Information Systems Security Policy Enforcement With Technological Agents: A Field Experiment

Frank Hadasch  
University of Mannheim, Mannheim, BW, Germany, hadasch@es.uni-mannheim.de

Ye Li  
University of Mannheim, Mannheim, BW, Germany, yeli@es.uni-mannheim.de

Benjamin Mueller  
University of Mannheim, Mannheim, BW, Germany, mueller@es.uni-mannheim.de

Follow this and additional works at: http://aisel.aisnet.org/ecis2013_cr

Recommended Citation
http://aisel.aisnet.org/ecis2013_cr/113
Information Systems Security Policy Enforcement with Technological Agents: A Field Experiment

Hadasch, Frank, Chair of Information Systems IV, Institute for Enterprise Systems, University of Mannheim, L 15, 1-6, 68131 Mannheim, Germany, hadasch@es.uni-mannheim.de

Li, Ye, Chair of Information Systems IV, University of Mannheim, L 15, 1-6, 68131 Mannheim, Germany, yeli@es.uni-mannheim.de

Mueller, Benjamin, Chair of Information Systems IV, University of Mannheim, L 15, 1-6, 68131 Mannheim, Germany, mueller@es.uni-mannheim.de

Abstract

This paper addresses the question of employees’ information protection behaviour from a socio-technical perspective. While prior information systems (IS) security research thoroughly examines the determinants of employees’ security motivation, this work investigates if information protection motivation translates into actual behaviour and how security policy-enforcing technological agents influence behaviour. For researchers this is important as measurements of actual security behaviour are scarce in the literature, the underlying cognitive process of information protection behaviour lacks detailed examination, and the design of information technology (IT) artefacts requires an understanding of what makes them effective in changing an employee’s information protection behaviour. In a scenario-based field experiment with 82 participants we test for direct and moderation effects of policy-enforcing technological agents on employees’ information protection behaviour. Captured screen recordings of the simulated work environment are analysed with principles of inductive reasoning to suggest a cognitive process model of users’ information protection behaviour. Shedding light into the black box of the post-motivational phase helps us to investigate when employees are specifically challenged to protect confidential information and how a policy-enforcing technological agent might help to prevent information leakage incidents.

Keywords: IS security behaviour, leakage prevention, policy enforcement, field experiment.

---

1 This work was funded by the DFG (German Research Foundation) under the project User-Centric Information Flow in Enterprise Systems (USIFES) as part of the priority program 1496 “Reliably Secure Software Systems - (RS3)”. One student, V. Lux, contributed to this work by developing the technological agent and supporting the data collection. We thank him very much. We would also like to thank the three anonymous reviewers and the track chair for their helpful comments.
1 Introduction

Employees using information technology (IT) to handle confidential company information face the challenge of distinguishing between secure and insecure actions. An accidentally performed insecure action may allow unauthorized individuals to access confidential information. An elucidating incident is the accidental disclosure of social security numbers affecting 3.5 million U.S. citizens (Shannon, 2011). In this case, an employee moved unprotected confidential information without any malicious objectives to a publicly accessible internet domain. Practitioners and researchers acknowledge the threat of accidental leakage and the wide-ranging societal implications (Warkentin and Willison, 2009).

In today’s digital economy with borderless operating environments the risk for accidental spreading of confidential information is increasing. Therefore, companies’ executives are interested in assessing the effectiveness of appropriate technical countermeasures. Information leakage prevention (ILP) tools represent IT artefacts that control how information is used after access is granted (Pretschner et al., 2006; Sandhu and Park, 2003). Also known as endpoint security, the primary objective of such tools is to monitor events on end users’ computers and interrupt in case critical events are detected under certain conditions. The practical relevance of these tools is highlighted in a recent global information security survey. Ernst & Young (2011) report that the top management of 1,700 surveyed companies plans to spend more on prevention technologies this year than in previous years. Although the effectiveness of such technologies in influencing end users’ behaviours is critical for executives’ investment decisions, it is under-examined.

In academic research the importance of information security in organizations is widely acknowledged. Several researchers highlight the need for studying security with socio-technical approaches (Anderson and Agarwal, 2010). To understand the social part of the socio-technical system, end users’ intentions have been extensively investigated. Based on the theory of planned behaviour (Ajzen, 1991), intentions are argued to summarize motivations to perform particular behaviours. In existing socio-technical security research, various forms of intentions (e.g., policy-compliance intention, policy-violation intention, intention to use protective information technologies) can be explained by drawing on theories such as deterrence theory (Straub, 1990), threat avoidance (Johnston and Warkentin, 2010; Liang and Xue, 2010), moral judgement (Myyry et al., 2009), and planned behaviour (Bulgurcu et al., 2010). The underlying assumption in existing work is that the higher the intention of an individual to perform a secure behaviour, the greater the likelihood that the behaviour is performed. However, does a high level of intention reliably lead to secure behaviours in natural situations? The answer seems to be rather unclear. Researchers from various fields discuss factors limiting explanatory power of intentions (Abraham et al., 1999; Cooke and Sheeran, 2004; Limayem et al., 2007; Sniehotta, 2009). In information systems (IS) security research, evidence exists that even if organizations prescribe security policies and employees have high motivations to process confidential information securely, their situation-specific assessment of the importance of different goals may steer them away from prescribed behaviours and self-reported policy-compliance intentions (Hedström et al., 2011).

While prior IS security research thoroughly examines the role of users’ security motivation, this work aims to advance our understanding towards actual behaviours. We study how IT artefacts influence the relationship between end users’ intentions and behaviours. Particularly, we focus on information protection intentions and corresponding behaviours performed by end users to prevent unauthorized persons from gaining access to confidential information. Drawing on discussions on the intention-behaviour gap and our conceptualization of ILP tools as policy-enforcing technological agents, our work addresses two research questions: How is the information protection intention-behaviour relationship influenced by the presence of a policy-enforcing technological agent? How does a policy-enforcing technological agent influence the cognitive process users undergo during a confidential information elimination task?
To address the research questions we conduct a field experiment with a simulated work environment. Employees from a manufacturing company participate in an e-mail forwarding task with presence or absence of a policy-enforcing technological agent. Our work contributes to theory by sensitizing security researchers’ awareness for factors that play a role in the post-motivational phase. This understanding is important to allow more accurate predictions of behaviours, which would be limited by only looking at the motivational phase (Schwarzer, 2008). In addition, our work provides insights on what constitutes the effectiveness of policy-enforcing agents in changing end users’ information protection behaviours. Based on an analysis of screen recordings captured during the simulation, we suggest a model describing the cognitive process users go through when eliminating confidential information. These insights can be leveraged to derive design theories for policy-enforcing technological agents and to better understand situations in which ILP tools are especially effective to prevent information leakage incidents.

2 Theoretical Foundations

Several researchers studied individual users’ security behaviour from a socio-technical perspective (for a review see Anderson and Agarwal 2010). Various theoretical lenses have been applied to examine the motivational phase of users to understand the antecedents of intentions. In the security context, intentions are conceptualized as policy-compliance intentions (Bulgurcu et al., 2010; Myyry et al., 2009), policy-violation intentions (Siponen and Vance, 2010), intentions to use protective information technologies (Dinev and Hu, 2007), and threat avoidance intentions (Liang and Xue, 2010). Drawing on the theory of planned behaviour (TPB), intentions are argued to be the main proximal determinant of behaviours mediating other motivational factors such as attitude, subjective norm, and perceived behavioural control. Intentions are indications of how hard people are willing to try to achieve their goal and how much effort they are planning to exert (Ajzen, 1991). While there is a lack of studies examining the intention-behaviour link in the IS security context, several studies in the health protection context have found intentions to be consistent predictors of a range of different health protection behaviours (Abraham et al., 1999). For example, Cooke and Sheeran's (2004) meta-analysis of psychosocial correlates of condom use reported average intention-behaviour correlations of .39 in several cross-sectional studies (N = 1,890). On average, intention accounts for approximately 20% of the variance in self-reported behaviours. Consistent with the TPB and prior empirical work we examine this relationship in the information protection context.

Hypothesis 1: An employee’s intention to protect confidential information positively affects the actual protection behaviour.

While we suspect employees’ motivations positively influence actual protection behaviour, we assume policy-enforcing technology agents have an effect on protection behaviour as well. To hypothesise this direct effect, we describe our conceptualization first. Policy-enforcing technological agents are designed to interpret observable events to decide if a user action is a violation of a policy or not. In case a violation is detected, an enforcement act is initiated by the technological agent. We define an enforcement act as an interaction between a policy-enforcing technological agent and a policy-violating human agent of evaluative nature (Al-Natour and Benbasat, 2009). Similar to social interactions, in which law enforcement officers interact with law violators, the sanction and the behavioural response of the law violator most likely depends on the course of this interaction. We consider evaluative interactions relate to outcome-based performance feedback, since in a similar way decision-makers confronted with an enforcement act, can infer information about their policy compliance performance. Hogarth and Gibbs (1991) thoroughly examined the effect of feedback on decision-making performance and found that in repetitive decision-making tasks outcome feedback increases performance. Thus, we hypothesise that protection behaviour is higher under presence of a policy-enforcing agent:

Hypothesis 2a: Employees working under presence of a policy-enforcing technological agent show higher levels of actual information protection behaviour than employees working under absence.
While hypothesis 2a describes a direct effect, policy-enforcing technological agents might also affect the relationship between intention and behaviour. If such a moderation effect would exist, the strength of how information protection intention translates into actual behaviour would, for example, be influenced differently by technology-enforcement of security policies if one compares employees with low intentions to employees with high intentions. Moderation effects can add explained variance beyond TPB variables and were intensively examined in existing work (Sheeran, 2002). For example, in their meta-review of 44 studies on the intention-behaviour and cognition-behaviour relationship, Cooke and Sheeran (2004) found accessibility as one of the most effective moderator. Accessibility refers to the strength of the association in memory between cognition (e.g., formed intention) and the object of cognition (e.g., protection behaviour). People who possess high cognitive associations can remember their intentions more easily even in demanding situations and exhibit therefore stronger intention-behaviour relations than people who have lower accessibility. Differences in accessibility might be one reason for the intention-behaviour discrepancy. Sheeran et al. (2005) explain intention-behaviour discrepancies with problems of intention activation. “The activation level of an intention refers to the extent to which contextual demands alter the salience, direction or intensity of a focal intention relative to other intentions.” (Sheeran et al., 2005 p. 278). For any particular time and context people are likely to have multiple and often conflicting goals. They may be especially vulnerable to more pressing alternatives when particular goals involve short-term affective or cognitive costs. Thus, the relative activation level of any particular intention may be reduced by situation-specific activation of intentions related to alternative goals.

Our work draws on these ideas from cognition research, which explains reasons for intention activation problems. These ideas give rise to our particular interest in if and how IT artefacts function as situational cues that could help with intention activation. Assuming situational cues have an effect is valid as prior research found cues reduce intention activation level. A reduced intention activation level would prevent issues of prospective memory failure (e.g., forgetting to perform an intended behaviour) or goal reprioritization (e.g., intention fails to attract sufficient activation). In several experiments researchers found that participants who fail to enact their intention report they were ‘too busy’ or other goals were ‘more important’ at that time (Abraham et al., 1999; Orbell et al., 1997). Hence, whether a formed intention is remembered depends on “(a) its time dependent accessibility (i.e., ease of remembering), (b) the situation-dependent cognitive resources available, and (c) habits that foster remembering and [...] on situational cues.” (Tobias 2009, p. 410). Empirical findings support this view. In several studies accessibility was increased by situational cues and a more stable intention-behaviour relationship was found for participants that were reminded of their goals by situational cues (Cooke and Sheeran, 2004).

Policy-enforcing technological agents are event-triggered, hence, situational. They present cues, since their messages use keywords referring to the observed policy violation (e.g., confidential information). Thus, we hypothesise that policy-enforcing technological agents facilitate a more stable intention-behaviour relationship assuming they function as situational cues increasing accessibility:

*Hypothesis 2b: The presence of a policy-enforcing technological agent positively moderates the relationship between an employee’s information protection intention and behaviour.*

3 Methodology

3.1 Experimental Design and Participants

We tested the hypotheses in a single-factor field experiment. Technology-enforcement of security policies is designed as a between-subject factor, being absent in the control group and being present in the treatment group. The experiment was conducted in the European sales headquarter of a global manufacturing company. The experimental site was chosen for two reasons. First, it grants us access to employees in both business and IT departments. The diversity in participants’ IT expertise is important to increase generalizability of the experimental results, since prior studies indicate that IT expertise influences security behaviours (Kraemer and Carayon, 2007). Second, the chosen company setting is
representative in terms of technology-enforcement of security policies. A policy for classifying and protecting confidential information is defined, but not enforced through technology. In this respect the company is similar to many others. According to Ernst and Young (2011) 74% of 1,700 surveyed companies define policies for classifying information and 66% do not adopt technological leakage prevention.

Participants were recruited through intranet e-mails after management’s approval. In total, 82 participants voluntarily took part in the experiment, including 66 from business departments (43 female; 23 male; M\(_{age}\) = 35.03 years; department: invoice handling, pricing and sales) and 16 from IT departments (2 female; 14 male; M\(_{age}\) = 41.38 years; department: application support, development). The experiment was introduced as a simulation, the goal of which was to test how a newly designed future work environment would increase the effectiveness of communication with external business partners. Participants were randomly assigned to the control or the treatment group.

### 3.2 Experimental Task and Manipulation

The experimental task was a scenario-based e-mail forwarding task, analogizing to work practices of communicating with external partners. In daily operations, sales representatives are in contact with different external partners to communicate individual agreements, invoices, and login credentials to a collaboration website. In the task, employees had to prevent that internal confidential information leaks to a wrong external domain (e.g., eliminate confidential information prior forwarding to externals). Participants were textually instructed to play the role of a sales representative forwarding e-mails to one external partner. Microsoft Outlook with a hidden e-mail generation engine was used as the task environment. Outlook is the e-mail tool used by all employees in the company. Each participant needed to forward six e-mails sequentially, among which four contained confidential information and two did not. The sequence of e-mails was counterbalanced. The contents of the e-mails were designed based on semi-structured interviews with three sales managers, representing typical communication cases with external partners. E-mails vary in number of attachments and number of words. By mixing both confidential and non-confidential e-mails, the set-up resembled real work practices and reduced the risks of exposing the real purpose of the experiment to participants. Participants had to read the e-mails and had to decide which part to eliminate prior forwarding. E-mails used for the task are summarized in table 1.

<table>
<thead>
<tr>
<th>E</th>
<th>Relevant information</th>
<th>Confidential information</th>
<th>AT</th>
<th>WC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sales performance for this recipient</td>
<td>Sales performance of other business partners</td>
<td>1</td>
<td>187</td>
</tr>
<tr>
<td>2</td>
<td>External login credentials</td>
<td>Internal login credentials</td>
<td>-</td>
<td>85</td>
</tr>
<tr>
<td>3</td>
<td>New discount for this recipient</td>
<td>Sales performance of other business partners</td>
<td>-</td>
<td>282</td>
</tr>
<tr>
<td>4</td>
<td>New invoice for this recipient</td>
<td>Second invoice of another business partner</td>
<td>2</td>
<td>639</td>
</tr>
<tr>
<td>5</td>
<td>New discount for this recipient</td>
<td>None</td>
<td>-</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>New invoice for this recipient</td>
<td>None</td>
<td>1</td>
<td>239</td>
</tr>
</tbody>
</table>

*E = E-mail type; AT = Number of attachments; WC = word count (e-mail and attachments)*

**Table 1. Relevant and confidential content of the 6 e-mails used for the forwarding task**

An Outlook plug-in was developed to carry out enforcement acts. For the treatment group, technology-enforcement was present: participants received an on-screen notification when they attempted to send an e-mail containing confidential information to the external partner. The notification faded in at the lower right corner of the screen, presenting the company’s logo and a text message “Please remove confidential data from e-mail before forwarding.” After clicking “OK” on the notification, participants were able to modify the e-mail and continue. For the control group, technology-enforcement was absent and participants did not receive any notification in case of a policy violation.
3.3 Experimental Procedure and Measurements

The experiment was carried out in an ordinary office. Participants arrived in the experimental room individually. Upon arrival, every participant was welcomed by an experimenter, who is an internal employee. The participant was asked to read a paper-based introduction explaining the objective of the simulation and the voluntariness of participation. And then, the participant filled in an entry survey on a computer, asking about demographic information and intentions. Afterwards, the participant was guided by a task instruction and asked to imagine being a sales representative who is in contact with an external partner. The instruction stated to process e-mails as efficient as possible while making sure that confidential information such as internal password, prices, discounts, and invoices does not leak to external partners. Then, the participant started the experimental task in the e-mail environment, processing six e-mails consecutively. Participants’ behaviours were captured by a screen recorder during the experimental task. Finally, the participant was asked to fill in an exit survey, asking about perceived realism during the task. The experiment duration averaged to 33 minutes per participant. To reduce noise, the same closed meeting room and computer was used for all the participants and all experiments were guided by the same experimenter.

The measurement for intention to protect confidential information was adapted from Johnston and Warkentin (2010). Three items were measured (e.g., “I intend to protect my organization’s confidential information from unauthorized access (e.g., manually remove critical e-mail parts”). Perceived realism was measured in the exit survey with two items (adapted from Webster and Sundaram, 1998). For each participant, actual information protection behaviour was measured by calculating a protection-coefficient: number of sent and correctly protected e-mails divided by total number of e-mails with confidential information. Six e-mails had to be forwarded (six trials), but only four of the six e-mails contained confidential content. Hence, participants can score 1, .75, .50, .25, or 0 for behaviour. Further details on measurements can be found in the online appendix.

Collected data was analysed with two objectives. First, hypotheses were tested to understand the role of a policy-enforcing agent in the intention-behaviour relationship. Second, captured screen recordings of the simulated work environment were analysed with principles of inductive reasoning to derive a cognitive process model of confidential information elimination, which is important to understand how a user’s cognitive process is being influenced by a policy-enforcing technological agent.

4 Data Analysis and Results

4.1 Hypothesis Testing

The descriptive results for the control group and treatment group are presented in Table 2.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>IR</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>39</td>
<td>6.333</td>
<td>1.268</td>
<td>.852</td>
<td>.8397</td>
<td>.2779</td>
<td>6.076</td>
<td>1.035</td>
<td>.905</td>
</tr>
<tr>
<td>Treatment</td>
<td>43</td>
<td>6.364</td>
<td>1.133</td>
<td>.852</td>
<td>.8372</td>
<td>.2307</td>
<td>5.860</td>
<td>1.114</td>
<td>.905</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics for the control and the treatment group

Hypothesis were tested with a stepwise linear regression, which is commonly used to study the intention-behaviour relationship and to test for potential interaction effects (McClelland and Judd, 1993). Scores were centralized prior to computing the interaction term (Aiken and West, 1991).

Behaviour was regressed on protection intention in step 1, on protection intention and technology-enforcement at step 2, and the protection intention × technology-enforcement interaction term was added to the equation at step 3. Results are shown in table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2a</th>
<th>Model 2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protection Intention (PI)</td>
<td>.049</td>
<td>.049</td>
<td>.045</td>
</tr>
<tr>
<td></td>
<td>.015</td>
<td>.015</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>.345**</td>
<td>.345**</td>
<td>.318**</td>
</tr>
<tr>
<td>Technology-Enforcement (TE)</td>
<td>-.003</td>
<td>-.003</td>
<td>-.067</td>
</tr>
<tr>
<td></td>
<td>.036</td>
<td>.035</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>-.009</td>
<td>-.008</td>
<td>-.239*</td>
</tr>
<tr>
<td>PL × TE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.119</td>
<td>.119</td>
<td>.175</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>.119</td>
<td>.000</td>
<td>.056</td>
</tr>
<tr>
<td>$F$ for change in $R^2$</td>
<td>10.792**</td>
<td>.007</td>
<td>5.331*</td>
</tr>
</tbody>
</table>

Table 3. Stepwise linear regression with actual protection behaviour as dependant variable

As hypothesised, information protection intention significantly influences actual protection behaviour ($r = .345, p = .002$). Thus, we found support for hypothesis 1. Protection intention on its own explained 11.9% of the variance in behaviour. While technology-enforcement does not have a direct effect on information protection behaviour (cf. model 2a), model 2b shows that the interaction term is associated with a significant improvement in the fit of the model ($\Delta R^2 = .056, p = .024$) and that the interaction term and protection intention have significant regression coefficients. Thus, we did not find support for hypothesis 2a. For hypothesis 2b we found a significant moderating effect of technology-enforcement, but in the reversed direction as hypothesised. Thus, hypothesis 2b was not supported, but the significant negative moderation effect required a more detailed examination.

A post-hoc analysis revealed that the negative moderating effect is caused by the weakened intention-behaviour relationship among participants, who reported low intentions. This means when technology-enforcement is present, low-intenders confronted with enforcement acts were influenced in a way that they carried out protective behaviours against their originally self-reported intention. Thus, policy enforcement seems to have effects that relate to questions on behaviour-changing situational factors: “The lack of correspondence between behaviour patterns predicted by intentions and measures of actual behaviour is accounted for by two groups: [...] (a) Intenders who fail to act raise questions of empowerment. [...] (b) Nonintenders who act raise questions concerning situational factors that can overcome cognitive reluctance to the adoption of new behaviours” (Abraham et al., 1999 p. 253).

A cognitive process model for confidential information elimination can account for these findings. The following section describes the suggested cognitive process model for confidential information elimination.

### 4.2 Cognitive Process Model for Confidential Information Elimination

To investigate cognitive processes of humans in a certain task, Pellegrino and Glaser (1982) recommend decomposing tasks into cognitive components. The task in our study was to send relevant information to a recipient while making sure confidential information is eliminated prior to transmission. Screen recordings of experimental participants were analysed by counting frequencies of observed behaviours (e.g., whether e-mail attachments were opened or not). In addition, illustrative examples were identified, which can be seen as evidence for theoretical concepts extractable from observations (Mao and Benbasat, 2000). Two main principles of inductive reasoning were used: constant comparison and iterative conceptualization (Urquhart et al., 2010). Constant comparison is the process of constantly comparing observations labelled as a particular category with other instances of data. Iterative conceptualization focuses particularly on relationships between categories. These principles for theory-
development are originally employed in qualitative research, but similar principles have been employed to analyse observational data collected in controlled environments (Mao and Benbasat, 2000).

Derived concepts were integrated with existing theories close to the phenomenon of interest. Guthrie’s cognitive process model on information location in documents (Guthrie, 1988) and the theory on technology-supported decision-making (Todd and Benbasat, 1999) were used as theoretical scaffolds. With this approach we propose a cognitive process model for confidential information elimination composed of three phases: (1) goal formation, (2) information categorisation, and (3) information elimination. These phases emerged from the theoretical integration of the inductively analysed data.

1. **Goal formation:** Based on Guthrie’s cognitive process model on information location in documents we adopt the idea that users form a particular goal on how to process information before or during reading e-mails. In our analysis we compared information processing effort exerted per e-mail type before any enforcement act was triggered. Evidently, for e-mail type 4 significantly less information was processed by participants from different departments: 70% of participants from the invoice handling department processed all available information by opening both attached invoices. This compares to 100% of participants from the IT development department, who processed all available information. A t-test revealed a significant difference between these two groups, \( t(29) = 3.53, p = .001 \), two-tailed.

Thus, we suggest in line with decision-making theory that in information elimination tasks users choose an information processing strategy, which is determined by their prior knowledge. This prior knowledge is used for category selection and assessment of information processing effort. The category selection determines ‘what’ to search for and the perceived effort determines ‘how much’ information to process. Specifically users with high domain specific knowledge seem to be more susceptible to put less effort in processing all available information based on knowledge-based inferences (e.g., ‘I do not need to read attached invoices in detail, as invoices are non-confidential and relevant for customers’). Not processing all available information increased the risk of leaking confidential information and was thus wrong in our setting.

2. **Information categorisation:** Whether confidential information can be successfully identified does not only depend on a user’s appropriate information processing strategy, but also on the salience of the information category. Salience is the amount of attention a perceived object attracts (Tobias, 2009). We observed that relatively more participants were able to identify internal credentials (username and password) as confidential. For example, 89.0% of all participants were able to protect e-mail type 2 whereas only 65.9% were able to protect e-mail type 4, \( t(132) = 3.24, p = .002 \), two-tailed. We suggest this observation can be explained with search heuristics users can employ. For information with higher salience users can employ a search heuristics, which requires lower cognitive effort. Passwords are more salient as they match a specific pattern, are single words, and do not require interpretation of context.

3. **Information elimination:** Whether confidential information can be successfully eliminated does not only depend on a user’s category identification performance, but also on the entanglement of relevant and confidential information. We propose that confidential information, which can be visually and contextually separated from relevant information, can be eliminated with decision heuristics requiring lower cognitive effort. Hence, the category elimination performance is higher for such cases. No variance was introduced in our experimental task on category entanglement. However, we base this proposition on observed elimination strategies employed by participants. Several participants used a conservative approach in terms of not only eliminating confidential but also irrelevant information en bloc prior to forwarding. We speculate that in case multiple instances of relevant and confidential information are present in an e-mail with high alternation and with uncertain occurrence, category elimination performance can be different from category identification performance.

The cognitive process depicted in Figure 1 summarizes the conceptualization and describes how outcome-variables of each process phase are being formed and act as inputs to other phases. We propose goal formation, information categorisation, and elimination as distinct phases. This is based on the observation that even if information was processed completely, confidential information mingled into the context was not identified successfully in several cases and especially if a simple search heuristics was
not possible. For example, participants that successfully identified and eliminated passwords, were less successful in removing financial information from a table that displayed entangled sales information of the recipient and of competing partners.

Figure 1. Cognitive process model for confidential information elimination

Based on the suggested cognitive process model we analysed how participants behaviourally responded if an enforcement act was triggered.

4.3 Responses to Policy-Enforcing Technological Agents

In case a participant tried to send confidential information via e-mail, an enforcement act was triggered. The send action was prevented and the participant was notified. The participant had to decide if and how to change the course of action. If the participant repeated the action without removing the confidential information, the notification was shown a second time and the action was prevented again. In case of a second prevention for a specific e-mail, the notification provided an additional functionality to set an ignore checkbox. Through this ignore functionality it was possible to send the e-mail with confidential information. For the next e-mail, the ignore function reset to zero. Table 4 summarizes the observations on the final outcomes for each e-mail type that contained confidential information (cf. table 1).

<table>
<thead>
<tr>
<th>Group</th>
<th>Behaviour</th>
<th>EA</th>
<th>E-Mail 1</th>
<th>E-Mail 2</th>
<th>E-Mail 3</th>
<th>E-Mail 4</th>
<th>Overall</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Protection No</td>
<td>33</td>
<td>35</td>
<td>33</td>
<td>25</td>
<td>126</td>
<td>80.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Protection No</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>14</td>
<td>30</td>
<td>19.2%</td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>Protection No</td>
<td>37</td>
<td>38</td>
<td>40</td>
<td>29</td>
<td>144</td>
<td>93.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protection Yes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>16</td>
<td>7.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Protection Yes</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>12</td>
<td>7.0%</td>
<td></td>
</tr>
</tbody>
</table>

EA = describes if an enforcement act was triggered for this case (Yes/No); Percentage = divide overall behaviour outcome to respective experimental group (control or treatment group)

Table 4. Protection behaviour outcome for all 328 cases (4 confidential e-mails; 82 participants)

No protection behaviour was observed for 7.0% of e-mail forwarding cases in the treatment group. That means in the simulated environment 12 cases were observed, in which participants forwarded confidential information even if they were informed by a policy-enforcing technological agent. In doing so, they had to choose the ignore functionality beforehand. On the other hand, in 16 cases, technology-enforcement helped to prevent the simulated leakage of confidential information. Participants were able to successfully eliminate confidential information prior forwarding. We compared these 28 cases in detail. We found that in 13 of the observed 28 cases technology-enforcement successfully influenced the information processing strategy. In these cases, additional effort to process all available information was exerted as a response to the enforcement act. For two of these 13 cases more information was processed, but still the categorisation was not successful. For further cases, no systematic responses were observed (e.g., relevant information eliminated, e-mail subject modified). It seems that the design of the policy-enforcing technological agent had a low effectiveness in making participants infer the appropriate protection action. In some cases, participants were confronted with up to 6 notifications and still were not able to appropriately protect confidential information prior e-mail transmission. To explain what determines effectiveness of a policy-enforcing agent we relate these observations to the cognitive process
model and suggest these capabilities may increase a technological agent’s effectiveness: (1) enables users to perform a knowledge-independent category selection, (2) increases the salience of relevant, irrelevant, and confidential information, and (3) resolves information category entanglement.

5 Summary

While prior socio-technical IS security research thoroughly examines the role of employees’ security motivation, our work aims to advance our understanding towards explaining actual behaviours. Specifically, we investigate the relationship between end users’ intentions and behaviour in the context of information protection behaviour to address the question of ILP tool effectiveness.

In this study, we tested hypotheses with 82 participants in a field experiment. Results support the view that there is a positive relationship between information protection intention and behaviour (H1). No evidence was found for H2a, which posits that the presence of a policy-enforcing technological agent systematically increases information protection behaviour. In reverse to the hypothesised direction, we found a negative moderation effect of technology-enforcement influencing the intention-behaviour relationship (H2b). Finally, we relate observations on employees’ responses to enforcement acts to our suggested cognitive process model. With this approach we explain what capabilities of policy-enforcing technological agents might increase their behavioural change effectiveness.

5.1 Limitations

The following limitations need consideration when assessing our work’s findings and contributions. First, employees’ behaviours were observed in a scenario-based simulated environment. Scenarios describe hypothetical situations in which behaviours are based on interpretations which might be different from real circumstances. However, in balancing realism versus artificial, self-reported versus observed, such simulations allow to control realistic stimuli and have advantages over methods such as questionnaire surveys. In surveys the variance in respondents’ interpretations of items is more likely to be a source of error (Snow and Thomas, 1994). Second, despite our careful experimental design and three rounds of pre-tests the data shows that the overall information protection behaviour is quite high. Potential group differences may have been obscured by a ceiling effect, eventually caused by too simple e-mails used for the task. Third, technology-enforcement is an intervention, which most likely influences intention formation as well. We did not measure if intentions changed during the experimental task as their stability was not of interest to our study. This is comparable to cross-sectional survey methods.

5.2 Contributions and Future Research

Contributions of our work to research are threefold. First, our work addresses calls of researchers to study actual behaviours even if they “are difficult to study in the security context” (Anderson and Agarwal, 2010 p. 628). Existing research is limiting data collection and analysis in order to investigate motivational variables, mainly intentions. While it should not be questioned that intention exerts a direct effect on actual behaviour, circumstances might exist under which this effect is partly or even entirely influenced by other variables (Limayem et al., 2007). Second, our work extends existing work on usable security (Payne and Edwards, 2008), which mainly focuses on how to overcome users’ cognitive limitations in the context of password remembering or e-mail encryption. Prior research studied why users do not encrypt certain e-mails and found that misunderstanding or forgetting occurred most frequently (Whitten and Tygar, 1999). While existing research does not provide any model of the users’ thought process, our work provides such a model. Third, our findings provide evidence that even with situational support users require decisional guidance to successfully protect confidential information. The knowledge base on decisional guidance can be employed to understand if and how improvements in information protection behaviour materialise through corresponding IT artefacts design.

Insights into the user’s cognitive process and on how policy-enforcing IT artefacts influence a user’s decision can be used to understand the underlying human-artefact interaction and corresponding IT
artefact design. For example, the triggering of an enforcement act might imply an existence of diverging goals between the policy-enforcing technological agent and the policy-violating human. Such an interaction might be fundamentally different from interactions studied in the domain of decision support systems. In decision support systems research users use artefacts according to their perception of associated utilitarian benefits (Al-Natour and Benbasat, 2009). This might be different for policy-enforcing IT artefacts. Diverging goals between a user and an artefact raise questions of what role users’ associate with policy-enforcing artefacts and how perceived legitimacy influences this understanding of roles. Such aspects require exploration.

This work’s practical contributions are illustrated with two questions executives might ask: “Considering the high level of security awareness and motivation we measured among employees, is it a good or bad idea to implement ILP tools in our organization? How effective are leakage incidents being prevented?” Our simulated ILP tool was found not to directly increase overall protective behaviours. However, low-intenders were influenced in a way that they behaved more securely that they reported they would. Thus, ILP tools might particularly help low-intenders. If leakage incidents can be prevented effectively depends on explanations used by ILP tools. Event-triggered ILP tools without context-awareness, as they have been simulated in our study, seem to have limited effectiveness. Executives should consider these findings for ILP tool investment decisions.

References


