees who do not comply with security policy as "criminals". This is not to say that the criminality assumption has not been challenged in existing cybersecurity research (see Siponen et al., 2022; Willison et al., 2018), but rather the point is that approaching the cybersecurity problem from the DT lens could readily make sanction the preferred solution, even in situations where punishment may be doing more harm than good, as in the case of punishing employees for falling victims to a phishing attack (B. Kim et al., 2020).

In this paper, we identify and scrutinize two metaphors that seem to enjoy wide popularity in the cybersecurity discourse; one implicit and the other explicit. First, the *"human-as-frightened-animal"* is an implicit metaphor that capitalizes on depicting humans as "pain avoiders" who should be controlled by the fear of punishment to refrain from conducting unwanted cybersecurity behavior. Second, the *"human-as-weakest-link"* is an explicit metaphor that capitalizes on depicting humans as the biggest source of weakness and often the one to blame in cybersecurity incidents. We then argue that these negative and condescending perspectives should be questioned. These metaphors may become barriers to finding new narratives for the role of humans, employees, end-users, and managers in cybersecurity. We call for reframing how we depict the human in cybersecurity discourse to find a new solution space.

2 Problematization and Metaphorical Assumptions

The problematization methodology (Alvesson et al., 2011) provides a framework for advancing the body of knowledge in a research area by questioning the taken-for-granted assumptions in that research domain. While the problematization framework has been developed in the management field, it has shown great utility in the information systems (IS) field as well (Carroll et al., 2023; Mikalef et al., 2022; Recker et al., 2021). Problematization is a form of critical research and one of the framework's main objectives is to generate novel research questions inspired by identifying and challenging "the assumptions underlying existing literature" (Alvesson et al., 2011, p. 252).

There is no single way to articulate the assumptions in a given scientific domain (nor there should be). Rather, different scholars provide different typologies that generally address assumptions at macro-, meso- and micro-levels within a given domain. For instance, Morgan (1980) provides three sets of assumptions reflecting the *paradigmatic*, *metaphorical*, and *puzzle-solving* assumptions within the social sciences. Similarly, Alvesson et al.'s (2011) typology categorizes assumptions into *in-house*, *root metaphor*, *paradigm*, *ideology*, and *field* assumptions. In the current study, we focus primarily on the meso-level assumption: namely, metaphor level (Morgan, 1980), or what Alvesson et al. (2011) refer to as root metaphor. We focus primarily on metaphors since they serve as a core defining element for schools of thought and their adopted worldviews. To Morgan (1980), "schools of thought [are] … those communities of theorists subscribing to relatively coherent perspectives, are based upon the acceptance and use of different kinds of metaphor as a foundation for inquiry." (p. 607). Further, metaphors serve as a sensitizing device regarding how we conceive a problem and its potential solutions. In this sense, in identifying and challenging popular metaphorical assumptions in cybersecurity research, not only do we open the door to see the subject matter differently (at the macro level), but also, open the door to broadening the puzzle-solving landscape as well (at the micro level).

Broadly, a metaphor is an "information-processing tool" (Mio, 1997, p. 118), or "an illustrative device whereby a term from one level or frame of reference is used within a different level or frame" (Brown, 1976, p. 170). They are "ways of seeing things *as if* they were something else" (Manning, 1979, p. 661), and therefore, metaphor is described as "seeing something from the viewpoint of something else" (Brown, 1976, p. 170). From the symbolism perspective, a metaphor is a symbol or "an image used for, or regarded as, representing something else. Symbols give meaning to what is perceived; they act as the filter through which the … [world] 'is read.'" (Hirschheim & Newman, 1991, p. 31). Since metaphor "takes the form of saying that one object is another" (Kendall & Kendall, 1993, p. 150), in using metaphors we invoke correspondences or comparison between two domains: "The comparison involves a mapping of the elements of one domain onto those of the other." (Spiggle, 1994, p. 498).

Metaphors may be either *explicit* or *implicit* (Ritchie, 2003). An explicit metaphor makes a clear linkage between the two domains of comparison while stating explicitly that one *is* the other. For example, when one says "time is money" (Hekkala et al., 2018), they are making an explicit parallel between the two domains of 'time' and 'money', most probably as an indication that one's time is valuable. The explicit use of metaphor enables the speaker to strategically shape the worldview of those on the receiving end. For instance, Mio (1997) demonstrates how a politician may invoke a carefully thought-out metaphor to convince voters to pass a bill despite having potential flaws. Alternatively, an implicit metaphor makes the creative leap between the two domains without making an explicit connection between the two domains. This form of metaphor is more difficult to capture for those not fully aware of the context of the conversation. For instance, in their work on knowledge management in IS research, Schultze and Leidner's (2002) analysis revealed that the dominant metaphor for knowledge in the normative IS discourse resembles that of an *"asset"*, and as such, one may argue that this discourse's (implicit) aim is to portray knowledge as "an object that can exist outside an individual, that can be stored and manipulated in the absence of a human knower, and that can be transferred to others (humans or machines)" (Schultze & Leidner, 2002, p. 221). Table 1 illustrates some popular metaphors from different contexts to familiarize readers who might be unacquainted with the subject.

Metaphor	Description
Technology as slaver	The *"technology-as-slaver"* metaphor is often invoked to demonstrate that in some situations where technology is used to govern or control certain tasks, the human end-user turns into an enslaved operator "capable only of action, not thought" (Hirschheim & Newman, 1991, p. 39).
Technology as the dark side	The *"dark side"* metaphor, Sometimes also referred to as the "Darth side" (Venkatraman et al., 2018), is often used by IS researchers to frame something as abnormal or evil since darkness usually invokes images of a "small, dark corner hidden out of sight within which something sinister lurks" (Mikalef et al., 2022, p. 265).
User vs. developer battle	The *"user-developer-battle"* metaphor is often invoked to capture "offensive and defensive strategies" in the system development "battle" between developers and end-users for the sake of their "survival" (Hirschheim & Newman, 1991, p. 39).
Student as consumer	The *"student-as-consumer"* metaphor is rooted in the capitalist view of the "marketplace", as argued by (McMillan & Cheney, 1996), which symbolizes the educational institution's relationship with students as "an organization's reaching out beyond its own boundaries to adapt its 'product' … to the desires of a group of customers or consumers" (p. 3).
Social menace as disease	The *"social-menace-as-disease"* metaphor has dominated the U.S. public discourse as noted by Mio (1997). First, "germ" was the dominant metaphor before it was replaced by the "cancer" metaphor. Whereas the "germ" metaphor characterizes the social menace problem as an external issue such as "foreign invasion into the body" (Mio, 1997, p. 125); the "cancer" metaphor characterizes it as an internal issue such as an "uncontrollable growth" or "rebellion from within" (ibid, p. 125).

Table 1. Popular metaphors in IS research and beyond.

There is no doubt that metaphors are essential as demonstrated by the following testimonials. To Bronowski (1972), metaphors are "the essential core of human thought and creativity" (as cited in Mio, 1997, p. 119). Kendall and Kendall (1993) have argued that "[t]here is true power behind metaphors,

power to shape reality and structure the thoughts of the people who are caught up in a particular metaphor and its entailments" (p. 149). Similarly, Hekkala et al. (2018) argue that "metaphors have 'true power' to shape the reality and thoughts of the people who are caught up in a particular behaviour." (p.143). Considering their immense impact on shaping reality (*at the macro/paradigmatic level*), this (metaphorical) power needs to be exercised with caution since they serve as the "basis of schools of thought" (Morgan, 1980), and adopting metaphors uncritically may lead to situations where they do more harm than good (*at the puzzle solving level*). For example, McMillan and Cheney (1996) warn that the overuse of the now-popular 'student as consumer' metaphor could have dangerous consequences for everyone involved, from professors to administrators, and even the students themselves. Professors who internalize the consumer mentality may seek "to maximize income and minimize teaching responsibilities" (ibid, p. 7), and may display tendencies of "careerism and professional myopia consistent with the market imperative" (ibid, p. 6). Administrators may show tendencies to exercise control to maximize profit. As McMillan and Cheney (1996) put it, "the cost-conscious, market-wise college president and administration may become more heavy-handed in surveillance over all aspects of academic life." (p. 6). Probably the most devastating effect of the student-as-consumer metaphor manifests in the students themselves. The consumer mentality, McMillan and Cheney (1996) argue, can turn students into passive actors in the educational process waiting "to be 'acted upon' rather than to 'act' in the pursuit of their educational goal"; it can turn them into "academic bystander"; an individual who is "content to lay back and wait for the 'quick information fix'" (p. 9).

3 Two Popular Metaphors in Cybersecurity Discourse

In this section we present two seemingly popular metaphors in cybersecurity discourse, namely: (a) the *"human-as-frightened-animal"* metaphor, and (b) the "*human-as-weakest-link"* metaphor. After that, we identify their underlying assumptions and challenge their efficacy, and then we will argue for the need to develop new and more emancipatory metaphors in the cybersecurity discourse. Table 2 summarizes these two popular metaphors and their underlying assumptions. Next, we discuss each of them in more detail.

Metaphor	Traditional Assumption	Challenge to Assumption	New Assumption
(1) The frightened animal metaphor	Humans are directed by the twin goals of pleasure-seeking and pain avoidance, hence instilling in them the fear of punishment is the optimal motivator to steer their behavior in the desired direction.	Most behavioral cybersecurity research does not constitute a crime. Several reasons for non-compliance exist, for example, bad memory, security culture, and non-localized security policies.	Human behavior is much more complex than pleasure seeking and pain avoidance. Contemporary theories of humans can provide a foundation for understanding human drive beyond the extrinsic "carrot and stick" motivations. Outside the domain of clearly defined crimes, instilling fear punishment can be counter-productive.
(2) The weakest link metaphor	Humans (i.e., employees, and end-users) are the root cause of most cybersecurity incidents.	Managers and security experts are also humans, and they create security systems, that one group has to comply with without questioning them. Social engineering attacks are also very sophisticated and might be impossible to detect by end-users.	Technical security solutions and countermeasures are not perfect if they do not take end-users into the cybersecurity loop (Zimmermann & Renaud, 2019).

Table 2. Metaphorical assumptions in cybersecurity research.

3.1 The "frightened animal" metaphor

3.1.1 The metaphor's underlying assumption

The "*human as frightened animal*" is an implicit metaphor that assumes that "fear" is (or should be) the main mechanism to control human behavior. Von Hentig (1938) coined the term "frightened animal" to highlight that the deterrence doctrine is rooted in the reasoning that "[t]he whole animal kingdom may be divided into terrorizers and terrorized specimen. Every living being flees harm …" (von Hentig, 1938, p. 555). This premise serves as a core foundation for the criminological theory of deterrence which emphasizes the role of punishment in deterring (or scaring off) potential transgressors from committing their transgression (Becker, 1968; Gibbs, 1975; Tunick, 1992). Importantly, deterrence serves a dual function: one *specific* and the other *generic*. Specific deterrence posits that experiencing the pain of punishment firsthand will deter those who committed crimes in the past from committing crimes again in the future (Gibbs, 1975). General deterrence posits that applying punishment publicly will deter the observers from engaging in a similar crime in the future (Gibbs, 1975). In this sense, general deterrence might suggest that even though in some situations "… punishment does not seem necessary, or even useful, but we punish nevertheless because we believe that impunity might loosen the ties of discipline and obedience." (von Hentig, 1938, p. 556).

The discourse on cybersecurity deviance as a criminal (anti-social) behavior may be traced to Parker (1976) who introduced the term 'computer abuse'. At that time, the term was used to describe an emerging form of deviance, where savvy "computer operators" used their knowledge and skills to commit crimes, such as digital theft and computer sabotage. Later, this crime view became a major interest for cybersecurity scholars and was formally defined as the "misuse of [organization's] information system assets such as programs, data, hardware, and computer service", where it was mainly characterized as an "anti-social", "deviant act" and a "white collar crime" (Straub, 1989). Today, computer abuse/misuse has become an even more generic term to describe any form of security behavior that deviates from an organizational policy. For instance, D'Arcy et al. (2009) define IS misuse (intention) as "an individual's intention to perform a behavior that is defined by the organization as a misuse of IS resources … The domain of IS misuse is quite varied, ranging from behaviors that are unethical and/or inappropriate (e.g., personal use of company e-mail) to those that are illegal (e.g., accessing confidential company information)." (pp. 3-4).

Making a parallel between violating an organizational security policy and breaking the law may explain why many scholars pointed their attention to the field of criminology as the primary reference discipline to inform their theoretical basis. This is one major reason why deterrence is now recognized as the dominant management philosophy in cybersecurity research (Balozian & Leidner, 2017). A careful reading of the cybersecurity literature reveals that DT is one of the most influential theories to inform both academia and practice on cybersecurity. In fact, the cybersecurity management standard ISO 27002 also draws on DT and emphasizes the use of sanctions (Theoharidou et al., 2005). Moreover, DT has been recently labeled "the single most cited theory in information security literature" (Chen et al., 2018, p. 1049).

Interestingly, like DT, Protection Motivation Theory (PMT), which is equally popular in cybersecurity research, also operates on fear as the central mechanism to change human behavior. As a fear-arousing framework, PMT's central argument is that fear-inducing messages can trigger people to abandon unhealthy behaviors (e.g., smoking, drug abuse, etc.) and adopt more healthy alternatives (Boss et al., 2015; Herath & Rao, 2009; Johnston et al., 2015). While various versions of PMT exist (Haag et al., 2021), the application of PMT has been criticized in cybersecurity research and even labeled as a "misuse of the reference theory" (Johnston et al., 2015, p. 115), mainly because the fear component in most PMT-based cybersecurity studies lacked "personal relevance" (ibid, p. 115). To remedy this misappropriation, Johnston et al. (2015) suggest that adding a threat personal vector using the fear of punishment component advocated in Deterrence Theory. The fear of punishment, in this case, would remedy the missing personal relevance component (i.e., "fear for self") which is integral to PMT, Johnston et al. (2015) explain. Hence, the fear of punishment seems to be a central requirement in the

two most applied theories in cybersecurity research: (a) in DT, as a core mechanism to deter abuse; and (b) in PMT, as a borrowed element from DT to restore the missing fear-for-self component.

3.1.2 Challenging the assumption

We find the frightened animal metaphor and its roots in Deterrence Theory problematic on two levels. The first level relates to DT itself and the other relates to how the theory is applied in the cybersecurity context, as we show next.

First, regarding Deterrence Theory itself.

Gibbs (1975) is often credited for reviving DT as a formal theory, however, the intellectual foundation of DT dates back to the works of Enlightenment philosophers, such as Hobbes, Bentham, and Beccaria, who adopt a specific view on human nature (Siponen et al., 2022). For instance, Beccaria viewed humans by their very nature to be self-interested beings who are inclined to commit crimes (Paternoster, 2010). Closely related, Bentham believed that "human behavior is directed by the twin goals of the attainment of pleasure and the avoidance of pain" (Paternoster, 2010, p. 770). Based on this view, it is convincing to argue that in their pursuit of pleasure, humans will commit crimes, and the only way to deter them is to instill fear of punishment that is 'severe', 'certain', and 'swift' (Gibbs, 1975).

As compelling a view as it may be, DT has received extensive critique for various reasons. For instance, criminologists have long recognized crimes may be driven by a myriad of motives beyond the simplistic notion of pleasure and pain (Paternoster, 2010). Furthermore, even if we accept the argument that humans, being members of the animal kingdom, can be controlled by inducing the fear of punishment, it is not guaranteed that fear will produce the desired effect. On a conceptual ground, von Hentig (1938) has argued that "[f]ear does not in every case mean that the frightened animal will desist from its initial aims. Fear can produce, and often produces a mere change of direction, a clever detour, or an aggressive protective reaction." (p. 555).

Furthermore, and on more empirical grounds, Deterrence Theory has been criticized by various criminologists for making grand claims that are not backed by convincing empirical evidence. For instance, Kennedy (1983) noted nearly four decades ago that "[l]ittle proof, if any, can be mustered to support the proposition that punishment has ever deterred potential criminals from committing crimes" (p. 11), leading them to conclude that "[f]or too long criminal deterrence theory has been a small tail wagging a very large dog" (Kennedy, 1983, p. 12).

Second, regarding the application of Deterrence Theory in the cybersecurity context.

There have been various calls to exercise caution when applying DT in contexts that are nonmalicious, non-criminal, and/or unintentional (see Guo et al., 2011; Siponen et al., 2022; Willison et al., 2018). Yet, assuming equivalency between criminal violations and the violations typically studied in cybersecurity research contexts (e.g., creating a weak password, or using a USB memory stick) prevails. Empirical cybersecurity research demonstrates that sanctions may be effective in deterring undesired computer abuse (intentions); while other empirical research implies that sanctions will not deter such abuse. For instance, some findings suggest that DT would be an effective approach to deter employees' computer abuse/misuse (intentions) within organizations (D'Arcy et al., 2009; D'Arcy & Hovav, 2009; Siponen et al., 2010; Straub, 1990; Ugrin et al., 2008). On the other hand, there are empirical findings that would suggest otherwise (Guo et al., 2011; Hu et al., 2011; Johnston et al., 2015; Lowry et al., 2015; Siponen & Vance, 2010). For instance, sanctions were not found to be a significant factor in explaining/predicting compliance intentions, leading Moody et al. (2018) to remove deterrents from their unified model of ISP compliance. In light of these conflicting findings, D'Arcy and Herath (2011) have argued that such discrepancies may be explained by several contingency variables that could moderate the impact of sanction on computer abuse (intentions). Such variables may include self-control, moral beliefs, and employee position, among others. In a similar vein, Siponen and Vance (2010) have argued that sanctions might not be effective in deterring computer abuse because employees could use different self-justifying techniques (based on neutralization theory) that reduce the perceived harm of their wrongdoing, thus moderating the deterrent effect of sanction. One may argue that a dominant underlying assumption of this line of work is that the phenomenon of interest is justifiably

punishable wrongdoing. In such situations, the researchers take on the task of explaining (or predicting) the occurrence of such wrongdoing. What is often overlooked in studies adopting the crime view is that studying cybersecurity behavior from a criminological lens does not make it a crime. In other words, if in the worldview of an employee, it is not justifiable to call a deed a crime, then there is no reason to believe that sanction will have a deterring effect. Siponen & Baskerville (2018) have articulated this problem by noting that "[d]eterrence theory may apply to criminal behavior, but it does not make sense for all insecure behaviors. For instance, weak password selection can occur because some users cannot memorize a password. Sanctions can hardly improve human memory." (p. 255).

Closely related, the assumption that humans have a natural tendency to crime and applying a fear-based governance framework is very problematic since it can produce undesired negative consequences (Siponen et al., 2022). For instance, research investigating the application of Panoptic-like governmental surveillance of online activities is a good example. The stated objective of Panoptic surveillance is often to create a mental state of deterrence "wherein individuals continually assess the costs and benefits of legal and illegal activity and reform their behavior accordingly." (Stoycheff et al., 2019, p. 604). On the flip side, deterrence has also been shown to produce the so-called "chilling effect", which suggests that people may be discouraged from engaging in legal and healthy behavior (e.g., political discussions) out of fear that they may be punished (Stoycheff et al., 2019).

3.2 The "weakest link" metaphor

3.2.1 The metaphor's underlying assumption

The "human as the weakest link" is an explicit metaphor that assumes that humans are a problem in a production line and are often the first to blame in cybersecurity incidents. Finding the exact origin of the notion that "human is the weakest link" is difficult (if not impossible), although there is some indication that it could be rooted in writings that are influenced by Taylor's scientific management (Braverman, 1998). We would like to highlight that accurately tracing the origin of the metaphor is not our primary target here. Rather, we are more interested in tracing potential roots that give this narrative its dominance, which could help us better understand the implications it has on viewing the human in the cybersecurity loop.

For instance, one possibility is that the metaphor is rooted in what Schein calls the "engineering culture" (Schein, 1996). Based on Schein's (1996) extensive work on culture, he noted the contrast between the "operator culture" (e.g., regular employees) and the "engineering culture" (e.g., "designers of the technology underlying the work" (p. 14). According to Schein (1996), engineering education "reinforces the view that problems have abstract solutions and that those solutions can, in principle, be implemented in the real world with products and systems free of human foibles and errors" (p. 14). Further, Ebert et al. (2023, p. 2) have noticed similarities in cybersecurity and safety science discourses. They argue that during the heyday of Taylorism in the 1910s "Human operators were framed as a problem to be controlled by enforcing compliance with rules and penalizing violations", which bears a resemblance to the "human as the weakest link" metaphor.

Another early source, and perhaps more influential seems to be Schneier (2000, p. 149), who argues "People often represent the weakest link in the security chain and are chronically responsible for the failure of security systems.". He further explains "The mathematics are impeccable, the computers are invincible, the networks are lousy and the people are abysmal". Schneier is an influential cybersecurity expert (Mc Mahon, 2020), who has published several books on the topic and actively writes a blog on cybersecurity issues. It is therefore likely that even if the origin of this claim would not be him, he has influenced the wide distribution of the "human is the weakest link" metaphor. McMahon (2020) argues that this metaphor is prevalent in current cybersecurity discourse from blogs to government-sponsored panel discussions, as well as in scientific literature. In the organizational context, often the metaphor is translated as the employee (Boss et al., 2009; Bulgurcu et al., 2010; Dhillon et al., 2021; Johnston et al., 2016; Silic & Lowry, 2020) or the user (Alohali et al., 2018; Heartfield & Loukas, 2018; E. Kim, 2013;

Nguyen et al., 2021) is the weakest link. From this, we may conclude, that it seems that the "human" referred to in the metaphor is seldom the one designing the security controls or technical solutions.

Nevertheless, whether the origins of this root metaphor are in Taylorism, engineering culture, or the impeccability of mathematics, the metaphor implants the idea of security as a chain, with one weak link –the human– that shatters the chain.

3.2.2 Challenging the assumption

The powerful imagery of this metaphor may dazzle the audience in various ways (Weick, 1989, p. 529). However, the assumption of security as a chain, which can be broken where it is at its weakest, can be challenged and the simplified imagery questioned.

First, regarding the linearity assumption.

The weakest link metaphor portrays security as a linear chain, which can easily be broken on a single point. Yet, security is not a chain, but a complex management system containing technical and social parts, which includes organizational processes and people or actors (Malatji et al., 2020). This metaphor focuses on blaming a single factor, the human (end-user), for the incident and thus ignores the complexity of the situation and aims to find the root cause and fix that (Ebert et al., 2023). Therefore, the metaphor limits improving cybersecurity holistically in organizations, since the compelling simplified root cause has been found already: "the human". Instead of blaming the humans, we may do a better job addressing the challenges recognized in the cybersecurity literature, such as lack of training (Hagen et al., 2008), the complexity of cybersecurity policies (Karlsson et al., 2017), usability of security (Whitten & Tygar, 1999), and systemic problems (Ebert et al., 2023). Finding a root cause of a cyber breach to be the "end-user" (or human error) amounts to little more than scapegoating since it allows the organization to avoid investigating and fixing the more complex system, the security environment (Mc Mahon, 2020). Edwards (1997, p. 290 in Klimburg-Witjes & Wentland, 2021) argues that computer systems are inherently insecure, and thus focusing on merely the end-user is an oversimplification of the contextual environment. After all, there might be also other reasons for security problems, such as a lack of localization of cybersecurity policies (Niemimaa & Niemimaa, 2019), overconfidence in technical controls (Wang et al., 2016), or bad memory (Siponen & Baskerville, 2018).

Second, regarding technical supremacy.

If humans were the weakest link, it would then assume that technical security is the stronger part of the security chain. If we consider the people designing technical security controls and tools as humans, they are also fallible (Zimmermann & Renaud, 2019). Thus, the technology and tools also contain weaknesses, they do not detect every possible threat. If the technical security was infallible, it should be able to detect cyber-attacks (e.g., social engineering, data breaches, etc.) before the end-user even sees them, and the end-user interference would be unnecessary. However, the ever-changing social engineering and cyberattacks in the age of artificial intelligence are so difficult to detect that humans are required to be the last line of defense when the technical tools fail.

Third, regarding the loss of human agency.

Portraying humans as the "weakest link" transforms the human agency from an "actor" to an "object" in the security discourse, the object of training and social engineering, and for whom the policies are written to control. Klimburg-Witjes and Wentland (2021, p. 1320) notice similarities in the role of humans in the security or social engineering discourse as the layperson has in public understanding of science and technology. They point out further two problems "1) arrogance/institutional language of expert bodies, 2) overlooking valuable contributions from lay people.". One example of the arrogance of expert bodies is the complexity of cybersecurity policies (see e.g., Karlsson et al., 2017), which end-users are expected to comply with while disregarding any critical thinking, although they also might be poorly suitable for the current work context (Niemimaa & Niemimaa, 2019). Seeing end-users as deficient or fallible, who do not have sufficient training or knowledge to notice, for example, a social engineering attack, makes a stark contrast to the cybersecurity experts who can decide how the security controls are selected and implemented (Klimburg-Witjes & Wentland, 2021).

Further, portraying humans as the "weakest link" contains the idea that the human actor is somehow breakable or broken. Blaming the victim transfers shame from the actual cybercriminals and their advanced methods to the targets (Cross, 2015). Mc Mahon (2020) also reminds us that if a user clicks hundreds of hyperlinks during the day, it is questionable to deduce that one click on one day turns them into the origins of a cyber breach. It has been established that continuous or sustained mindfulness of, for example, security threats is impossible, since "attention is scarce" (Levinthal & Rerup, 2006, p. 509) as well as easily fleeting. But this is not a human weakness; rather a human nature. Prior research has also found that if employees feel that they are not trusted members of the organization, the degree of computer abuse incidents seems to grow (Posey et al., 2011). Further, shaming employees for cyber breaches seems to be counterproductive to compliance and employee-employer relationships, but a more positive, supportive response to breaches could in fact lead to better results that improve security (Renaud et al., 2022).

4 Discussion

So far, we have analyzed two root metaphors and scrutinized their content, imagery, and their potential impact on the discourse in the cybersecurity field. We demonstrated that these metaphors depict the human as a *'frightened animal'* and *'the weakest link'*. The pressing question now is: How can we improve the role of humans in cybersecurity in the future? First, we suggest that we recognize the power of language of the strong imagery metaphors can implant in our thinking. Second, we propose that there is a need for new metaphors that empower, rather than denigrate, the human in the loop. We briefly elaborate on these suggestions next.

4.1 Pay attention to the power of language

Problematizing these two root metaphors draws our attention to the imbued image of a flawed user in cybersecurity discourse and practice, hinting that on the other side is the perfect security system or technical security. When the human (end-user, employee) is seen as a frightened animal or as the weakest link, the powerful language of metaphors has consequences. As McMillan and Cheney (1996, p. 11) state "Metaphors bear watching and listening. […] Do they still mean what we want them to mean? A presumably useful metaphor […] when we turn our backs on it, can drag us where we really don't want to go."

The frightened animal metaphor presents the idea of fear appeals, which are a central part of the Deterrence Theory and Protection Motivation Theory (at least, as popularized by many in the cybersecurity field). When we draw a parallel between behavioral cybersecurity and criminology, we derive the idea that criminals break the law and then we start assuming that cybersecurity policies and other controls are effectively the "law". From that assumption, it is not implausible that sanctions and other punishments are the solution for non-compliance (Willison et al., 2018). Further, although the insider threat exists, our conception of the "insider" needs to be more nuanced. Balozian and Leidner's (2017, p. 14) classification points to different types of insiders, and that not all policy violations are intentionally malicious. For instance, some employees "are willing to comply but are not able" which can be due to naivety, lack of awareness, or improper training (Hagen et al., 2008; West, 2008). In addition, this line of thinking may lead to disregarding also external criminals, who can be performing cybercrimes in the eyes of the law.

Inherent in these framings is the assumption that rules written in an ISP are the authoritative voice setting the boundaries for what is (and is not) acceptable. Hence, violating ISP is often portrayed as a sign of "laziness" or "sloppiness" (Boss et al., 2009). There is nothing wrong with this assumption in contexts where the framing adopted in ISPs is congruent with that of the employees' framing. For instance, there is no reason to challenge this view if, the ISP of an organization mandates that sensitive data cannot be shared with outsiders, and the employees believe that sharing sensitive data with outsiders poses a major risk to the organization. In such situations, the ISP requirement is justified. Here, a top-down approach to cybersecurity is assumed, where top management dictates and enforces an ISP with the correct knowledge and the subordinates obey or face the consequences, such as sanctions and monitoring (Zhu

et al., 2023). But the question now is: what if the ISP requires a rule that clearly violates the contextual relevance? The reasons behind non-compliance are varied, such as usability of security measures (Sasse et al., 2001), security culture (da Veiga & Martins, 2017), lack of training (Hagen et al., 2008), and bad memory (Siponen & Baskerville, 2018). Treating employees as "lazy", "insider threats" or even "criminals" is hardly reasonable in these cases, and punishing them with sanctions for ISP violations is questionable.

Claiming users being the weakest link in the security chain (Schneier, 2000) obtains a linear view of security although it is far more complex socio-technical system (Ebert et al., 2023; Malatji et al., 2020). Also, it indicates an idea of unbreakable technical security in contrast to the breakable human, although in the end, users are required to notice cyberattacks and social engineering attacks when the technical controls have not detected them. Although mathematics is used in cryptography and those cryptographic solutions may be "impeccable", as Schneier (2000, p. 149) states, this does not result in impeccable technology, nor an impeccable security system.

However, it might be tempting to disregard the human element (end-user, employee) as something weak, as the taken-for-granted part of the organization. Companies seem to make significant efforts to develop extensive security controls including ISPs but raise their hands in the air as a sign of surrender and copy the sanction section to ISP from the existing template as a final "legal" guarantee. Therefore, they may forget about training the users (Hagen et al., 2008), localizing ISPs (Niemimaa & Niemimaa, 2019), or using clear language and structure in it (Karlsson et al., 2017). In current digital transformation literature, similar notions have been problematized: "[O]rganizations are spending vast amounts of resources on digital transformation projects and often assume that clear business goals and objectives will cascade down to the work floor" (Carroll et al., 2023).

In conclusion, many cybersecurity researchers using these two metaphors (frightened animal, and the weakest link) seem to overlook the borrowed imagery of these metaphors and treat the subject matter as if it were a perfect instantiation of the appropriated image. In so doing, the cybersecurity community, given the time, may start to treat these borrowed images as reality itself. Morgan (1980) describes how this has occurred among organization scientists: "… the metaphorical nature of the image which generated … [the research] concepts is lost from view, and the process of organizational analysis becomes over-concretized as theorists and researchers treat the concepts as a description of reality." (Morgan, 1980, p. 612).

4.2 New liberating theories, metaphors, and imagery

So far, we have discussed and challenged two metaphors that depict the human with a negative undertone (the 'frightened animal' and the 'weakest link'). In his seminal work, Morgan (1980) argued that "… new metaphors may be used to create new ways of viewing organizations which overcome the weaknesses and blindspots of traditional metaphors, offering supplementary or even contradictory approaches to organizational analysis." In this same spirit, we argue that we need new metaphors that empower, rather than, denigrate the human actor. The good news is, that examples of such imagery do exist, albeit not as vocal as they could. In the following lines, we briefly present a few ideas for future directions and hope that they act as inspiration for developing more empowering metaphors.

First, metaphors that promote the human's self-determined behavior.

What if, instead of relying on fear-driven solutions, we explored approaches that empower individuals through autonomy and self-determination? Theories of motivation, such as Self-Determination Theory (SDT), offer a broader perspective on human drives, moving beyond traditional concepts of punishment and reward (Ryan & Deci, 2000). SDT posits a spectrum of regulatory styles, ranging from external to intrinsic motivation, each influencing behavior differently. As behavior shifts from external controls like punishment and reward towards internal, intrinsic motivations, it becomes more sustainable and fulfilling (Ryan & Deci, 2000). SDT underscores the fundamental human needs for autonomy, competence, and relatedness. When these needs are met, individuals are more likely to internalize behaviors, fostering commitment, effort, and excellence (Ryan & Deci, 2000, p. 76). Interestingly, applying this framework in cybersecurity study reveals that addressing these basic human needs

enhances security behavior, both in terms of in-role compliance as well as extra-role volunteering (Davis et al., 2023). Therefore, emphasizing imagery that promotes competence, autonomy, and relatedness seems to offer an effective motivational strategy away from fear.

Second, metaphors that promote a sense of community.

Johnston et al (2019) have recently suggested imagery that capitalizes on the power of the collective, with slogans such as *"It Takes a Village"*. We are more familiar with the original idiom "it takes a village to raise children", which might be reformulated to this context perhaps "it takes a village to build cybersecurity". Being part of a village connects an individual to a community. It further assumes that the village is competent in building security and also the individual actors have the autonomy to behave as village people. Klimburg-Witjes and Wentland (2021) further call for collective-level consideration of security, or social engineering, which would emphasize the probabilities and impacts of risks instead of relying on the end-user. Yoo et al (2020) have noticed that the group level has significant effects on cybersecurity, they even raise the question in the article's title: "Is information security a team sport?", which is yet another empowering image. Athletes train to become top performers in their sport, and connecting sports and athletic training to cybersecurity would give a more positive spin and empowering message for users. Taking the team or community perspective in cybersecurity theorizing would bring new perspectives but also challenges to cybersecurity research. There is some community-level research, which has been now concentrating more on incident management teams or security operations center teams (e.g., Shah et al., 2023; Thangavelu et al., 2021) or on the entire cybersecurity culture (da Veiga & Martins, 2017), but understanding the roles of employee teams in organizational cybersecurity has not received much attention on behavioral cybersecurity agenda (see Mallmann & Soliman, 2023).

This kind of imagery is needed for the empowerment of users as an integral part of cybersecurity management, a central force, or a hub if you will. Zimmermann & Renaud (2019) propose moving from viewing a 'human-as-problem' to a 'human-as-solution' mindset and viewing users as valuable actors, and 'first responders' to security threats. A similar empowerment message is in the UK's National Cybersecurity Centre's (National Cyber Security Centre, 2019) "You shape security" guidance, or in the "security hero" notion (Pfleeger et al., 2014). Ebert et al. (2023) suggest that focusing on a single factor or root cause can occasionally be appropriate but there lies a risk for victim blaming. They continue that if an organization aims to study the situation holistically and improve its security with different measures, cultural, management, or systemic incident causation models are more suitable methods to follow.

We urge cybersecurity researchers to explore new approaches to motivate users to exhibit secure behavior. Future studies could focus on shaping security designs that resonate with human preferences, fostering a collective atmosphere for improved organizational cybersecurity, and promoting user competence and empowerment in security practices. For instance, comparative case studies or ethnographic research can examine how metaphors empower both humans and technology in organizational cybersecurity. Empowering imagery can be integrated into security training and experimental designs can be developed to assess their organizational impact and effectiveness. Moreover, design science research can develop tools to enhance intrinsic motivation for policy compliance, while exploring effective collective user training strategies tailored to different user groups. We advocate for research that empowers all stakeholders, from end-users to top management, in cybersecurity management systems. For practitioners, we issue a similar challenge to develop positive engagement practices, incentives, and empowering metaphors for end-users. Encouraging positive imagery and shifting perceptions about cybersecurity from a barrier to a facilitator of business can be impactful. Community building should become a priority in cybersecurity management, transforming the human role into an active and capable part of a cohesive security community within organizations.

4.3 Limitations

There are some limitations in this paper. First, this is a conceptual paper, and as such, we do not present our arguments as an indisputable truth, but rather as a cultivation of our reading of the cybersecurity literature. Second, we acknowledge that the metaphors (and their underlying assumptions, imagery, and

roots) that we presented in this article are not necessarily reflective of all cybersecurity researchers, and indeed, there is plenty of research that does not use these metaphors or even acknowledges their possible impact on, or relevance to, their own work. And last, we recognize that in some cultures, or organizations, authoritative leadership is dominant and user empowerment might not be suitable for that specific organization or culture.

5 Conclusion

In this article, drawing inspiration from the research problematization framework, we focused on the utilization of root metaphors in cybersecurity discourse. Our objective has been to stimulate critical reflection on the past and offer insights that can enhance future research endeavors. We focused on two popular root metaphors in cybersecurity discourse, namely *'human as a frightened animal'* and *'human as the weakest link'*. We challenged the conventional assumptions associated with employing fear appeals and sanctions to regulate cybersecurity behavior. Additionally, we questioned the notion of security as a linear chain and the belief that the technical security role outweighs that of humans. These metaphors, we argue, not only raise doubts but also have potentially adverse effects on both human participants and cybersecurity itself. Subsequently, we proposed alternative avenues for depicting humans in a more empowering light, suggesting new metaphors that emphasize self-determined behavior and a sense of community. Furthermore, we advocated for exploring fresh motivational approaches beyond the conventional dichotomy of reward and punishment (or the stick and carrot). Ultimately, our aim has been twofold: to critique prevailing metaphors and to chart a course toward a more nuanced and inclusive approach to cybersecurity discourse and practice. We hope that this article contributes to reshaping the narrative surrounding cybersecurity, paving the way for a more holistic understanding and implementation of cybersecurity.

References

Alohali, M., Clarke, N., Li, F., & Furnell, S. (2018). Identifying and predicting the factors affecting end-users' risk-taking behavior. *Information and Computer Security*, *26*(3), 306–326.

Alvesson, M., Sandberg, J., & Orgen, J. (2011). Generating research questions through problematization. *Academy of Management Review*, *36*(2), 247–271.

Balozian, P., & Leidner, D. (2017). Review of IS security policy compliance: Toward the building blocks of an IS security theory. *Data Base for Advances in Information Systems, 48*(3), 11–43.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy*, *76*, 169–217.

Berente, N. (2020). Agile Development as the Root Metaphor for Strategy in Digital Innovation. In *Handbook of Digital Innovation* (pp. 83–96). Edward Elgar Publishing.

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., & Polak, P. (2015). What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviours. *MIS Quarterly*, *39*(4), 837–864.

Boss, S. R., Kirsch, L. J., Angermeier, I., Shingler, R. A., & Boss, R. W. (2009). If someone is watching, I'll do what I'm asked: Mandatoriness, control, and information security. *European Journal of Information Systems*, *18*(2), 151–164.

Braverman, H. (1998). *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. NYU Press.

Brown, R. H. (1976). Social theory as metaphor: On the logic of discovery for the sciences of conduct. *Theory and Society*, *3*(2), 169–197.

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, *34*(3), 523–548.

Carroll, N., Conboy, K., Hassan, N., Hess, T., Junglas, I., & Morgan, L. (2023). Problematizing assumptions on digital transformation research in the information systems field. *Communications of the Association for Information Systems*, *53*(1), 508–531.

Chen, X., Wu, D., Chen, L., & Teng, J. K. L. (2018). Sanction severity and employees' information security policy compliance: Investigating mediating, moderating, and control variables. *Information and Management*, *55*(8), 1049–1060.

Cross, C. (2015). No laughing matter: Blaming the victim of online fraud. *International Review of Victimology*, *21*(2), 187–204.

da Veiga, A., & Martins, N. (2017). Defining and identifying dominant information security cultures and subcultures. *Computers & Security*, *70*, 72–94.

D'Arcy, J., & Herath, T. (2011). A review and analysis of deterrence theory in the IS security literature: Making sense of the disparate findings. *European Journal of Information Systems*, *20*(6), 643–658.

D'Arcy, J., & Hovav, A. (2009). Does one size fit all? Examining the differential effects of IS security countermeasures. *Journal of Business Ethics*, *89*, 59–71.

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research*, *20*(1), 79–98.

Davis, J., Agrawal, D., & Guo, X. (2023). Enhancing users' security engagement through cultivating commitment: The role of psychological needs fulfilment. *European Journal of Information Systems*, *32*(2), 195–206.

Dhillon, G., Smith, K., & Dissanayaka, I. (2021). Information systems security research agenda: Exploring the gap between research and practice. *The Journal of Strategic Information Systems*, *30*(4), 101693.

Ebert, N., Schaltegger, T., Ambuehl, B., Schöni, L., Zimmermann, V., & Knieps, M. (2023). Learning from safety science: A way forward for studying cybersecurity incidents in organizations. *Computers & Security*, *134*, 103435.

Gibbs, J. P. (1975). *Crime, punishment, and deterrence*. Elsevier.

Guo, K. H., Yuan, Y., Archer, N. P., & Connelly, C. E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems*, *28*(2), 203–236.

Haag, S., Siponen, M., & Liu, F. (2021). Protection Motivation Theory in Information Systems Security Research: A Review of the Past and a Road Map for the Future. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *52*(2), 25–67.

Hagen, J. M., Albrechtsen, E., & Hovden, J. (2008). Implementation and effectiveness of organizational information security measures. *Information Management and Computer Security*, *16*(4), 377–397.

Heartfield, R., & Loukas, G. (2018). Detecting semantic social engineering attacks with the weakest link: Implementation and empirical evaluation of a human-as-a-security-sensor framework. *Computers & Security*, *76*, 101–127.

Hekkala, R., Stein, M., & Rossi, M. (2018). Metaphors in managerial and employee sensemaking in an information systems project. *Information Systems Journal*, *28*, 142–174.

Herath, T., & Rao, H. R. (2009). Protection motivation and deterrence: A framework for security policy compliance in organisations. *European Journal of Information Systems*, *18*(2), 106–125.

Hirschheim, R., & Newman, M. (1991). Symbolism and information systems development: Myth, metaphor and magic. *Information Systems Research*, *2*(1), 29–62.

Hu, Q., Xu, Z., Dinev, T., & Ling, H. (2011). Does deterrence work in reducing information security policy abuse by employees? *Communications of the ACM*, *54*(6), 54.

Johnston, A. C., Di Gangi, P. M., Howard, J., & Worrell, J. (2019). It takes a village: Understanding the collective security efficacy of employee groups. *Journal of the Association for Information Systems*, *20*(3), 186–212.

Johnston, A. C., Warkentin, M., McBride, M., & Carter, L. (2016). Dispositional and situational factors: Influences on information security policy violations. *European Journal of Information Systems*, *25*(3), 231–251.

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), 113–134.

Karlsson, F., Hedström, K., & Goldkuhl, G. (2017). Practice-based discourse analysis of information security policies. *Computers & Security*, *67*, 267–279.

Kendall, J. E., & Kendall, K. E. (1993). Metaphors and methodologies: Living beyond the systems machine. *MIS Quarterly: Management Information Systems*, *17*(2), 149–168.

Kennedy, K. C. (1983). A critical appraisal of criminal deterrence theory. Michigan State University College of Law. In *Dick. L. Rev.* (Vol. 88, pp. 1–13). Michigan State University College of Law.

Kim, B., Lee, D.-Y., & Kim, B. (2020). Deterrent effects of punishment and training on insider security threats: A field experiment on phishing attacks. *Behaviour & Information Technology*, *39*(11), 1156–1175.

Kim, E. (2013). Information Security Awareness Status of Business College: Undergraduate Students. *Information Security Journal*, *22*(4), 171–179.

Klimburg-Witjes, N., & Wentland, A. (2021). Hacking Humans? Social Engineering and the Construction of the "Deficient User" in Cybersecurity Discourses. *Science, Technology, & Human Values*, *46*(6), 1316–1339.

Levinthal, D., & Rerup, C. (2006). Crossing an apparent chasm: Bridging mindful and less-mindful perspectives on organizational learning. *Organization Science*, *17*(4), 502–513.

Lowry, P. B., Posey, C., Bennett, R. (Becky) J., & Roberts, T. L. (2015). Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: An empirical study of the influence of counterfactual reasoning and organisational trust. *Information Systems Journal*, *25*(3), 193–273.

Malatji, M., Marnewick, A., & von Solms, S. (2020). Validation of a socio-technical management process for optimising cybersecurity practices. *Computers & Security*, *95*, 101846.

Mallmann, G., & Soliman, W. (2023). The Collective Violation Talk Show: How Do Workgroups Account for Cyberdeviance? *14th Scandinavian Conference on Information Systems*. https://aisel.aisnet.org/scis2023/9

Manning, P. K. (1979). Metaphors of the field: Varieties of organizational discourse. *Administrative Science Quarterly*, *24*(4 Qualitative Methodology), 660–671.

Mc Mahon, C. (2020). In Defence of the Human Factor. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01390

McMillan, J. J., & Cheney, G. (1996). The student as consumer: The implications and limitations of a metaphor. *Communication Education*, *45*(1), 1–15.

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems*, *31*(3), 257–268.

Mio, J. S. (1997). Metaphor and Politics. *Metaphor and Symbol*, *12*(2), 113–133.

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a unified model of information security policy compliance. *MIS Quarterly: Management Information Systems*, *42*(1), 285–311.

Morgan, G. (1980). Paradigms, metaphors, and puzzle solving in organization theory. *Administrative Science Quarterly*, *25*(4), 605–622.

National Cyber Security Centre. (2019, February 12). *You shape security*. https://www.ncsc.gov.uk/collection/you-shape-security

Nguyen, C., Durcikova, A., Jensen, M. L., & Wright, R. T. (2021). *RESEARCH ARTICLE A comparison of features in a crowdsourced phishing warning system*. *January 2020*, 1–41.

Niemimaa, M., & Niemimaa, E. (2019). Abductive innovations in information security policy development: An ethnographic study. *European Journal of Information Systems*, *28*(5), 566–589.

Parker, D. B. (1976). *Crime by computer*.

Paternoster, R. (2010). How much we really know about criminal deterrence? *The Journal of Criminal Law and Criminology*, *100*(3), 765–824.

Pfleeger, S. L., Sasse, M. A., & Furnham, A. (2014). From Weakest Link to Security Hero: Transforming Staff Security Behavior. *Journal of Homeland Security and Emergency Management*, *11*(4), 489–510.

Posey, C., Bennett, R. J., & Roberts, T. L. (2011). Understanding the mindset of the abusive insider: An examination of insiders' causal reasoning following internal security changes. *Computers & Security*, *30*(6), 486–497.

Recker, J., Lukyanenko, R., Jabbari, M., Samuel, B. M., & Castellanos, A. (2021). From representation to mediation: A new agenda for conceptual modeling research in a digital world. *MIS Quarterly*, *45*(1), 269–300.

Renaud, K., Searle, R., & Dupuis, M. (2022). Shame in Cyber Security: Effective Behavior Modification Tool or Counterproductive Foil? *Proceedings of the 2021 New Security Paradigms Workshop*, 70–87.

Ritchie, D. (2003). 'ARGUMENT IS WAR'-Or is it a Game of Chess? Multiple Meanings in the Analysis of Implicit Metaphors. *Metaphor and Symbol*, *18*(2), 125–146.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sasse, M. A., Brostoff, S., & Weirich, D. (2001). Transforming the 'Weakest Link'—A Human/Computer Interaction Approach to Usable and Effective Security. *BT Technology Journal*, *19*(3), 122–131.

Schein, E. H. (1996). Three cultures of management: The key to organizational learning. *Sloan Management Review*, *38*(1), 9–20.

Schneier, B. (2000). *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons.

Schultze, U., & Leidner, D. E. (2002). Studying knowledge management in information systems research: Discourses and theoretical assumptions. *MIS Quarterly*, *26*(3), 213–242.

Shah, A., Ganesan, R., Jajodia, S., Cam, H., & Hutchinson, S. (2023). A Novel Team Formation Framework Based on Performance in a Cybersecurity Operations Center. *IEEE Transactions on Services Computing*, *16*(4), 2359–2371.

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., Pahnila, S., & Mahmood, M. A. (2010). Compliance with information security policies: An empirical investigation. *IEEE Computer*, *43*, 64–71.

Siponen, M., Soliman, W., & Vance, A. (2022). Common misunderstandings of deterrence theory in information systems research and future research directions. *The DATA BASE for Advances in Information Systems*, *53*(1), 25–60.

Siponen, M., & Vance, A. (2010). Neutralization: New insights into the problem of employee information systems security policy violations. *MIS Quarterly*, *34*(3), 487–502.

Spiggle, S. (1994). Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research*, *21*(3), 491.

Stoycheff, E., Liu, J., Xu, K., & Wibowo, K. (2019). Privacy and the Panopticon: Online mass surveillance's deterrence and chilling effects. *New Media and Society*, *21*(3), 602–619.

Straub, D. (1990). Effective IS security: An empirical study. *Information Systems Research*, *1*(3), 255–276.

Straub, D. W. (1989). Validating instruments in MIS research. *MIS Quarterly*, *13*(2), 147–169.

Thangavelu, M., Krishnaswamy, V., & Sharma, M. (2021). Impact of comprehensive information security awareness and cognitive characteristics on security incident management – an empirical study. *Computers & Security*, *109*, 102401.

Theoharidou, M., Kokolakis, S., Karyda, M., & Kiountouzis, E. (2005). The insider threat to information systems and the effectiveness of ISO17799. *Computers and Security*, *24*(6), 472–484.

Tunick, M. (1992). *Punishment: Theory and Practice*. University of California Press.

Ugrin, J., Pearson, J., & Odom, M. (2008). Cyber-slacking: Self-control, prior behavior And the impact Of deterrence measures. *Review of Business Information Systems*, *12*(1), 75–88.

Venkatraman, S., Cheung, C. M. K., Lee, Z. W. Y., Davis, F. D., & Venkatesh, V. (2018). The 'Darth' side of technology use: An inductively derived typology of cyberdeviance. *Journal of Management Information Systems*, *35*(4), 1060–1091.

von Hentig, H. (1938). Limits of deterrence. *Am. Inst. Crim. L. & Criminology*, *29*, 555.

Wang, J., Li, Y., & Rao, H. R. (2016). Overconfidence in Phishing Email Detection. *Journal of the Association for Information Systems*, *17*(11).

Weick, K. E. (1989). Theory Construction as Disciplined Imagination. *Academy of Management Review*, *14*(4), 516–531.

West, R. (2008). The psychology of security. *Communications of the ACM*, *51*(4), 34–40.

Whitten, A., & Tygar, J. D. (1999). Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. *Proceedings of USENIX Security Symposium*.

Willison, R., Lowry, P. B., & Paternoster, R. (2018). A Tale of Two Deterrents: Considering the Role of Absolute and Restrictive Deterrence to Inspire New Directions in Behavioral and Organizational Security Research. *Journal of the Association for Information Systems*, *19*(12), 1187–1216.

Yoo, C. W., Goo, J., & H. Raghav Rao. (2020). Is Cybersecurity a Team Sport? A Multilevel Examination of Workgroup Information Security Effectiveness. *MIS Quarterly*, *44*(2), 907–931.

Zhu, J., Feng, G., Liang, H., & Tsui, K.-L. (2023). How Do Paternalistic Leaders Motivate Employees' Information Security Compliance? Building a Climate and Applying Sanctions. *Journal of the Association for Information Systems*, *24*(3), 782–817.

Zimmermann, V., & Renaud, K. (2019). Moving from a 'human-as-problem" to a 'human-as-solution" cybersecurity mindset. *International Journal of Human Computer Studies*, *131*(April), 169–187.