Pretty Good Persuasion: A First Step towards Effective Password Security in the Real World

Dirk Weirich
Department of Computer Science
University College London
Gower Street
UK - London WC1E 6BT
(+44) 207 679 3033
D.Weirich@cs.ucl.ac.uk

Martina Angela Sasse
Department of Computer Science
University College London
Gower Street
UK - London WC1E 6BT
(+44) 207 679 7212
A.Sasse@cs.ucl.ac.uk

ABSTRACT

In the past, research on password mechanisms has focussed almost entirely on technical issues. Only in recent years has the security research community acknowledged that user behavior plays a part in many security failures, and that policies alone may not be sufficient to ensure correct behavior. We argue that password mechanisms and their users form a socio-technical system, whose effectiveness relies strongly on users' willingness to make the extra effort that security-conscious behavior requires. In most organizations, users cannot be forced to comply; rather, they have to be persuaded to do so. Ultimately, the mechanisms themselves, policies, tutorials, training and the general discourse have to be designed with their persuasive power in mind. We present the results of a first study that can guide such persuasive efforts, and describe methods that can be used to persuade users to employ proper password practice.

KEYWORDS

security, passwords, user-centered design, mental models, cognitive task analysis, user training, motivation

1. INTRODUCTION

Password mechanisms are the first line of defense of most computer systems, and therefore affect almost every user on a daily basis. Research on security mechanisms in general has in the past focused almost exclusively on technical issues. Only in recent years has the security community recognized that user behavior is a part of many security failures, and started to consider the effect of human factors in security (see, for example, [12,4,10,3]). [1] and [9] have shown that current password mechanisms have largely failed to consider usability, and that – given the increasing number of systems and passwords - most users cannot cope with the demands imposed on them. In addition, many users are not sufficiently educated about security issues. Thus, many usersm construct their own, often wildly inaccurate models of security threats and the importance and effective deployment of security measures. All this has led to a situation where a large number of users consistently behave in a manner that undermines the security of the systems they are using: they choose cryptographically weak passwords, write them down, and readily disclose them to other people. It is exactly these kinds of behaviors that are exploited by hackers and industrial spies, many of whom use social engineering [11,6].

We believe that the usability of password mechanisms will be improved, and that users will become better educated. However, even once this is achieved, there is an additional issue that needs to be addressed: the willingness of users to behave in accordance with proper password practice. In most cases, authentication to a system is an *enabling task*, which means it creates an overhead for the user, who is using that system as a tool to achieve a *primary*, real-world *task*. It is predictable that most users will cut corners to reduce that extra load given a chance, unless they are *motivated* to make the effort to behave in a security-conscious fashion – an argument [10] have put forward for security mechanisms in general. Oversimplifying for the sake of argument, users of password mechanisms can be divided into two groups: those that face personal damage if they do not behave in a security-conscious fashion; and those that do not put themselves, but others, at risk by cutting corners. Self-employed and home users fall into the first category – users in this group can, if educated about the possible consequences of their behavior, make an informed choice about their behavior, based on an assessment of the risks and the effort required to reduce these risks. Users in an organizational context fall into the second category, and for them education will very often not be sufficient. [9] found that users who had access to systems essential to the operations of their company - which had experienced a number of break-ins - had an attitude set towards security that could at best be labeled as

NSPW'01, September 10-13th, 2002, Cloudcroft, New Mexico, USA.

'unconcerned'. Organizations put themselves at risk if they have employees with such attitude sets, which are likely to prevail even with further education.

Work on computer security has always been strongly influenced by its roots in the military environment, where users can be expected to comply with rules given to them. [1] demonstrated that this approach does not work in modern business organizations with skilled, empowered knowledge workers, who do not work under constant supervision and are supposed to use their own initiative. These users cannot be expected to comply with security practices that they perceive as obstacles on the path to getting their job done. They will be aware that it is impossible to monitor their password behavior constantly, and are therefore likely to ignore such rules. One of the studies in [9] found that the vast majority of users did not follow company rules for passwords. In many corporate environments, the highest-ranking executives are those least likely to comply with security rules because they *"don't have time"* to bother with procedures that *"get in the way of more important things"*. Monitoring staff closely to enforce compliance would be prohibitively expensive and unacceptable from a human resources point of view. Since employees cannot simply be forced to comply, they have to be *persuaded* to do so. We believe that in the long run, any persuasive effort will only be fully effective if the password mechanisms are usable, integrated with real world tasks, and designed from the very beginning with their persuasive power in mind. However, in this paper we will present a first demonstration of 'pretty good persuasion' without changing the mechanisms. Some of these methods rely on changes to the policies and the way they are enforced, and some rely purely on changing the discourse about passwords mechanisms, supported by a social marketing campaign. Ultimately, only a combination of all these methods will achieve maximum persuasive power.

Our research was originally motivated by a simple set of questions. In large organizations, many users have similar jobs to do, and access information with the same degree of confidentiality. How can it be that some of them are motivated to behave in a security-conscious fashion, and others are not? Is this due to general personality differences, or can it be traced back to their mental constructs, e.g. their knowledge, beliefs and attitudes? And if it can be traced back to their mental constructs, would it be possible to entice users who behave improperly to take on the constructs of users that behave well, thus changing their behavior? In an initial investigation of these questions, we carried out semi-structured in-depth interviews on password security with 17 participants. Ten of these worked for a technology company, 6 were doctoral candidates, and one was a systems administrator working in a bank. The interviews lasted 30-60 minutes and were subsequently transcribed for analysis. Interviews (rather than questionnaires) were chosen in order to allow exploratory questioning, and since it has been reported that a lot of people will answer questions on security in interviews that they will not answer in a questionnaire [2]. We kept the interviews as open as possible, but were broadly guided by concepts taken from Rogers' protection motivation theory [8]. The theory is concerned with the use of *fear appeals* to change the behavior of people. It states that fear appeals will be effective if they convince the recipient that

1. the problem is serious;
2. it may affect her/him;
3. it can be avoided by taking appropriate action; and
4. the recipient is capable of performing the necessary behavior required to avoid the problem.

We initially analyzed the interviews looking for beliefs, attitudes and knowledge items, but subsequently found the concept of *interpretative repertoires* (IR) extremely useful. Our use of this concept draws on Potter and Wetherell's formulation of discourse analysis [5], and its application in Human-Computer Interaction [7]. Discourse analysis argues that language *constructs* reality, rather than representing or reflecting it. There is always more than one way to describe things, and our choice of how to describe particular aspects of reality has an immense power to shape the way we experience the world and behave in it. Interpretative repertoires are the shared linguistic resources we draw on to construct aspects of reality.

In analyzing the interviews, we made a number of discoveries that we believe to be important for anybody wanting to persuade users in an organizational context to behave in a security-conscious fashion. Section two of the paper will describe these findings in detail, but the following is a high-level summary:

1. A large number of the participants in the interviews had mental constructs that make it almost impossible to use fear appeals effectively to change their behavior. The good news is that there were also a few participants with mental constructs that can assist us in creating powerful fear appeals.
2. We found that participants quite freely disclosed their passwords to other members of their organization. The interesting point is that there is a strong social element in sharing passwords – it is seen as a sign of *trust* among co-workers. In addition, the criteria for who to share with, and when, directly play into the hands of hackers, industrial spies and social engineers.
3. Another way of persuading users to behave properly would be an advertising approach of associating 'positive qualities' with the desired behaviors. We found that, currently, the exact opposite is the case. People who behave in a security-conscious fashion are often described as 'paranoid' – even by themselves.

The third section of this paper, will present initial ideas on solving these problems. We are currently applying some of these approaches to establish their effects, whilst the others are promising avenues for future research. In particular, we suggest a three-tiered approach to address the current state of affairs:

1. In the short run, users' willingness to comply with existing regulations can be improved by changing the discourse about password mechanisms, and by using techniques from social marketing.
2. Where possible, additional changes to policies and the way they are enforced will increase compliance.
3. Ultimately, only password mechanisms that have been designed with their 'persuasive power' in mind will achieve the maximum level of compliance, in conjunction with the previous two methods.

2. UNDERSTANDING USERS

2.1 Who's Afraid of the Big Bad Wolf? Why Conventional Fear Appeals Don't Work for Most Users

In the introduction, we stated the conditions that have to be met for a *fear appeal* to be successful. In this subsection, we will show that beliefs held by many participants, and the interpretative repertoires they draw on, mean they are effectively immune to conventional fear appeals because not all of the necessary conditions can be met.

2.1.1 Who tries to get into other people's accounts – and why?

The interpretative repertoires participants draw on to describe the people who try to break into other people's accounts, and their motivation to do so, bear a direct relation on whether they perceived themselves as likely targets for such attacks. The repertoires used by most participants lead them to believe that this likelihood was small, as will be shown in 2.2.2.

The most prominent repertoire was **kids**[1], with **vandals** and **criminals** a distant second. The other repertoires reported here were only mentioned by individual participants.

Kids were described as sad little kids (*spotty little s***s, basically, that have nothing better to do than to fascinate themselves by writing programs about how to get into things which they're not supposed to.*) or curious kids (*Curiosity, just saying "This is secret, can I break into it, that would be fun." Like, basically, kids playing around.*). Some technically-minded participants even expressed a certain amount of admiration for them (*Very technically literate, very capable technically, with a devious mind. [laughs]*.). Their motivation is to prove they can do it, to get a buzz, to get a sense of achievement, or to be better than someone else and impress their friends. They target security-conscious organizations, prestigious ones, or the rich and famous. Once they have broken into a system, they might deface a web page, or leave a message, but they don't do any serious harm.

Vandals are seen as abnormal (*I don't know how to describe them. They're obviously not normal people.*). They want to have a pop at the establishment or are just plain mad (*but all the destructive stuff is like a cat burglar that's just having an episode in a place, you know, they lose their rag, they go completely mad and start racking the place, that seems a bit unnecessary to me*). They have the same targets as **kids**, but unlike them, they do serious damage in the systems they break into.

Criminals were seen almost exclusively as trying to carry out activities related to online banking - which none of the participants had direct access to from their company account - with only one participant considering the threat of their account being used as a base to commit fraud undetected.

Vengeful people are vengeful against a specific individual, whereas **disgruntled employees** want to get revenge on an organization. The final repertoires that appeared where **industrial spies**, **terrorists**, and **jokers** (*And they might even know the person that they're targeting, where it's just a joke, where they then send of an email purporting to come from some individual, saying outrageous things.*).

[1] The special font denotes the interpretative repertoires.

2.1.2 Whom do they target?

The likely targets for attackers are a direct result of the repertoires introduced above. The important point to make here is that none of the participants perceived themselves as falling under one of the repertoires that make up potential targets, except for the **weakest link** repertoire. Mostly, the targets are **security-conscious organizations** or **high-profile organizations** (*They're high-profile. Some of those are supposed to be very secure, like the Pentagon is supposed to be unbeatable, so if you can get into that, it's like a big thing, a big macho-thing, but also if you tell your mates you hacked into some system that nobody has ever heard of, they won't be very impressed, however secure it was. They won't be very impressed. You hacked into X, who is X? No-one's heard of them, so it's not very impressive.*). In addition, **people with important information** are targeted, as are **people who have annoyed the attacker**. Only few participants, and only after further questioning, drew on the **weakest link** repertoire (*Ahh... yeah, probably not, I think it's unlikely that anyone from outside would choose, you know, that their aim would be to get into specifically my account, but I think they could end up targeting me, you know, like I was saying, searching for a weak link in a corporate organi-, yeah, they want to get into some part of {company name}'s network, and I'm one way in, so they might, I might get targeted in that sense.*).

2.1.3 There is no personal danger

As shown above, the likelihood of being targeted personally was seen as small by the participants. In addition, we could confirm the results of the previous studies [1] and [9]: the severity of the negative consequences of someone breaking into the account is regarded as small as well. Participants did on the whole not believe that the information in their account was of importance or use to anybody (*but, but I mean, the sort of information that you, that's passworded is not of any interest to anybody. The number of man-hours that have been working on my project, who cares? There are items there that are important to me, and that I would know how to exploit them, but if somebody had a look at them, I think they would have great difficulty, first of all, in understanding them, and secondly, finding a market for them.*).

In addition, a number of mechanisms that organizations employ in order to reduce the possible negative consequences of break-ins directly lead to participants regarding the danger as less strong (*Q: Would there be any potential harm to you personally? A: Only if they send emails on my behalf, I think, that's the only scenario I could think of. They could destroy my work, but I use the mainframe as a backup, so everything that's on there is stored elsewhere anyway.*). Participants in the commercial organization also showed a strong belief in the security of their Intranet (*Ah, perhaps not so important is, to me, is the passwords dealing with computing security in terms of files, file storing places because, mostly, because we're inside an intranet, it's mainly secure from outside.*).

2.1.4 Hackers can always find a way in

We have shown that participants did not consider themselves as under threat. We also wanted to know whether they believed they could prevent someone from breaking into their system, and came across a repertoire that clearly diminishes that belief – **hackers**

can always find a way in *(Ahm, I think if somebody is determined enough to break into a system, they will expand the effort, either guessing the passwords or rampaging through bins to find those torn-up envelopes or, or whatever. I think if somebody is determined enough, they'll break in. [passwords] add another level to make it more difficult for people who aren't particularly experienced to access your account.).*

2.1.5 Personal accountability

We asked participants what they would do if their superior confronted them with the fact that someone had accessed other parts of the system from their account, causing considerable harm to the organization. Most participants did not regard this as a problem, since they could always rely on the **trust in me** (*people that know me personally would know that I wouldn't do things like that*) and the fact that **passwords are not infallible** (*Ahm, I mean, most, we work on, ah, in a company like {company name}, in such a big company, ahm, that sort of stuff may happen, and people are aware that passwords are not infallible and therefore there is kind of a, a trust among people, and if I said 'I didn't do it', then I would expect people to trust me, because, ahm, basically, the, the, it should be clear that systems are not totally infallible and some systems can be compromised.*).

2.2 Is There Hope? Why Fear Appeals Could Work Better for Some Users

None of the participants followed regulations on password security to the letter. However, some of them made more of an effort to behave in a manner that *they* regarded as security-conscious, or at least were aware of the dangers to them of not complying with regulations. The important point about the repertoires we present in this section is that some of them are direct 'antidotes' to the ones presented in previous sections. We present all of the repertoires we have found, even though some of them are likely to be more useful in persuading users to participate in the required manner than others:

1. **Allegiance**: *Basically, the way I see it, obviously, my main allegiance is to the department at the moment, rather than the College, because that's where I've been for god knows how long, so from my point of view protecting that account and the email that comes to and from that account is more important than the college facilities that I use.*
2. **Previous break-ins**: *I maintain the highest level of personal security I can on that because that has been hacked before.*
3. **Following policy**: *But, ahm, and I think I have this sort of back in my head, I have this sort of feeling 'Oh. I'm in the office. It's office policy. You know, there is a, there is a culture here, a security culture, definitely.*
4. **Avoid personal embarrassment**: *or if it was the network password, it could certainly be embarrassing in terms of, first of all people getting the impression that {name of speaker} was doing illegal things within the network and maybe sending out bizarre emails or viruses to people which would then be, probably, succeeded by the realization that {name of speaker} actually wasn't a malicious individual himself, but he'd been stupid enough to let his account being hacked into by someone who has, I mean, there would be an amount of sympathy, but also, people would get a bit tired with hearing about it, and they'd probably assume that it was your lax security that somehow allowed people to do this. So, embarrassment, probably, it's unlikely that it would bankrupt me, or lose me my job, for example.*
5. **Respect for other people**: *if, if something, something dealing with people matters, say, an appraisal, a performance review, all those things have to be kept confidential.*
6. **Privacy**: *I don't really have anything on it that I would die if someone gained access to but it's just the thought of it, you know it's my, it's my desktop, it's my setup, it's my files, I don't want anybody reading them. But they're not of a highly confidential manner, no.*
7. **Passwords are actually secure**: *I think an eight-digit, sort of arbitrary password is, I think it's pretty secure, to be honest with you.*
8. **Confidence in data**: *At the moment, because I've got, I don't, I'm not working on anything that I would deem to be confidential, I wouldn't mind anybody actually looking in to what I'm doing. I think I'd be upset if anybody else actually went in and wrote over my work, which is probably why the password is actually a benefit. But apart from that, that is...*
9. **Financial matters**: *No, yeah, because it's personal, you know, it has my bank balance, my bank transactions, the money I earn.*
10. **It would be hard to defend myself**:
    - *Well, that would be really hard, because, ahm, I have got some of them written down, in a, in a sort of place that, if somebody really wanted to, they could find, so, I think it would be quite hard to defend myself. Q But which angle would you take? A Mmmm... I really don't, I, gosh, I don't know. I don't know how I could defend myself. The fact that I've written them down for anybody to find... I, I just don't know, don't know. I'd have to think quite hard about that one. Q Okay. I think what a lot of people just say is 'I would just say: okay, that's what everybody does.', so... A Oohh. Yeah, but it, just because everyone else does it...*
    - *First of all, if somebody hacked in through my account into somebody else's account, then my account name will appear on the hacking record or whatever, and therefore I will be blamed for it. I won't be held responsible if my system was too easy to get into or if I had a easy-to-guess password. I'm sure the regulations say my password should be changed frequently and should be hard to guess. And it isn't. And therefore if somebody had broken in through it, I could be held responsible, I guess.*

2.3 The Sharing Culture: Why Social Engineering Is a No-Brainer

In this section we will investigate the actual situations in which participants shared passwords, and show that there is a social component that currently makes it difficult for many people to refuse a request to disclose their password. In addition, there are common criteria that determine whether a request to disclose one's password is successful or not. The point we want to make here is that the reasons for sharing passwords, and the criteria

underlying the decision to disclose passwords, offer ideal entry points for hackers and industrial spies using social engineering techniques. Finally, we point out repertoires that increase resistance to disclosing one's password.

2.3.1 *In what situations do people disclose passwords?*

There are a number of situations that lead to password disclosure. Industrial spies can easily exploit some of these, and hackers can attempt to engineer situations that allow them to ask someone else to disclose their password:

1. Have somebody access your account: *I've also had to give my password to another colleague, because I had to go home and had some urgent email, but I couldn't, I don't have access to email at home, so I gave that person my password and they checked my email for me.*
2. Necessary for work: *There's been, when we do experiments, it's often to set, to set up the computer in order to do an experiment, we sometimes have to give each other passwords, I mean I've had another colleagues password as well. Over the last few days, in order to do an experiment because my home directory didn't have the, the correct paths in it. So she gave me hers.*
3. Following higher orders: *Right. Okay. Ahhmm, I'd do that if, if my group leader phoned me up and told me that...*
4. Informal support: *Ahm. Well, ahm, because I'm computer-illiterate, ahm, I have to have a trusted friend who can help me out, so, ahm, one of the young people in the team who is very, very good at jiggling around with PCs has sort of taken me under her wing. And I usually manage to find someone who does that for me, wherever I go, so, ahm, so of course I let her have my password, so she can get onto my PC and change things and do things.*
5. Organized sharing: *what we do is, we actually, ahm, inside our group, we write, write down passwords that are deemed to be important, and we put them in sealed envelopes, and they're in our head of group's filing cabinet locked away.*

2.3.2 *'Don't you trust me?' The social component*

One important finding is that password disclosure is seen as a sign of trust between colleagues – and the refusal to disclose as a sign of lack of trust (*I'm dodgy. Like I'm dodgy. Would they have had a good reason? I think intellectually I could understand why someone would want to not tell anyone their password but I think I'm trustworthy and I would take it as a personal insult if a situation arised... had arisen... I think I would... a situation arose where I would need to someone's machine to achieve something that was important and I couldn't do it because they refused to give me their password, I would consider that to be a little over-protective. And I think I'd feel a little bit insulted about their views about my ability to use that password sensibly. Ahm, probably because it comes down to 'Don't you trust me?' Since we work together, sort of, on a daily basis.).* Someone unwilling to participate in this social activity can easily be seen as hiding something *(Somebody that has something which he's not supposed to have, or just very secretive by nature without having any reason for it.).*

2.3.3 *Criteria underlying the decision to disclose one's password*

As we have seen, there are certain situations in which people disclose their passwords, and in which it might even be a disadvantage in the social context not to. In addition, there seems to be a common decision-making process that is based on all or some of the following criteria:

- Trust is the key criterion: *But usually it would just be one or two people that I trust, trust more or completely, I suppose. Ahm, that's not what I mean. With one or two people who, I suppose, yeah, I suppose trust is the right word. Ah, it would probably, yeah. Yeah, I think that's what I mean.*
- Trust is often related to proximity: *Aaah, trust, I suppose and proximity to my, I mean, I'd choose someone from the group, you know, the group that I work in.*
- The danger of sharing is considered: *Basically, I see it as a bit of an equation, really. Depending on the degree and severity of the information, depending on how serious the consequences of disclosure are and there's another variable, which is obviously the amount of trust. If you sort of put it all together, you know, that's my implicit sort of mechanism for disclosure.*
- The importance of sharing to the other person is a criterion – do they have a reason to ask for the password? *Even if I trust them, then the second question will apply, whether they are doing anything dodgy for a start. I don't want to break the law unwittingly. But you know, if it was a fairly reasonable request, i.e., I need to be able to print something out because my password, I've forgotten my password, [department's name] expects some written documentation beforehand and it's an emergency, and I've got this floppy disc and I just need to print it, you know that probably wouldn't bother me as much, if I knew and trusted them, that would be fine.*
- An additional criterion can be whether nobody else can help: *I probably would, if, if I couldn't be there to do it myself, on their behalf, or there's actually nobody else they could go to and it was a particularly important piece of information they needed to get at.*

2.3.4 *Repertoires that increase resistance to sharing*

There are a number of repertoires that reduced participants' willingness to disclose their passwords – or even completely obliterated it. Again, some of these are direct antidotes to repertoires encouraging password disclosure:

- **Can always find an alternative way**: *so, I wouldn't, if I needed really to read my email or something then I would find another way to do it, not by giving somebody my password to access the system.*
- ***I don't want to become a suspect*** : and if something happens to that other person's account, then you could be somebody who would become, would be a suspect in that situation, so I don't try to get information about other people's security information or password information other than...
- **I use this password for several systems**: *No, it's more a case of, I think, what it, I think the reason is because that*

password, I use, I will use the same password in different places. Effectively I am giving away, I'm giving somebody else the ability to do things as me. Ahhmm, by giving them my password. Now, if I used a different password for every system, that wouldn't necessarily be a problem, because I would know the limit of the damage, as it were.

2.4 *"I'm not a nerd"*: About Proper Behavior and Negative Self-Image

One of the most interesting issues we discovered in our study was the relationship between *self-image and security-conscious behavior*. People who care about security often carry a negative image because of this. This section summarizes repertoires that could be filed in the above sections, but we wanted to keep them together because they all carry a very personal tone:

- A common perception is that only technically-oriented people understand security issues and care about them. Obviously, this is a disincentive for people who don't want to be seen as **nerds**: *There's a lot of people who are technologists, and, and they tend to know about things like security issues, and they care about them.*
- People who are concerned about security are often regarded as **paranoid** – even by themselves: *I suppose general personality types. People who would want to be more secure. I don't know. That's really a question for psychologists. What sort of people keep their desks tidy. What sort of people comb their hair in the morning. Probably the same sort of people who would not give their passwords away. People who are very sort of... either people who are very paranoid about breaches of security*
- They might also be regarded as **anal** and pedantic: *Mmm... I'd just think they were very diligent in following the site's security policy. They're more worried about not to be seen to be breaching any security rules. I mean, some colleagues, even though you might work with them, might be particularly pedantic on that kind of thing, or...*
- People not disclosing their passwords can be seen as unsociable, or might even get the image of **not being team players**: *Completely closed and shuttered down and, not, don't want to give away, share, not, not team players, as they say. I would say they're those sort of people. But yes, but I think people who are like that as part of their nature, I think that's just how they are as people, and they're, they're just not team players at all, just very shuttered and closed, and I'd probably think they're a bit weird, to be honest.*
- People not following regulations can be seen as **pragmatic**: *I think it's, it's interesting, we're all given hold of these passwords, and we're not supposed to share them, but I think people are more pragmatic about things, so I wouldn't be surprised if it happened, so, ah.*

3 . APPLICATIONS AND FUTURE RESEARCH

In section two, we have presented a large number of interpretative repertoires that undermine security-consciousness, as well as some that increase it. The aim of any intervention must be to make users abstain from the former and employ the latter. The approaches to achieving this we present here are changes to the mechanisms itself, policies, tutorials, training and the general discourse about passwords. We believe that a combination of all of these will prove most effective, though 'pretty good persuasion' can be achieved without changing the mechanism itself.

3.1 Methods Not Requiring Changes to Policies or the Mechanism Itself

3.1.1 Changes to the discourse about password mechanisms

The interpretative repertoires we have presented co-exist as a complex, entangled web within individual users. Any discourse about password mechanisms, for example in tutorials and training, should obviously introduce and reinforce the desired repertoires. In addition, it should use those repertoires that act as antidotes to undesirable ones. An example would be to point out that any break-in into an employee's account might result in personal embarrassment (**avoid personal embarrassment**) in order to combat the general belief that no personal danger can be caused by such break-ins.

A further interesting area of future research would be the deployment of an adequate metaphor for the whole password mechanism that counteracts some of the repertoires that undermine proper security behaviors. One metaphor we are currently investigating in the context of private users is the 'burglar alarm'. As with password mechanisms, users of burglar alarms are aware of the fact that they can ultimately not keep out a highly determined intruder. Still, most house owners install burglar alarms in order to make it as difficult as possible for the intruder to get in. In the scenario we are currently investigating, we are pointing out that attackers of computer systems will ultimately go for the easiest target – which means that a person employing proper password practice does not fight the intruder, but competes with other users to be better-protected then them, so the intruder attacks them, not her/him. This idea is equivalent to the situation with burglar alarms, and might be conveyed easily by using this metaphor.

3.1.2 Social marketing for social people

An important result of our study are the social and self-image issues we have discovered. An interesting and promising area of future research is the possible use of concepts and methods from *social marketing* in order to associate positive qualities with proper password practice, and negative ones with bad password practice. One example would be an advertising campaign depicting people behaving properly as professional and caring about their organization, and those behaving improperly as highly unprofessional and anti-social in that they put their colleagues at risk.

3.2 *If There Is no Reason to Be Security-Conscious, Create One:* A Different Way of Using Fear Appeals

The findings in section 2.1. show that many users do not expect to suffer personal consequences from improper password behavior. Current security policies tend to threaten punishment for improper password practice, but these are hardly ever enforced. It is likely that the actual enforcement of these policies would meet with resistance among users, considering that most of them do not

believe there to be any *reason* to be security-conscious in the first place. The challenge then is to find a way of creating such a reason in a way that meets their acceptance. One such way, which we are currently investigating, is based on a change of policies and the way they are enforced, intertwined with a justification for this change that stresses the danger to the organization rather than the individual. The change we are investigating is based on the following ideas:

1. Present the danger as one of the organization's reputation being tarnished if it were to be known to the outside world that its employees did not behave in a security-conscious fashion. Depending on the type of the organization, this might focus on issues such as ensuring that customers' data is kept secure. This gives the fear appeal (and its associated punishment) a rational motivation that will raise users' acceptance of it.
2. Punish non-compliant behavior if it is careless, rather than due to a lack of knowledge and support.
3. Be seen to punish such behavior.

3.3 Changes to the Password Mechanism Itself

The following is a radical scenario that we are currently investigating in focus groups in order to determine the effectiveness of its individual elements:

1. The system hands out to each user a unique password that can not be changed.
2. In addition, the user is given instructions at the time of receiving the password on how to memorize it.
3. The user can log into his system using the password alone – no user_id is needed.
4. In case the user forgets his/her password, it takes 24 hours to be allocated a new one.
5. The password is changed only at long intervals, e.g. every six months or more.

The aim of these changes is to associate the password closer with its user – since s/he can log in with the password alone, anyone finding a written copy of it can abuse it. Since it is changed only at long intervals, anyone this password is disclosed to has access to the system for a long time. In addition, it is made inconvenient to get a new password, thus increasing the importance of the password, putting it on par with a key that is not replaced instantly either.

4. CONCLUSION

We have put forward an argument that can be summarized as follows:

1. Password mechanisms and their users form a socio-technical system whose aim it is to achieve security.
2. Users' willingness to make the extra effort that security-conscious behavior requires is a vital variable influencing the effectiveness of this system.
3. Users cannot be forced to behave in a proper fashion, but an effort to *persuade* them to do so has to be made.
4. Systems, policies, tutorials, trainings and the general discourse about password mechanisms have to be designed with their persuasive power in mind.
5. *Pretty Good Privacy* can be achieved without changing the mechanisms themselves, though optimal results will only be obtainable by complete redesign.

We have given the results of a first study that can be used to guide the development of persuasive methods. In addition, we have given first ideas on which methods might deserve specific research attention in the future. Finally, we would like to stress that the applicability of 'pretty good persuasion' is not restricted to password mechanisms, but is likely to increase the effectiveness of other security mechanisms as well.

5. REFERENCES

[1] Adams, A. and Sasse, M.A. (1999), "Users are not the enemy", Communications of the ACM, Vol. 42, No. 12.

[2] Adams, A., Sasse, M.A. and Lunt, P. (1997), "Making passwords secure and usable", In H. Thimbleby, B. O'Conaill & P. Thomas (Eds.), "People & Computers XII (Proceedings of HCI'97)", Springer, pp. 1-19.

[3] Dufft, Cornelius C., Espey, Juergen, Neuf, Hartmut, Rudinger, Georg and Stapf, Kurt (1999), "Usability and Security", In Guenter Mueller & Kai Rannenberg (Eds.), "Multilateral Security in Communications, Volume 3 - Technology, Infrastructure, Economy", Addison Wesley.

[4] Holmström, U. (1999), "User-centred design of secure software", Proceedings of Human Factors in Telecommunications. Copenhagen, Denmark.

[5] Potter, J. and Wetherell, M. (1987), "Discourse and social psychology. Beyond attitudes and behaviour", Sage Publications Ltd., London.

[6] Poulsen, K. (2000), "Mitnick to lawmakers: People, phones and weakest links", http://www.politechbot.com/p-00969.html.

[7] Rimmer, J., Wakeman, I., Sheeran, L. and Sasse, M.A. (1999), "Examining users' repertoire of Internet applications", In M.A.Sasse & C.Johnson [Eds.], "Human-Computer Interaction – Proceedings of INTERACT '99".

[8] Rogers, R.W. (1983), "Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation", In J. Cacioppo & R. Petty (Eds.), Social Psychophysiology. New York: Guilford Press.

[9] Sasse, M.A., Brostoff, S. and Weirich, D. (2001): Transforming the "weakest link": a human-computer interaction approach to usable and effective security. BT Technical Journal, Vol 19 (3) July 2001, pp. 122-131.

[10] Whitten, A. and Tygar, J.D. (1999), "Why Johnny can't encrypt: A usability evaluation of PGP 5.0", 8th USENIX Security Composium. Washington.

[11] Winkler, I. (1997), "Corporate Espionage: what it is, why it is happening your company, what you must do about it", Prima Publishing, CA.

[12] Zurko, M.E. and Simon, R.T. (1996), "User-centered security", New Security Paradigms Workshop, CA