Fixing Security Together:

Leveraging trust relationships to improve security in organizations

Iacovos Kirlappos, Martina Angela Sasse
Department of Computer Science
University College London
United Kingdom
{i.kirlappos, a.sasse}@cs.ucl.ac.uk

Abstract. **Current approaches to information security focused on deploying security mechanisms, creating policies and communicating those to employees. Little consideration was given to how policies and mechanisms affect trust relationships in an organization, and in turn security behavior. Our analysis of 208 in-depth interviews with employees in two large multinational organizations found two trust relationships: between the organization and its employees (*organization-employee trust*), and between employees (*inter-employee trust*). When security interferes with employees' ability to complete work tasks, they rely on inter-employee trust to overcome those obstacles (e.g. sharing a password with a colleague who is locked out of a system and urgently needs access). Thus, non-compliance is a collaborative action, which develops inter-employee trust further, as employees now become "partners in crime". The existence of these two relationships also presents employees with a clear dilemma: either try to comply with cumbersome security (and honor organization-employee trust) or help their colleagues by violating security (preserving inter-employee trust). We conclude that designers of security policies and mechanisms need to support both types of trust, and discuss how to leverage trust to achieve effective security protection. This can enhance organizational cooperation to tackle security challenges, provide motivation for employees to behave securely, while also reducing the need for expensive physical and technical security mechanisms.**

Keywords — Trust; Information security management; Compliance; Security design

I. Introduction

Traditionally, information security seeks to mitigate security risks by implementing policies and technical mechanisms that specify employee behavior; policies also may threaten sanctions in case of non-compliance. The impact of this "comply-or-die" approach on day-to-day functioning of an organization is significant: organizations not only pay a cost for security mechanism operations, but also create constraints for honest employees seeking to perform well [1]. It slows down their production tasks, sometimes even completely blocking them, mostly due to security mechanisms and processes not being designed around employee needs and priorities [2-5]. This type of security fails to deliver effective protection even though it drains productivity. Recent industry data suggests that security breaches increase year on year (48% increase in the past year [6]), with a large part of those attributed to employee behavior [7].

The effects and potential benefits of trust on employee security behavior has not been explored. We already know that employees are emotionally attached to the organizations they work for [8][9]; they are also motivated and capable to protect them [4][10][11]. Despite this, in current security approaches, they are often treated as untrustworthy. Excessive monitoring and restrictions are put in place, "just in case employees turn bad in the future" [12]. But strict policies that cannot be followed means employees create *ad-hoc* security deployments that spin out of organizational control [10]. Also, by increasing restrictions, the organization and its security managers cannot reap the second-order benefits of trust, such as enhanced cooperation, goodwill development, and creativity to address organizational security challenges [13]. We argue that organizations' productivity and security can be improved by shifting thinking towards incentivizing trustworthy behavior, rather than restricting and controlling employee actions.

In this paper we present insights about security-related trust development in organizations. We analyze a set of interviews on employee security behavior, and we compare our findings to a framework of trust development [15]. We identify two different trust relationships emerging in organizational environments and present the effect of those on security behaviors. We then present the emerging conflicts between keeping the organization secure and preserving established trust relationships in the organizational environment. Finally we discuss how the resulting improved understanding on the role of trust in security behaviors can be used to provide effective information security management and contribute to the design and deployment of effective information security solutions.

II. Treating Users as a Problem: The Quest to Eliminate "Human Risk"

Information security breaches originating from human behavior lead people being described as the "weakest link" in the security chain [16]. The main human-related threats to security can be attributed to three areas: (1) human error leading to data leakages or creating vulnerabilities that can be exploited by attackers [17], (2) social engineering, where attackers psychologically manipulate people into performing security-compromising actions or divulging information [18] and (3)

USEC '15, 08 February 2015, San Diego, CA, USA
Copyright 2015 Internet Society, ISBN 1-891562-40-1
http://dx.doi.org/10.14722/usec.2015.23013

insider attacks, when employees intentionally exceed or misuse authorized levels of access to networks, systems, or data to steal confidential or proprietary information from the organization [19]. To counter these threats, organizations implement various assurance mechanisms (e.g. access control restrictions, anti-virus software, data loss prevention systems), combined with policy formulation and communication, aiming to render employees aware and able to behave in a secure way. When control is impossible (e.g. employee handling of confidential documents), monitoring and sanctions are introduced to deter misbehavior. Employees violating policies are threatened with potential reprimands that can be as serious as losing their job and legal action taken against them.

In theory, implementing extensive security controls and sanctions should prevent both intentional insider breaches and reduce the impact of erroneous behaviors that increase organizational security exposure. In practice, however it has been shown to have number of drawbacks:

- Human-related security risks are not well-defined. Exhaustively covering a wide range of behavioral scenarios requires implementation of architectural means to identify and eliminate of all potential vulnerabilities an organization may face is expensive and practically impossible. In a time when security budgets becoming tight [6], this can lead to organizations having to compromise with suboptimal solutions [1][14].
- Strict controls take away employee flexibility to respond to changing environments, reducing their ability to respond to non-predictable situations [20]. When employees engage with tasks where flexibility is required (e.g. remote or home working) control becomes time-wasting, inefficient or even impossible to implement. As a result, in order to maintain this flexibility, organizations end up relaxing security policies, weakening the system [21].
- Excessive assurance and sanctions can lead to employee dissatisfaction creating a value gap between the organization and its employees [22]. This hinders the development of social capital and shared values [23], leading to minimal incentive for secure behavior, and increased probability of insider attacks [19]. It also impacts the ability of the organization to retain its valuable employees; dissatisfaction can lead to them eventually leaving the organization [24].

All the aforementioned problems suggest that attempts to eliminate human-related security risks through assurance and sanctions essentially weaken organizational defenses. Security design needs to stop treating employees as intrinsically untrustworthy. Recent research has shown that most employees want to participate in security, and even create their own security solutions when the existing ones do not work [10] (e.g. sharing passwords to aid colleagues in urgent need for system access, changing those afterwards to reduce potential security risks). Building on this employee "propensity to do good", in the remainder of this paper, we make a case for trust as an important element of security design. We start by identifying two security-related trust relationships in organizations, presenting their impact on employee security behavior, how conflicts between the two relationships arise, leading to insecure behaviors, and how these conflicts can be addressed to create more effective security implementations.

III. A Framework for Trust

Mayer at al. [25] define trust as "*willingness to be vulnerable based on positive expectations about the actions of others*". It is only required in transactions where *risk* and *uncertainty* about the outcome exist, leaving a *trustor* (trusting actor) vulnerable to the *trustee's* (trusted actor) actions. Risk usually arises from the potential losses a trustor stands to suffers if the trustee does not behave as expected. Uncertainty arises from the lack of information available to the trustor about the ability and motivation of the trustee to fulfil in the transaction [15]. Despite the risk and uncertainty, a trustor makes themselves vulnerable to the trustee due to potential benefits by the trustee's later fulfilment. On a single transaction basis a trustee would be better off defecting after receiving the benefits of the trusting action: they have already received all the potential rewards from the transaction, so they are at a point of maximum gain, having invested minimal effort. Any effort to fulfil their part requires investment of additional resources that will reduce their net benefit compared to the pre-fulfilment state (Fig. 1, [26]).

$$Trustee's\ net\ benefit = Transaction\ Rewards - Effort\ invested\ for\ fulfilment$$

Fig. 1. Trustee benefit equation [26]

A. *Trust-warranting properties*

Fulfilment motivation for the trustee comes from the existence of *trust-warranting properties* [27], the long-term effects of which outweigh immediate non-fulfilment gains. Trust-warranting properties can be distinguished between *intrinsic* and *contextual* (Fig. 1).

1) Intrinsic properties: These are relatively stable attributes of the trustee that affect their *ability* and *motivation* for fulfilment in the trust transaction.

- *Ability:* Trustee's possession of the resources required for fulfilment (e.g. knowledge required on performing required security actions)
- *Motivation:* Fulfilment incentivised by factors internal to the trustee (e.g. personal costs of breaking trust). It is driven by *internalized norms* or *benevolence* that dictate doing what a trustee perceives to be "the right thing" and provide non-monetary fulfilment rewards to the trustee, like personal satisfaction.

2) Contextual properties: These constitute attributes of the context of the interaction (*temporal, social* and *institutional*) that provide motivation for trustworthy behavior by dis-incentivising non-fulfilment, leading to self-interested trustees fulfilling:

- *Temporal embeddedness:* The prospect of future interactions becomes an incentive for fulfilment [28]:

Non-fulfilment can damage future trust shown towards the trustee.

- *Social embeddedness*: When performance information about a trustee's past behavior can be shared amongst trustors, potential reputational damage due to fulfilment failure in a transaction can act as a fulfilment incentive [29].
- *Institutional embeddedness*: The presence of external enforcement third parties penalizing non-compliance also acts as a non-compliant deterrent for the trustee [30].

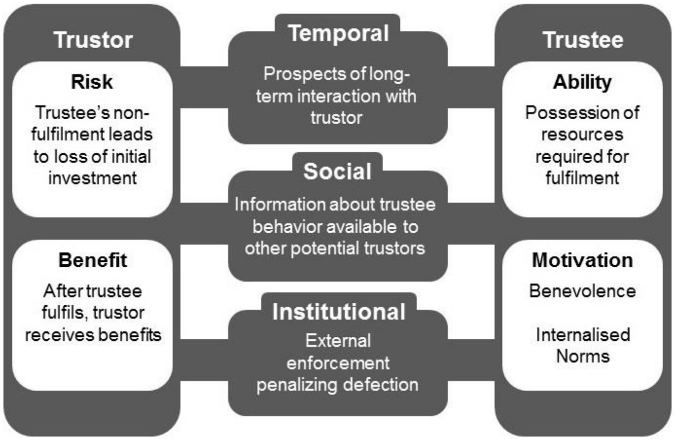
Trustor
Risk
Trustee's non-fulfilment leads to loss of initial investment
Benefit
After trustee fulfils, trustor receives benefits
Temporal
Prospects of long-term interaction with trustor
Social
Information about trustee behavior available to other potential trustors
Institutional
External enforcement penalizing defection
Trustee
Ability
Possession of resources required for fulfilment
Motivation
Benevolence
Internalised Norms

Fig. 2. Model of a trust interaction (adapted from Riegelberger at al. [15])

The important difference between intrinsic and contextual properties is to whom the trustor's trust is placed. Intrinsic properties result in *party trust* (i.e. to the trustee), while contextual lead to *control trust*: to the mechanisms that dis-incentivize non-fulfilment by the trustee [31]. Fulfilment due to control trust does not imply a trustee is trustworthy, but in most cases it is "*good enough*" – as it still leads to successful transactions [15].

IV. TRUST AND ORGANIZATIONAL SECURITY

Security (and the systems that come with it) is part of a wider socio-technical environment, where the primary goal is effective and efficient completion of production tasks [32][33]. Trust is a key element of that environment: it aids the development of social norms amongst employees, leading to improved collaboration and more effective production task completion [25]. Current uses of the term *trust* in security have a very narrow meaning though. *Trusted components* have been defined as *"systems or components whose failure can break the security policy"* [18]. But this definition refers to system components (hardware or software), certified to exhibit a specific behavior under specific conditions; this is assurance not trust. For trust to exist the organization should have no or little control over what employees can do, and, instead of stringent enforcement, choosing only to trust and encourage them to behave in a secure manner [21]. Based on the definition of trust as risk and uncertainty, the only component whose failure can break the security policy is the people who use those systems, as total control over them is impossible.

An attempt to move away from the restrictive understanding of the role of trust in security management was made by Flechais et al [21]. They used Riegelsberger et al.'s model [15] to discuss current trust placement in security implementations and identified a number of ways in which trust-warranting properties affect employee security behavior. Their first suggestion was that an organization should not aim to achieve total assurance if employees exhibit the intrinsic properties required to behave trustworthy. Well-trained employees (*ability*) that understand the risk mitigation effects of trustworthy behavior (*motivation*) are more likely to act in ways that protect the organization. On the other hand, intrinsic properties can also lead to trust violations: benevolence, social norms and expectation of future relationships can break the security policy. An employee's willingness to help a colleague locked out of a system by sharing their password may be stronger than their motivation to adhere to the security policy. This evolves over time, after a number of successful trust exchanges, and can be dangerous for the organization. Social engineering attackers, for example, can exploit this by pretending to be benevolent (e.g. selflessly helping to fix a problem on the target's PC which they created in the first place) [21].

Social embeddedness can also have a very powerful effect on security behavior: Weirich and Sasse [32] report that newcomers' security behavior follows that of members of their immediate work team, even after security training as part of their induction, as desire to "fit in" is usually stronger.

Temporal embeddedness, on the other hand, is used by the organization to reduce risks from employee behavior. Employees ready to leave a company, for example, may be willing to vandalize and cause damage to systems they have access to, since they have no expectation of future benefit from their employer and their colleagues. If they are leaving to join a competitor, they may even have reasons to break trust (e.g. stealing sensitive intellectual property information). Organizational "exit protocols" aim to eliminate this risk, making sure that people who are leaving the organization cannot exploit trust that was extended to them as employees.

Institutional embeddedness also acts as a non-compliance deterrent. The presence of organizations or institutions with power to sanction untrustworthy behavior (e.g. ethics committees or legislation) acts as a deterrent, usually depending on the type, strictness and severity of punishment (e.g. the threat of being excluded from a professional group).

Based on the aforementioned analysis, Flechais et al. [21] made a number of suggestions for improving security:

1) Simplify security: When a degree of flexibility is required, rigid policies cannot work because they are too complex, constraining or expensive, eventually exhausting employees' security *compliance budget* [3]. The only available option is to encourage and trust employees to behave in a secure manner, complementing this with monitoring to detect whether employees are actually complying with the policy.

2) Improve education: Security awareness and training should be given continuously to all employees, as opposed to just giving it to newcomers to improve motivation and ability.

3) Promote security culture: Ensure security policy is neither excessive nor unfair.

4) Participative security: Involve various stakeholders to increase perceived responsibility.

5) Foster group cohesion: Group people into security groups to improve social responsibility to contribute to security.

Despite their potential usefulness, the above suggestions are mostly empirical, based on "common knowledge" discussed in the context of Riegelsberger et al.'s framework, but not based on any data. As a result, the authors conclude that further research is required to examine the impact of trust on security behaviors and its effect on security implementations. This suggestion acted as a motivator for the research presented in the remainder of this paper.

V. Study Description

To understand the trust relationships developed through the interaction of employees with security systems, we conducted a secondary analysis on a set of interviews with employees of two large multinational organizations. We had 120 interview transcripts available form the first organization and 88 from the second one. The interviews were semi-structured and conducted on a one-to-one basis by a team of seven researchers (including both the authors). Each lasted approximately 50 minutes, allowing for elicitation of a suitably rich representation of the employee experience of security. Participants were recruited via the company email newsletter, sent to all employees. They held various lower-level and lower to middle management positions within a number of organizational divisions, including customer service, marketing, administration, finance, procurement and IT. Participation was anonymous and participants were given an informed consent form, assuring that they would not be identified or followed up. After the interview, participants were paid the equivalent of $40. We did not encourage participants to tell us about security infractions, but simply asked about their awareness of, and experience with, a set of corporate security policies. The structure of interviews touched upon aspects of security awareness and compliance, including:

a. What is the employee perception of how security impacts their role?
b. What do employees appreciate in terms of organizational support for security?
c. Where employees exercise non-compliance as a response to shortcomings or frictions in the organizational security experience, what conditions led to those behaviors divergent from organization policy?

Interviews were recorded, transcribed, and a Grounded Theory analysis [34] was conducted, using Atlas Ti. The two authors independently coded an initial set of ten interviews and a codebook was devised. This was then used for the full analysis of all the interviews by one of the authors, aiming to capture the different roles trust plays in the deployment of organizational security. The results of this analysis are presented in the next section (primary analysis of the interview data for one of the two organizations was published in [10]). It is also important to note here that employee responses and the corresponding emerging behavioral patterns, presented in the next section, were consistent and of similar nature in both the organizations we studied.

VI. Understanding Organizational Trust Relationships

From our analysis we identified the presence of two different security-related trust relationships that affect the design, implementation and evolution of security behaviors in an organization: (1) *organization-employee trust*: organization's dependency on employees behaving securely and (2) *inter-employee trust*: employee dependency on each other that affects their security behavior.

A. Organization-employee trust

Employees reported that, based on the existing security implementation, the organization's security appears to significantly depend on them behaving in a trustworthy way:

P143: "*We tend culturally to be allowed more freedom and responsibility than some people might do*."

P200: (talking about security implementers) "*so long as we tell them that we're going to do something, they'll trust us to do it, they won't necessarily come along and sort of sit behind you and make sure you're actually implementing that piece of design.*"

They also recognized organizational vulnerability to potential misbehaviors of its employees:

P143: *"It's almost impossible in security terms to stop a human actually attaching a document when they shouldn't it's very difficult to get round that."*

Based on the above, we define organization-employee trust as: "*The level of organizational dependency on the actions of employees that the existing security implementation creates".* This type of trust was present in situations when the organization trusted its employees to behave in a trustworthy way, thus not restricting their actions, essentially remaining vulnerable to their potential misbehaviors. The development of employee-organization trust is based on both intrinsic and contextual properties (Fig. 3):

1) Intrinsic: Employees possess the knowledge and risk awareness required to take actions to protect the organization (*ability*) (e.g. P137: "*we are using a lot data and we know the impact that has on the company and the customers if that gets into the wider domain*") but also show a propensity to do good (*motivation*): P178: "*So maybe I'm not the right person to take those risks and make those choices, but I think we all have to share that that's part of the ethos of the company*."

2) Contextual: Employees need to comply with organizational policies to avoid any sanctions (*temporal incentives*). P4: "*…They do it only because "Oh, I might get into trouble if I don't do it*".

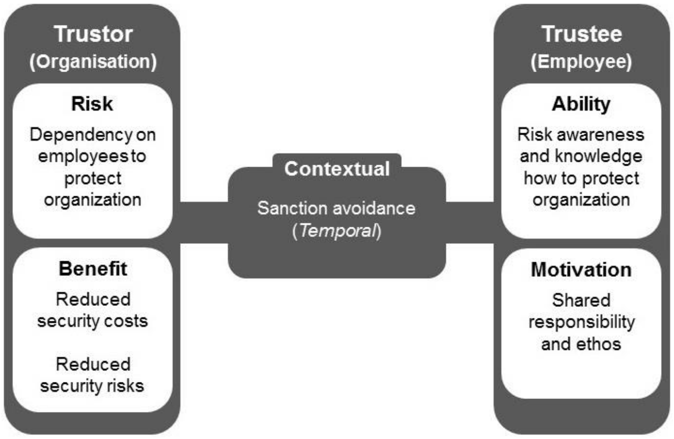

Fig. 3. Organisation – employee trust development incentives

Despite the existence of both intrinsic (*ability, motivation*) and temporal incentives to comply with security, employees reported a number of security violations driven by the current configuration of organizational systems not allowing efficient production-related task completion. For example:

P2: "*…they'll take their company computer off the proxy, and while you're off the proxy, go home, log in, grab the files, save them, come back in, you're back on the proxy, you're okay*."

In general employees appeared to recognize the reliance of the organization's security implementation on their behavior. In addition they were motivated to behave in a secure way, based on both intrinsic and contextual incentives that appeared to be driving this behavior. Despite that, we identified a wide range of security violations by employees, routed back to three main reasons: (1) problematic security implementation, (2) inaccurate user risk perception and awareness, and (3) the need to develop or preserve existing inter-employee trust relationships prevailing over the need to act in a way that preserved employee-organization trust. In the next section we focus on how the third factor (preserving inter-employee trust) leads to security violations (for more examples of a wider range of violations relating to (1) and (2) please refer to [10]).

B. *Inter-employee trust as a non-compliance motivator*

We identified many cases where employees either explicitly reported the presence of trust as a driver for non-compliance, or discussed how the close relationship they have with their colleagues lead to them not following the prescribed security practices.

P123: "*You work with them so much. God, the engineers that I work with for our company I spent hours with them on a daily basis, so you do get to know them very well.*"

I: "*So if someone asked you to share your password with them, you'd have no problem with that?*" P126: "*No, as long as it's a trusted colleague.*"

Existence of trust amongst employees acts as an enabler to productivity:

P57: "*I mean SharePoint is good, but you have to be trusting for users to use it properly*."

Based on this, we define inter-employee trust as: "*The willingness of employees to act in a way that renders themselves or the organization vulnerable to the actions of another member of the organization".* It can be developed both inside and outside the security domain, and leads to behaviors that diverge from the security policy. A few examples:

P2 (on password sharing): *I have some level of trust with them, it's more if they have enough level of trust with me to be "Okay, here's the thing so you can log in and do it quick. I'll change it as soon as I come back so that we're secure and all that but I need you to keep working to get the job done*."

P149 (on sharing documents through non-official communication channels): "W*ell if someone's into the company and they need a certain document they know where to find it then pass it on.*"

P120 (on not locking their screens): "*…because when you comment on it and say "Well you should actually be locking your screen when you walk away", the comment you get back is the fact that "Well you know we should be able to trust people around.*"

As the interview extracts show, employees appear willing to knowingly diverge from recommended practices: they disclose information or perform actions for which they could be held accountable, either because they need to help a colleague in need (e.g. share a password or information) or due to trusting people around them (e.g. leaving their laptops unlocked or letting them tailgate). Essentially, when security creates problems, employees turn to their trusted colleagues. They use a resource readily available (inter-employee trust) to cope with over-restrictive mechanisms that hinder both their individual ability to do their job but also improve ability to collaborate with their colleagues. An employee who was locked out of a system by entering their infrequently used password incorrectly and who cannot access the helpdesk immediately can easily borrow a trusted colleague's password to fetch some information they need urgently. Willingness to help a colleague, recognition that they may end up in the same situation in the future and the overall desire to be part of the overall organizational social environment provide enough incentives for a colleague to help them, even if that means breaking the security policy.

Similarly to organization-employee trust, the development of inter-employee trust is also based on both intrinsic and contextual properties (Fig. 4):

1) Contextual: Successful employee collaboration results in increased willingness to collaborate in the future (*temporal embeddedness*) P123: *"I spent hours with them on a daily basis, so you do get to know them very well.",* but also in increased feeling that collaborators are members of the same social group *(social embeddedness)* P191: "*Yeah if it's someone within the team then they can be trusted, yeah. I wouldn't do it for anyone external to the company*".

2) Intrinsic: Employees feel the need to help someone in need within their social environment (P31: *"…there's a policy that, shortly after we moved to this building they made a big* *deal out of “Don’t allow following access through doorways.” […] it seems kind of impolite to say, Sorry, I can't let you through, I’m going to have to slam the door in your face. Human nature tends to be I’ll hold the door for you.*”. They even perceived their colleagues as collaborators that collectively preserve organization-employee trust, affirming their ability to protect the organization (P80: “*So when you’re an employee and I speak to another employee, it’s not a problem, because everybody knows what the security policies are with the company*”).

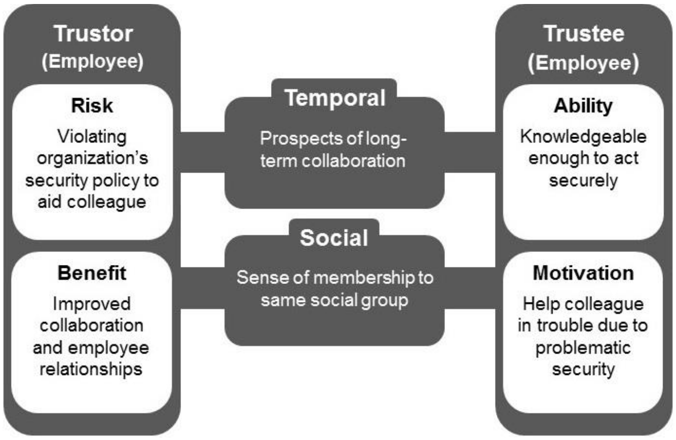
Trustor
(Employee)
Risk
Violating organization’s security policy to aid colleague
Benefit
Improved collaboration and employee relationships
Temporal
Prospects of long-term collaboration
Social
Sense of membership to same social group
Trustee
(Employee)
Ability
Knowledgeable enough to act securely
Motivation
Help colleague in trouble due to problematic security

Fig. 4. Inter–employee trust development incentives

C. *Conflicts of two trust relationships*

Our findings suggest that employees appear sufficiently able and motivated to behave securely, but sometimes their relationship with their colleagues prevails over the need for security. Many of the non-compliance scenarios we identified were related to security policies essentially asking employees to distrust their colleagues (e.g. no password sharing, lock screen, no tailgating). Employees then need to prioritize on which of the two trust relationships they need to preserve, ending up breaking organization-employee trust to help improve or preserve inter-employee trust. This leads to the emergence of two different types of organizational security: one defined in the policy and one devised by employees on an ad-hoc basis, based on their interaction with information security mechanisms and perceived risks (*Shadow Security*, [10]). When this happens, security inevitably spins out of organizational control, leaving the organization vulnerable to behaviors out of sight of security managers. We identified a few cases of this behavior. For example the self-devised security mechanism of this employee when working with contractors that needed wider access than what they already had:

P12: “*…as far as “Oh, this contractor wants to get something done quick, here use my ID for doing that. You know, and then I’ll switch the password after. Okay, he’s sort of protected it, but really is, you’ve just shared your ID, you just shared a password, and with a non-company person, you know, violation, but you need to get your work done.*”

Another example highlighting the presence of a trust conflict is contractor access and interaction with permanent employees. Contractor motivation for trustworthy behavior can sometimes be lower:

P70: “*…before when we had like our group meetings, even though we were contractors, we were also allowed in those meetings, but as an employee, I felt myself to be more a part of that group now, because now we belong to the company*”

Despite the potential of this resulting to insecure behaviors and organization-employee trust violations, contractors are treated by employees like everyone else:

P9: “*…when that contractor was still there I wasn’t told to treat contractors differently.*”

In addition, employees talked about colleagues leaving and rejoining as contractors – should they now treat them differently?

P41: “*I saw this co-worker who was being hired as a contractor and I called the woman […] and she calls me back in a couple of days from now and she says “I checked it out and found out that that wasn’t really another person, that was me…*”

Expecting employees to treat a colleague differently in such a case would be unrealistic. The existing trust relationships amongst them are inevitably going to prevail, despite awareness that this may increase organizational exposure to security risks.

Flechais et al. correctly predicted another problem of inter-employee trust development: over long periods of time, trust turns to reliance, with employee work processes *depending* on collective trust violations (e.g. P90: “*…a lot of times the field guys, they won’t tend to trust you initially you’ve got to be there for a while. Like now that I’ve been here three years, “Oh, I’ve worked with him a lot. Not a problem, I like working with him.*””).

D. *The need for assurance*

Our analysis also confirmed another suggestion by Flechais et al. [21]: despite violating security to help their colleagues, employees appear to appreciate the need for the organization to put some controls and limitations in place. For example:

P102: “*…I think there’s a balance to be struck between giving people trust and appreciating their common sense and their intelligence and also protecting one’s system from the occasional stranger who walks through the area.*”

Employees also appeared to recognize the consequences of breaking organization-employee trust:

P193: “*The trust has always been there, but the consequences are also there if it’s broken.*”

In summary, employees appear sufficiently motivated and able to behave in a trustworthy way, honoring the trust shown to them by the organization. They even recognize the need for security mechanisms that limit their actions in order to protect the organization and the existence of consequences for trust violations. But when security becomes over-restrictive, impacting their ability to proceed with their day-to-day tasks or help their colleagues, they are willing to break this trust relationship to help develop, improve or maintain inter-employee trust that appears to be important and widely prevalent amongst them.

VII. Leveraging trust relationships to improve security implementations

Trust-driven security violations create significant risks for the organization. Dependency on collaborative non-compliance accentuates the perceived need to preserve inter-employee relationships, leading to security violations becoming a norm negatively impacting the development of long-term information security culture [35]. It also leads to the development of social capital [30] amongst employees grounded on collective security violations increasing organizational exposure to social engineering: employees willing to violate security to help their colleagues can be more vulnerable to attacks by impostors that rely on their willingness to share information through informal channels.

Security designers need to understand that - in a highly social environment like a large organization - inter-employee trust is often more important to employees than complying with security. This should not be used as a pretext to treat employees as untrustworthy though: employees possess both the ability and motivation to exhibit trustworthy behavior as long as their ability to complete their primary tasks is not significantly hindered by security. Organizational security design should recognize the existence of the trust relationships identified in this paper and their importance for the success of any security implementation. In the remainder of this section, based on our improved understanding, we discuss what organizational security designers could do in order to accommodate for organizational security trust relationships in security design.

A. *Know when to trust and when assurance is necessary*

Despite our findings suggesting that current security implementations heavily rely on employees behaving in a trustworthy way, this has not been formalized by organizations and it is not reflected in current information security management strategies. Organizational insistence on using technical mechanisms and sanctions to deter untrustworthy behavior has a number of negative effects:

1. It damages employee ability and motivation to comply with security: the perceived lack of organization-employee trust, materialized through unrealistic security expectations, decreases employee ability to comply with security policies. It also creates resentment, increasing employee incentive for collaborative non-compliance, based on inter-employee trust. Prolonged resentment is dangerous [36]: it increases the risk for insider attacks and loss of valuable human capital, with disgruntled employees leaving the organization.

2. Collaborative non-compliance also encourages disregard for security in general and non-compliance becoming the habitual behavior amongst employees. The emerging security behaviors, may not manage risks effectively, due to inaccurate employee understanding of security risks and countermeasures, but are the best available actions employees can use to proceed with production tasks. Low appreciation for the need for security can create a non-compliant security culture within the organization [37], also resulting in the organization losing track of employee actions, thus increasing the security risks it is exposed to.

3. Once a security culture is developed based on collaborative security violations, new employees that try to "fit in" and participate in inter-employee trust development are more likely to follow suit to their colleague's non-compliance.

To reduce the impact of the aforementioned problems, information security should take advantage of both users and technology to achieve effective protection. To achieve this, organizations need to know "when to trust" and "when to assure", supporting correct trust development through:

1) Simplification of security mechanisms: Security design should be designing around *security hygiene* [23]; rules should not be broken for productivity reasons. Security mechanisms need to be designed around employee production tasks to reduce the need for trust violations for the sake of productivity, but also reflecting the trustworthiness an organization should show towards its employees. To eliminate the need for password sharing for example, the organization should create mechanisms that provide quick account creation for employees that need to access to new systems.

2) Include trust in security communication: When security mechanisms are implemented in a way that encourages trustworthy behavior, Security Awareness, Education and Training campaigns (SAET) should be used to create new behaviors better connected and adjusted to the actual risks the organization faces not employee-perceived ones [38]. For example, in one of the two organizations we examined, home working is quite prevalent, with many people being either full time home workers or working from home two or three times a week. This makes it impossible for the organization to restrict employee actions, so they need to be trusted not to violate security. This organizational dependency on employee behavior should be communicated to them to explain their importance and responsibility in keeping the organization secure. Communication should: (1) Make it clear to employees that they are trusted and supported in their security decisions (to improve *motivation*), also explaining the "It's business, not personal" need for security vigilance and (2) Include information on current threats and how real-world trust development signals break down when using computer systems (improving *ability*).

3) Knowing when to assure: Once usable mechanisms and effective security communication are in place, clever controls are required to balance trust and assurance. Risk-aware employees that interact with well-designed security mechanisms, no longer have reasons to violate security, unless they are doing so for malicious purposes (e.g. when the rewards from not playing by the rules are significantly higher than the consequences of not doing so). In such cases, violations can be detected by improving current monitoring implementations to include contextual information on users to detect employee trust abuse and precursors of insider attacks. An employee that has full, uninterrupted access to corporate fileservers, should not be downloading vast amounts of information locally on their computer. This action on its own may not constitute an offence, but it provides sufficient grounds for further investigation. Malicious actions should then be followed up with serious consequences that are visibly enforced. Visible enforcement can act both as a deterrence for future misbehavior and as a motivation improver, reminding to employees that they are trusted and responsible to keep the organization safe. It also has to be clear to employees that it is not themselves who are deemed as untrustworthy, but assurance is in place to protect the organization.

B. *Once developed - don't enforce it!*

The first assessment of an employee's trustworthiness comes even before they join the organization through background checks and vetting procedures during recruitment. These process uses past behavior as an indicator of potential future actions. The need for this confirms our suggestions that total assurance is impossible and ineffective; otherwise screening would be unnecessary. When the organization establishes that ability and motivation to behave in a trustworthy way are present, there's no need to over-assure. Employees that pass the screening process should be considered trustworthy and treated as such instead of being subject to continuous restrictions. The consequent visible presence of trust towards employees can inject trust in the organization-employee psychological contracts that dictate organizational employee behavior, leading to cooperation that benefits everyone in the organization:

- People in an organization develop shared values and a shared-sense of responsibility for the well-being of the organization based on shared formal or informal norms promoting cooperation [39][40], which also affect their security behavior [37]. Secure behavior should be driven by a feeling of contribution to common organizational interests, rather than rule-driven actions to avoid sanctions.

- The need for enforcement of problematic security will be reduced. This will then reduce the 'noise' introduced by productivity-driven 'legitimate' violations. The resources saved from reduced noise can be reinvested in implementing other more effective security mechanisms, enabling the implementation of clever monitoring to identify serious malicious activity (insider or outsider attacks) [41]. Precursors of serious attacks (e.g. intellectual property theft, currently accounting for less than one per cent of all cybercrimes, also resulting in more than 50% of the monetary losses [42]) may be currently lost in false-positive alarms when employees violate security for productivity or collaboration reasons.

- Flexibility strengthens employee ability to defend the organization. Attackers are likely to adapt to new technologies, but attacks are much harder to succeed with suspicious employees, motivated to protect the organization and a culture that favors secure behavior This is not uncommon in other security implementations: for example biometrics at passport control are considered to be more effective than individuals but when a problem is identified a human can take over and use a much richer and broader set of factors from the context of the environment to assess a passenger's trustworthiness [21].

Potential for increased organizational reliance on employee security creates another problem though: regulation and international standards currently advocate against it. Suggestions and regulatory standards to organizations include security practices that have proven to be insecure (e.g. ISO27002 security standard advocates for frequent password changes to improve security [43], despite sufficient evidence for the contrary [5]). Both researchers and practitioners need to push for changes in regulation and standards to be up to date with latest security research findings and the needs of modern organizations. No organizational resources or user effort should be expended on implementing solutions that offer no security benefits.

C. *Accommodate urgency, encourage self reporting and follow it up*

Organizations also need to create mechanisms for unusual circumstances. Non-fulfilment cannot be totally eliminated, as this is both uneconomical and prohibitive for productivity, but enhancement of the employee-organization trust relationship can ensure that it happens less often and does not go unreported. Organizations who acknowledge the fact that employees may, under *rare* and *unusual* conditions, have to circumvent security for productivity reasons, should also implement mechanisms where employees can report those non-compliance instances. Clear instructions should then exist how to deal with any potential vulnerabilities. For example, an employee who shared their password with a colleague in an emergency situation should recognize this as a violation, then login to a non-compliance log system and report the behavior. The same could apply to physical access control: an employee who forgot their pass should be easily able to get a daily pass through a simple verification process. In both cases, the organization should encourage self-reporting by communicating that no action will be taken against employees who self-report, while those who do not should be susceptible to sanctions. The organization should also ensure adequate measures were taken to close the resulting loophole (e.g. employee changed their password within two hours). Accommodating for urgency should not be implemented as a substitute to usable systems though. Violations, even reported ones, need to be infrequent enough to avoid non-compliance becoming part of employee security culture, also avoiding introducing significant overheads in terms of resources required to close the loopholes created by frequent circumventions.

D. *Promote collective and participative security*

The improved understanding of the role of trust that emerges from our findings, together with a participatory approach to security design, can enable more effective design of security mechanisms. Inclusion of users in security design can increase their motivation to comply [21][44]. It is definitely possible to get help with tailoring security by smaller organizational sub-divisions, taking advantage of the already existing inter-employee trust. Line managers, who have a considerable influence on their staff's security decisions can also elicit feedback from them on security challenges [10]. Bringing security to the table at group meetings can lead to increased awareness amongst employees and ability to connect with the risks presented by their managers or colleagues. The emergent *participatory security* environment can increase their sense of contribution and ownership of security implementation, triggering internalized norms and benevolence-related compliance.

VIII. CONCLUSION

Effective security needs to strike a productive balance between trust and assurance. Our findings suggest that employees possess both the ability and motivation to behave securely, honoring the trust shown towards them by the organization (*organization-employee trust*), also aided by contextual motives to do so. But when security comes to conflict with other parts of their work and relationships with their colleagues, non-compliance becomes their only option to preserve the existing trust relationships in the social environment of the organization (*inter-employee trust*). To reduce this conflict security management needs to take advantage of trust and aid in its development, refraining from overly assuring once trust is developed. Employees that have been screened, trained and understand the risks of insecure behavior should not need to choose between organization-employee trust or inter-employee trust when interacting with security mechanisms: both trust relationships contribute to the organization achieving its productivity targets while remaining secure. Security design that accommodates for this can lead to the creation of a high-trust/low-assurance environment which can introduce significant economic benefits for organizations: compliance coming from employees motivated to behave securely, not forced to do so, reduces the need for expensive physical and technical mechanisms.

IX. RELATED WORK

The productivity benefits of trust have been identified in non-security related contexts: trust between members of an organization leads to highly cooperative behaviors, acting as a substitute for control [45]. In addition the more connected employees feel with an organization, the more committed and involved they are [46]. Our findings strongly suggest that effective security design can achieve similar benefits in a security context, leveraging the identified existing trust relationships to provide effective security. Our findings also provided support for the suggestions made by Flechais et al. [21] (presented in Section IV): trustworthy behavior can be supported through improved employee-centric security design and effective communication, improving organizational security culture and creating an environment where employees collaborate to keep the organization secure.

X. FURTHER WORK

We believe our findings present a good starting point to understand how trust plays a crucial role in protecting an organization. There are further areas to be explored though. Outsourcing, for example, is increasingly popular amongst large organizations, in an attempt to save money and also get better service from bespoke providers. To our knowledge, its impact on trust development and its effects on security have not been studied to date. In the future we aim to investigate further how outsourcing and other changes in the organizational environment (e.g. increased amount of employees working from home and schemes like Bring Your Own Device) affect security-related trust relationships. Another interesting suggestion for further research in the one by Flechais et al. that creating simple, reliable means of mutual authentication for employees to authenticate to each other can solve the problem of social engineering. Unfortunately we found no evidence to support that in our analysis, but it definitely deserves to be part of further future research on the subject. Security self-reporting by employees also deserves to be researched further: it is interesting to evaluate whether employees will be willing to report violations, if an organization provides sufficient assurance that it will have no negative impact on them.

XI. ACKNOWLEDGEMENTS

We thank IT security managers at our partner organizations for facilitating the interviews with employees and the researchers who took part in data collection, conducting some of the interviews.

XII. REFERENCES

[1] F. Pallas. "Information Security Inside Organizations-A Positive Model and Some Normative Arguments Based on New Institutional Economics." Available at SSRN 1471801, 2009

[2] A. Adams and M. A. Sasse. "Users are not the enemy". In Communications of the ACM, 42(12), pp. 40-46, 1999.

[3] A. Beautement, M. A. Sasse and M. Wonham. "The compliance budget: managing security behaviour in organizations". In Proceedings of the 2008 New Security Paradigms Workshop pp. 47-58. ACM, 2008.

[4] I. Kirlappos, A. Beautement, and M. A. Sasse. ""Comply or Die" Is Dead: Long live security-aware principal agents." In Financial Cryptography and Data Security, pp. 70-82. Springer Berlin Heidelberg, 2013.

[5] C. Herley. "So Long, and No Thanks for the Externalities". In New Security Paradigms Workshop (NSPW), 2009.

[6] PWC. "The Global State of Information Security Survey 2015". Retrieved from: http://www.pwc.com/gx/en/consulting-services/information-security-survey/index.jhtml2013

[7] "US State of Cybercrime Survey," In CSO Magazine, U.S. Secret Service, Software Engineering Institute CERT Program at Carnegie Mellon University and Price Waterhouse Cooper, April 2014

[8] D. M. Rousseau. "Psychological and implied contracts in organizations." In Employee responsibilities and rights journal, no. 2: 121-139, 1989.

[9] L. F. Love and P. Singh. "Workplace branding: Leveraging human resources management practices for competitive advantage through "best employer" surveys." In Journal of Business and Psychology, 26(2), 175-181, 2011.

[10] I. Kirlappos, S. Parkin, and M. A. Sasse. "Learning from "Shadow security": Why understanding non-compliant behaviors provides the basis for effective security", in press, 2014.

[11] J. Blythe, R. Koppel, and S. W. Smith. "Circumvention of security: Good users do bad things." In Security & Privacy, IEEE, 11(5), 80-83, 2013.

[12] G. J. Silowash, D. M. Cappelli, A. P. Moore, R. F. Trzeciak, T. Shimeall, and L. Flynn. "Common sense guide to mitigating insider threats", 2012.

[13] C. Handy. “Trust and the virtual organization.” In Harvard Business Review 73 (3), 40–50, 1995.

[14] F. Björck. “Security Scandinavian style”. PhD diss., Stockholm University, 2001.

[15] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy. “The mechanics of trust: a framework for research and design.” In International Journal of Human-Computer Studies, 62(3), 381-422, 2005.

[16] B. Schneier. “Secrets and lies: digital security in a networked world.” Wiley, 2000.

[17] “IBM Security Services 2014 Cyber Security Intelligence Index.” Retrieved from: http://media.scmagazine.com/documents/82/ibm_cyber_security_intelligenc_20450.pdf

[18] R. J. Anderson. "Security Engineering: A Guide To Building Dependable Distributed Systems”. Wiley, 2010.

[19] A. P. Moore, D. Cappelli, T. C. Caron, E. D. Shaw, D. Spooner, and R. F. Trzeciak. “A preliminary model of insider theft of intellectual property”, Technical Report, Carnegie Mellon University, 2011.

[20] C. Vroom and R. Von Solms. Towards information security behavioural compliance. In Computers & Security, 23(3), 191-198, 2004.

[21] I. Flechais, J. Riegelsberger, and M. A. Sasse. “Divide and conquer: the role of trust and assurance in the design of secure socio-technical systems.” In Proceedings of the 2005 workshop on New security paradigms (pp. 33-41). ACM. 2005.

[22] E. Albrechtsen and J. Hovden. “The information security digital divide between information security managers and users”. In Computers & Security 28(6), pp.476-490, 2009.

[23] I. Kirlappos and M. A. Sasse. "What Usable Security Really Means: Trusting and Engaging Users." In Human Aspects of Information Security, Privacy, and Trust. Springer International Publishing, 2014. 69-78.

[24] Ken Blanchard, “Building Trust”, Ken Blanchard companies, 2010, retrieved from: http://www.kenblanchard.com/img/pub/Blanchard-Building-Trust.pdf,

[25] R. Mayer, J. Davis and F. D. Schoorman. “An integrative model of organizational trust.” In Academy of Management Review, 20 (3), 709-734, 1995.

[26] M.A. Sasse, I. Kirlappos. “Design for Trusted and Trustworthy Services: Why We Must Do Better”. In *Trust, Computing, and Society* (pp.229-249). Cambridge University Press, 2014.

[27] M. Bacharach and D. Gambetta. “Trust in signs.” In Trust in society, 2, 148-184, 2001.

[28] R. Axelrod. “More effective choice in the prisoner's dilemma.” In Journal of Conflict Resolution, 24(3), 379-403, 1980.

[29] C. L. Corritore, B. Kracher and S. Wiedenbeck. “On-line trust: concepts, evolving themes, a model.” In International Journal of Human-Computer Studies,58(6), 737-758, 2003.

[30] B. Schneier. “Liars and outliers: enabling the trust that society needs to thrive." John Wiley & Sons, 2012.

[31] Y. H. Tan and W. Thoen. “Toward a generic model of trust for electronic commerce.” In International Journal of Electronic Commerce, 5, 61-74, 2001.

[32] D. Weirich and M. A. Sasse. “Pretty Good Persuasion: A first step towards effective password security in the real world.” In NewSecurity Paradigms Workshop 2001.

[33] S. Brostoff and M. A. Sasse. “Safe and Sound: a safety-critical approach to security design.” In New Security paradigms Workshop 2001.

[34] A. Strauss and J. Corbin. “Basics of qualitative research: Techniques and procedures for developing grounded theory”. Sage Publications, Incorporated, 2007.

[35] A. Da Veiga and J. H. P. Eloff. “A framework and assessment instrument for information security culture.” In Computers & Security, 29(2), 196-207, 2010.

[36] E. W. Morrison and S. L Robinson. “When employees feel betrayed: A model of how psychological contract violation develops”. In Academy of management Review, 22(1), 226-256, 1997.

[37] S. L. Pfleeger, M. A. Sasse, and A. Furnham. "From Weakest Link to Security Hero: Transforming Staff Security Behavior." In Journal of Homeland Security and Emergency Management 11.4 (2014): 489-510

[38] B. Bulgurcu, H. Cavusoglu and I. Benbasat. “Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness”, In *MIS Quarterly*, (34: 3) pp.523-548, 2010.

[39] F. Fukuyama. “Social capital, civil society and development.” In Third world quarterly, 22(1), 7-20, 2001.

[40] P. Resnick. “Beyond bowling together: Sociotechnical capital.” In HCI in the New Millennium, 247-272, 2001.

[41] D. Caputo, M. Maloof and G. Stephens. “Detecting insider theft of trade secrets”. In Security & Privacy, IEEE, 7(6), 14-21, 2009.

[42] R. R. Rantala, “Cybercrime against Businesses, 2005”, Bureau of Justice Statistics Special Report, Sept. 2008; http://bjs.ojp.usdoj.gov/content/pub/pdf/cb05.pdf

[43] ISO and I.E.C. Std. “ISO 27002: 2005. - Information Technology-Security Techniques-Code of Practice for Information Security Management”. ISO (2005).

[44] S. Bartsch and M. A. Sasse. “Guiding Decisions on Authorization Policies: A Participatory Approach to Decision Support”. In ACM SAC 2012, Trento, Italy, 2012.

[45] A. C. Costa, R. A. Roe and T. Taillieu. “Trust within teams: The relation with performance effectiveness”. In European journal of work and organizational psychology, 10(3), 225-244, 2001.

[46] A. Bussing. “Trust and its relations to commitment and involvement in work and organisations”. In SA Journal of Industrial Psychology, 28(4), p-36, 2002.