Human-centred security starts with understanding humans and their interaction with technologies, controls and data. By discovering how and when humans ‘touch’ data throughout the working day, organisations can uncover the circumstances where psychological-related errors may lead to security incidents.

For millennia, attackers have been using methods of psychological manipulation to coerce humans into making errors. Attack techniques have evolved in the digital age, increasing in sophistication, speed and scale. Understanding what triggers human error will help organisations make a step change in their approach to information security.

This briefing paper helps security professionals to:

- understand how psychological vulnerabilities in humans can lead to errors in decision making
- identify methods and techniques used by attackers to exploit psychological vulnerabilities
- manage psychological vulnerabilities to improve information security.

“Information security should put humans at the centre. If they truly are security’s ‘weak point’ then start by understanding their vulnerabilities and managing them.” – ISF Member

Humans have frequently been regarded as the ‘weakest link’ in information security. In 2018, 88% of data breaches reported to the UK Information Commissioner’s Office were attributed to human error, rather than vulnerabilities in the underlying technologies. However, organisations have historically relied on the effectiveness of technical security controls, instead of trying to understand why people are susceptible to mistakes and manipulation. A new approach is clearly required: one that helps organisations to understand and manage psychological vulnerabilities, and adopts technology and controls that are designed with human behaviour in mind. That new approach is human-centred security.

ABOUT THIS BRIEFING PAPER
Human-centred security starts with understanding humans and their interaction with technologies, controls and data. By discovering how and when humans ‘touch’ data throughout the working day, organisations can uncover the circumstances where psychological-related errors may lead to security incidents.

For millennia, attackers have been using methods of psychological manipulation to coerce humans into making errors. Attack techniques have evolved in the digital age, increasing in sophistication, speed and scale. Understanding what triggers human error will help organisations make a step change in their approach to information security.

This briefing paper helps security professionals to:

- understand how psychological vulnerabilities in humans can lead to errors in decision making
- identify methods and techniques used by attackers to exploit psychological vulnerabilities
- manage psychological vulnerabilities to improve information security.

PUTTING HUMANS AT THE CENTRE
Many different terms are used to describe human-centred security, including human-centric security, people-centric security or people-focused security. They all relate to the aim of mitigating or reducing the risk of human error.

ISF research identified that organisations are struggling to manage the risk of what is called “the accidental insider” – the authorised member of staff making accidental errors. Equally, traditional security controls are proving to be less effective at preventing external malicious attacks. Attackers are transitioning away from malware-based attacks to more targeted, social engineering-based attacks designed to coerce or influence the accidental insider into making exploitable errors, as illustrated in Figure 1. Some ISF Members have struggled to attract business buy-in and investment for the type of security awareness campaigns and training that can address these risks and threats.

Such failures in security controls, combined with under-investment in awareness, highlight a growing need to realign existing approaches to information security – and to put humans at the centre of future plans.

Figure 1: Email-based attacks blocked in 2018

<table>
<thead>
<tr>
<th>10% of attacks contained malware:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Viruses</td>
</tr>
<tr>
<td>- Ransomware</td>
</tr>
<tr>
<td>- Spyware</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>90% of attacks were more targeted:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Impersonation</td>
</tr>
<tr>
<td>- CEO fraud</td>
</tr>
<tr>
<td>- Spear phishing</td>
</tr>
</tbody>
</table>

1 Identifying human vulnerabilities

Human-centred security acknowledges that employees interact with technology, controls and data across a series of touchpoints throughout any given day. These touchpoints can be digital, physical or verbal. During such interactions, humans will need to make decisions. Humans, however, have a range of vulnerabilities that can lead to errors in decision making, resulting in negative impacts on the organisation, such as sending an email containing sensitive data externally, letting a tailgater into a building or discussing a company acquisition on a train. These errors can also be exploited by opportunistic attackers for malicious purposes.

### Errors

There is considerable debate regarding what constitutes a true error in decision making and what constitutes negligent behaviour. This paper focuses on errors as the result of decision making, rather than wilful, malicious acts. The following definition has been applied:

Errors are genuine accidents or mistakes made without malicious intent.

In some cases, organisations can put preventative controls in place to mitigate errors being made, e.g. preventing employees from sending emails externally, strong encryption of laptops or physical barriers. However, errors can still get through, particularly if individuals decide to subvert or ignore these types of controls to complete work tasks more efficiently or when time is constrained. Errors may also manifest during times of heightened pressure or stress.

By identifying the fundamental vulnerabilities in humans, understanding how psychology works and what triggers risky behaviour, organisations can begin to understand why their employees might make errors, and begin managing that risk more effectively.

### HEURISTICS AND COGNITIVE BIASES IN DECISION MAKING

When making decisions the human mind has to process a tremendous amount of information, but humans are limited by the amount of time they have to make decisions, as well as the amount of information at their disposal. Some information is consciously and rationally digested, whereas other information is subconsciously processed.

“People are often unaware of their own unawareness.”

— Thomas Gilovich, Heuristics and Biases: The Psychology of Intuitive Judgment

Weaknesses in the capacity of the subconscious human mind mean that humans look for cognitive shortcuts to reduce the effort the human brain expends during decision making, especially during times of heightened pressure. These shortcuts are known as heuristics. Many individuals do not know or acknowledge that their decisions are influenced by heuristics but they typically shape human interactions.

**Heuristics**

Also called “mental shortcuts”, heuristics are efficient mental processes that help humans solve problems and learn new concepts. These processes make problems less complex by ignoring some of the information that is coming into the brain. Heuristics are an essential part of quick decision making.

Heuristics can, however, result in cognitive biases, which lead to poor judgements or errors in decision making. There are several cognitive biases that impact everyday decisions. In information security a cognitive bias that leads to an error in decision making could cause significant damage to an organisation.

**Cognitive biases**

Cognitive biases are systematic errors in reasoning that lead to failures in producing correct decisions and judgements.

---

Figure 2 highlights how pressure can result in humans using heuristics in the decision-making process, and demonstrates the relationship between heuristics, cognitive biases and errors.

**Figure 2: Relationship between heuristics, cognitive biases and errors**

![Diagram showing the relationship between pressure, heuristics, cognitive biases, and errors.]

“There’s no way we can be robotic about human perception [...] as security specialists we need to mitigate these risks [cognitive biases] by understanding them better.”

— Dr Margaret Cunningham, Forcepoint

For many cognitive biases the information security implications are far reaching. If cognitive biases are not understood and management of them isn’t integrated into formal training and awareness campaigns, human error may continue to pose a significant information security risk to organisations. Figure 3 illustrates a range of cognitive biases that are most relevant to information security. These are outlined in more detail on pages 4 and 5 with examples of where and how they might manifest.

**Figure 3: Cognitive biases relevant to information security**

[Diagram of cognitive biases: Affect heuristic, Anchoring, Availability heuristic, Bounded rationality, Choice overload, Decision fatigue, Ego depletion, Herd behaviour, Licensing effect, Optimism bias, Over-justification effect, Polarisation.]

---

### Cognitive Biases in Information Security

<table>
<thead>
<tr>
<th>Cognitive Bias</th>
<th>Definition</th>
<th>Real-World Example</th>
<th>Information Security Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect heuristic</td>
<td>Humans make decisions by quickly relying on an emotional response.</td>
<td>An individual may buy a specific mobile phone because they liked the salesperson’s pitch and attitude. The individual may attribute these positive feelings with the mobile phone, resulting in the purchase.</td>
<td>An assistant receives an email from a hacker masquerading as their managing director, demanding that they send money to a fake supplier immediately. An emotional response drives the assistant into sending the money without proper consideration of the threat.</td>
</tr>
<tr>
<td>Anchoring</td>
<td>Humans tend to rely heavily on ‘anchor’ on the first or most profound piece of information they are given.</td>
<td>The initial price offered during the negotiation of a second-hand vehicle tends to set the tone for the rest of the deal. Subsequent counter-offers will typically be impacted negatively by the first offer.</td>
<td>During the NotPetya attack, some organisations mistakenly classified the attack because earlier media reports incorrectly referenced a different strain of malware.</td>
</tr>
<tr>
<td>Availability heuristic</td>
<td>Humans make judgements about the likelihood of an event based on how easily an example, instance or case comes to mind.</td>
<td>Investors may judge the quality of an investment based on recent information in the news, potentially ignoring other relevant factors.</td>
<td>Individuals hold open doors to buildings for delivery people typically because they have never experienced a tailgating incident before and therefore do not perceive them to be a threat.</td>
</tr>
<tr>
<td>Bounded rationality</td>
<td>Humans make ‘good enough’ decisions based on the time they have to make the decision.</td>
<td>A parent searching for a product with a crying baby will rush to find a product that will suffice but might not be the most suitable.</td>
<td>During a cyber attack, tensions run high, and a security analyst may make ‘good enough’ decisions based on the information they have and the tools at their disposal.</td>
</tr>
<tr>
<td>Choice overload</td>
<td>Humans struggle to make decisions when faced with too many options. This can cause significant decision delay, or complete avoidance.</td>
<td>A study found that consumers were 10 times more likely to purchase jam on a display when the options were reduced from 24 to 6.</td>
<td>Security analysts in a Security Operations Centre (SOC) may be overwhelmed by a large volume of alerts, impeding their ability to quickly and accurately detect a potential security incident.</td>
</tr>
<tr>
<td>Decision fatigue</td>
<td>Repetitive decision-making tasks drain mental resources. Therefore, if humans become fatigued, they tend to make easier decisions, but not necessarily the best decisions.</td>
<td>Academics studied factors affecting the probability that prisoners who were put before a judge for a parole hearing would be set free. They found that prisoners seen in the morning and immediately after lunch were more likely to be set free, and those seen later in the afternoon were not.</td>
<td>Towards the end of the working day employees may be mentally drained and therefore more likely to disregard good security practices or policies.</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>COGNITIVE BIAS</th>
<th>DEFINITION</th>
<th>REAL-WORLD EXAMPLE</th>
<th>INFORMATION SECURITY EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ego depletion</td>
<td>Humans have a limited supply of willpower and it decreases with use.</td>
<td>Individuals that are dieting throughout the day and diligently avoid snacks in the office may fall victim to ego depletion in the evening once they expend all mental energy. This could result in ordering a takeaway or snacking on a favourite chocolate bar.13</td>
<td>Employees may follow information security policy after receiving awareness training, but at the end of the month their willpower may have diminished and their security behaviour lapses.</td>
</tr>
<tr>
<td>Herd behaviour</td>
<td>The tendency for humans to mimic the actions (rational or irrational) of a larger group. Herd behaviour can manifest itself across many group sizes.</td>
<td>If an individual enters a lift and the rest of the group are facing one direction, they are likely to follow suit.</td>
<td>Poor password hygiene can manifest if one member of a team writes passwords down on post-it notes and the behaviour spreads throughout the entire team.</td>
</tr>
<tr>
<td>Licensing effect</td>
<td>Humans allow themselves to indulge after doing something positive first.</td>
<td>Studies showed that those who took multivitamin tablets were more likely to engage in unhealthy activities, such as drinking alcohol and smoking.14</td>
<td>Employees that shred sensitive materials may perceive that they have ‘done a good deed’ for the day and therefore be more complacent and click on a phishing email.</td>
</tr>
<tr>
<td>Optimism bias</td>
<td>Humans believe that they are at lesser risk of experiencing a negative event compared to others.</td>
<td>Smokers tend to feel they are less likely than non-smokers to get lung cancer.15</td>
<td>Many company boards believe they will not be hit by a cyber attack, despite statistical evidence to the contrary.16</td>
</tr>
<tr>
<td>Over-justification effect</td>
<td>The loss of motivation and interest as a result of receiving excessive external rewards.</td>
<td>Children that are told they will receive ice cream if they “eat their vegetables” are unlikely to be motivated to eat them in the future without the reward of ice cream.</td>
<td>Some organisations reward their employees for ‘good security behaviours’, such as monetary rewards or extra annual leave. However, this may actually reduce the desire to perform good security behaviours without the rewards.</td>
</tr>
<tr>
<td>Polarisation</td>
<td>Polarised thinking places humans or situations in ‘either/or’ categories, with no shades of grey. An individual will see things only in extremes,17 which can result in catastrophising or minimising.</td>
<td>Opinion on climate change tends to be polarised with both sides ignoring evidence provided by the other and very few adopting a middle position.</td>
<td>Employees who do not work in security or IT may perceive information security as something that does not concern them. This can manifest in these employees not paying attention to policies, training or phishing campaigns.</td>
</tr>
</tbody>
</table>

Once organisations accept that decision making can be significantly disrupted by cognitive biases they can begin to recognise what triggers them and thus take steps to manage human vulnerabilities and the related risk to the business. From an information security perspective, the risk to information is most likely to occur at touchpoints where humans interact with technology, controls and data.

---

HOW COGNITIVE BIAS IMPACT THE WORKING DAY

An individual’s working day involves many touchpoints with technology, controls and data, creating a number of potential risks to the organisation. Figure 4 illustrates a working day where cognitive biases may have impacted decision making, leading to errors that attackers could take advantage of, with adverse consequences to the organisation.

Figure 4: Cognitive biases and their influence on decisions in the working day

Employee X has been a senior accounting executive at a multinational public organisation for over 10 years. Employee X has two children and is unfortunately going through a difficult divorce. The organisation is preparing for a high-profile acquisition, with employee X playing a key role.

1. Employee X wakes up late, checks emails and hastily clicks on an email containing a ‘confirm parcel delivery’ button. Employee X clicks on the button and rushes out of the house to catch the train.

2. Whilst on the train, employee X logs onto the company system, checking finances for the acquisition meeting later. There is no privacy screen but the employee comes from an affluent area and is not concerned about shoulder surfers.

3. The train is delayed. Employee X has to run to the office. Upon arrival, a delivery person is waiting at the door with a large box of goods, fumbling for a key. Employee X enters, holding the door open.

4. Employee X enters the acquisition meeting. All attendees take notes on paper and at the end of the meeting a few attendees take photos of the plans on the board with their mobiles. Employee X copies their behaviour.

5. All attendees go to the bar after the meeting. Employee X drinks far too much and stumbles out of the bar at the end of the night, leaving the bag with the notes and phone on the table.

- Anchoring: Trusted the ‘confirm parcel delivery’ button in the email was legitimate.
- Decision fatigue: Stressed and lack of sleep from the divorce.
- The ‘confirm parcel delivery’ button contained malicious content, infecting the employee’s device.
- Availability heuristic: Never experienced a shoulder surfer before.
- Licensing effect: Followed password policy procedures but did not use a privacy screen.
- Shoulder surfers saw account details on employee X’s screen and noted them down.
- Bounded rationality: Made a judgement call that the delivery person couldn’t possibly be an attacker.
- An attacker was masquerading as a delivery person and gained access to the building, stealing data.
- Affect heuristic: Only thought about being late for work, not the consequences of letting someone in without a pass.
- Herd behaviour: One attendee took photos, so the ‘herd’ followed suit.
- Optimism bias: Never considered that taking pictures of the acquisition would be a problem.
- Acquisition plans were stored inappropriately on personal devices, disregarding policy.
- Ego depletion: The supply of willpower had diminished as it was the end of the day. Concern for security was at an all-time low.
- Polarisation: Minimal concern for security was considered as drinks were flowing and all attendees were having a good time.
- Acquisition plans are stolen, and attackers begin manipulating the share price of the company.

Cognitive biases can occur under normal working conditions with no undue external influence, but psychological pressure can be used deliberately to trigger and exploit cognitive biases further. For centuries attackers have used attack techniques to trick, manipulate and coerce people’s underlying cognitive biases in decision making for their own gain.
2 Exploiting human vulnerabilities

Psychological vulnerabilities present attackers with opportunities to influence and exploit humans for their own advantage. The methods of psychological manipulation used by attackers have not changed since humans entered the digital era but attack techniques are more sophisticated, cost-effective and expansive, allowing attackers to effectively target individuals or to attack on considerable scale.

Attackers use the ever-increasing volume of freely available information from online and social media sources to establish believable personas and backstories in order to build trust and rapport with their targets. This information is carefully used to heighten pressure on the target, which then triggers a heuristic decision-making response. Attack techniques are used to force the target to use a particular cognitive bias, resulting in predictable errors. These errors can then be exploited by attackers, as illustrated in Figure 5.

“The whole idea is why invest hundreds of thousands of dollars to build your own malware when you can just convince someone to do something stupid?”  
– Adam Meyers, Crowdstrike

There are several psychological methods that can be used to manipulate human behaviour; one such method that attackers can use to influence cognitive biases is social power.19

---

TRIGGERING HEURISTICS AND COGNITIVE BIASES WITH SOCIAL POWER

There are six unique types of social power – degrees of influence that an individual or organisation has over others. Attackers can use social power to trigger heuristic responses and cognitive biases, thereby manipulating targets into making errors that can be exploited for malicious purposes.

Below is a list of the six types of social power, with examples of how attackers might use them effectively during an attack.

**Reward power**

Reward power incentivises targets to complete a task by promising that there will be a ‘reward’ if the task is complete, e.g. ‘419 scams’ typically incentivises targets by promising financial rewards.

**Coercive power**

Coercive power is the opposite of reward power and leverages fear and punishment to manipulate human behaviour, e.g. during CEO scams attackers may induce stress, increase the sense of urgency or heighten fear by using threatening language to force victims into performing tasks.

**Referent power**

Referent power typically uses the ‘cult of personality’ to manipulate individuals into following their idols blindly. Celebrities, sports stars and other high-profile personalities are held in high regard by their followers and attackers masquerading as them can manipulate their targets easily, e.g. some individuals have been messaged by celebrity impersonators on social media as part of a scam.

**Informational power**

Informational power uses privileged information to convince targets that the attacker is credible because they know things that only a credible source would know, e.g. during a fraud campaign, an attacker may target an employee using specific company information to convince their target to share privileged company financial details.

**Legitimate power**

Legitimate power influences individuals to carry out tasks if they are ordered to do so by those in positions of power, e.g. masquerading as a manager, law enforcement or delivery person.

**Expert power**

Expert power is used by attackers to impersonate an individual who is seen to possess in-depth information, knowledge or expertise, e.g. using technical jargon to give the impression that the attacker is an expert and should be trusted.

Attackers may use any or multiple types of social power as part of an attack technique to coerce their targets and induce the use of cognitive bias when making decisions.

ATTACK TECHNIQUES USED TO EXPLOIT HUMAN VULNERABILITIES

There are many attack techniques that use the method of social power to exploit human vulnerabilities. Attack techniques can be highly targeted or conducted on scale but they typically contain triggers which are designed to evoke a specific cognitive bias, resulting in a predictable error. Whilst untargeted ‘spray and pray’ attacks rely on a small percentage of the recipients clicking on malicious links, more sophisticated social engineering attacks are becoming prevalent and successful. Attackers have realised that it is far easier targeting humans than trying to attack technical infrastructure.

---

According to FireEye’s 2018 Email Threat Report, 91% of all cyber attacks still enter via email, but 90% of emails that were blocked and investigated in their sample set did not contain malware – they were more targeted in nature, using attack techniques that leverage social power to trigger cognitive biases such as impersonating the CEO. This highlights the growing need to understand attack methods and techniques that exploit human vulnerabilities.

Below is a list of some of the most common attack techniques and how they leverage social power to trigger cognitive biases and force predictable errors.

**Spear phishing**

Spear phishing is an increasingly popular attack technique and is essentially a highly targeted phishing email.²⁶ Attackers will aim to establish credibility and trust and will have a specific outcome in mind, such as getting the target to trigger a malware link or enter credentials. The spear phishing exchanges will often include elements of informational power or referent power to persuade the target that their requests are legitimate and will try to exploit cognitive biases related to emotion, such as the affect heuristic.

**Whaling**

Whaling is a type of phishing email targeting a single high-value target, typically senior executives or those with privileged access. Attackers will typically play the ‘long game’, using methods of social power over a long period of time, often using expert power to establish credibility, attempting to trigger herd behaviour or using coercive power in the form of blackmail in return for privileged information.

**Baiting**

Baiting is in many ways similar to phishing attacks. However, what distinguishes baiting from other types of social engineering is the promise of an item that hackers use to entice targets. Baiters may offer users free music or movie downloads if they surrender their login credentials to a certain site. This type of attack uses reward power to draw targets in, before triggering the affect heuristic.

**Tailgating**

Tailgating is a more opportunistic form of attack, also known as ‘piggybacking’. These types of attacks involve someone who lacks the proper authentication following an employee into a restricted area. Attackers can also leverage legitimate power to persuade the target to let them into restricted areas, wearing official uniforms or creating branded logos and fake passes. The availability heuristic tends to be triggered during tailgating as targets make a judgement on the likelihood of the attacker causing a security incident.

**Smishing**

Smishing is social engineering via text message,²⁷ and will likely become more popular as users have become more aware of traditional email-based phishing attacks, but less aware of text message attacks.²⁸ Usage of mobile phones has increased exponentially worldwide, providing another means for attackers to use coercive power or reward power, potentially triggering optimism bias. Smishing attacks are currently less common and targets are unlikely to perceive this type of attack technique as a real threat.

**Vishing**

Vishing is social engineering using voice, and increased 350% between 2014 and 2018.²⁹ Attackers can apply any number of social power techniques using their voices to build rapport and trust instead of using email or text message, triggering the affect heuristic or bounded rationality biases, through the intimacy of a direct phone call. As the capabilities of artificial intelligence grow more complex, AI chatbots will become increasingly believable.³⁰ This means that artificially intelligent vishing attacks are likely to be on the horizon.

---


The way in which the attack technique uses social power to trigger cognitive biases will differ for each scenario. In some cases, a single email may be enough to trigger one or more cognitive bias resulting in a desired outcome. In others, the attack may gradually manipulate the target over a period of time using multiple techniques. What is consistent, however, is that the attacks are carefully constructed and sophisticated. By knowing how attackers use psychological methods, such as social power, to trigger cognitive biases and force errors, organisations can deconstruct and analyse real-world incidents to identify their root causes and therefore invest in the most effective mitigation.

Figure 6 demonstrates how a CEO fraud email exchange uses social power to convince the target that they are a credible source, triggering specific cognitive biases to fulfil the attacker’s desired outcome.

Figure 6: Example of CEO fraud email exchange

| Bob are you available to do this? The CEO is insistent that this is completed by close of play today! |
| Bob Adams |
| Alice Acres |
| Chief Operating Officer |
| Safehaven Financial |

<table>
<thead>
<tr>
<th>This message is high priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject: URGENT RESPONSE NEEDED</td>
</tr>
<tr>
<td>Alice Acres</td>
</tr>
<tr>
<td>Bob Adams</td>
</tr>
<tr>
<td>25 May 16:45</td>
</tr>
</tbody>
</table>

| I need a payment transaction completed by close of business, can you do this? I will send the account details over. |
| Alice Acres |
| Chief Operating Officer |
| Safehaven Financial |

<table>
<thead>
<tr>
<th>This message is high priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject: URGENT RESPONSE NEEDED</td>
</tr>
<tr>
<td>Alice Acres</td>
</tr>
<tr>
<td>Bob Adams</td>
</tr>
<tr>
<td>25 May 16:51</td>
</tr>
</tbody>
</table>

| Bob are you available to do this? The CEO is insistent that this is completed by close of play today! |
| Bob Adams |
| Alice Acres |
| Chief Operating Officer |
| Safehaven Financial |

| Hi Alice, |
| Bob Adams |
| Financial Accounts Admin |
| +44 (0) 211 872 882 |
| Safehaven Financial |

| Affect heuristic: |
| By making the email look urgent it provokes an emotional rather than a rational response. |

| Legitimate power: |
| An email from the COO, with the company signature legitimates the request. |

| Referent power: |
| Referencing the CEO confirms the demands in the eyes of the target. |

| Bounded rationality: |
| Placing a time constraint pushed towards a ‘good enough’ decision rather than a well thought-out one. |
Information Security Forum
Human-Centred Security: Addressing psychological vulnerabilities

For information security programmes to become more human-centred, organisations must become aware of cognitive biases and their influence on decision making.\(^{31}\) They should acknowledge that cognitive biases can arise from normal working conditions but also that attackers will use carefully crafted techniques to manipulate them for their own benefit. Organisations can then begin to readdress information security programmes to improve the management of human vulnerabilities, and to protect their employees from a range of coercive and manipulative attacks.

---

3 Managing human vulnerabilities

Human vulnerabilities, whether triggered through work pressure or by a malicious attacker, can lead to errors that can significantly impact an organisation’s reputation or even put lives at risk. Organisations can strengthen information security programmes in order to mitigate the risk of human vulnerabilities by adopting a more human-centred approach to security awareness, designing security controls and technology to account for human behaviour, and enhancing the working environment to reduce the impact of pressure or stress on the workforce. Organisations should consider any of the below emerging security solutions when addressing human vulnerabilities.

DEVELOP A HUMAN-CENTRED SECURITY CULTURE

Reviewing the current security culture and perception of information security should give an organisation a strong indication of which cognitive biases are impacting the organisation. Increasing awareness of human vulnerabilities and the techniques attackers use to exploit them, then tailoring more human-centred security awareness training to account for different user groups should be fundamental elements of enhancing any information security programme.

Review and improve the culture and perception of information security

Carry out a review of the organisation’s culture, starting from the most senior positions, considering the organisation’s mission statement and values regarding information security. This should enable a better perspective of how different areas of the organisation value information security and identify areas that have significant human vulnerabilities that require training and development.

Prequalify knowledge and tailor employee awareness

Gain perspective of current employee awareness and understanding of information security behaviours, threats and values by testing knowledge of cognitive biases. Results may be unique to one area of the business or even to individual teams or departments. Individuals will also exhibit different cognitive biases during different scenarios. This could indicate where training and awareness campaigns may need to be tailored to specific audiences and activities.

Identify threats and run exercises

Identify key risks and threats to the business and/or industry area. Targeted scenario planning and training (e.g. table-top exercises or phishing campaigns) can then be tailored to improve responses to specific threats and enable employees to better react to stressful situations that may trigger cognitive biases.

Cognitive biases that may be mitigated:

- Herd behaviour
- Optimism bias
- Over-justification effect
- Licensing effect
- Optimism bias
- Polarisation
- Affect heuristic
- Availability heuristic
- Bounded rationality

**Promote critical thinking**

Produce and deliver materials that promote critical thinking and awareness of cognitive biases. This type of training should develop analytical, communication and problem-solving skills. This reinforces creativity and open-mindedness, reducing the risk of errors being made.

**Cognitive biases that may be mitigated:**
- Anchoring
- Ego depletion
- Polarisation

---

The ISF briefing paper *Delivering an Effective Cyber Security Exercise* provides an overview on how to prepare, run and follow-up on cyber security exercises.

**DESIGN A HUMAN-CENTRED APPROACH TO SECURITY**

The design of technologies, controls, policies and procedures has to be suitable for human interaction. It should guide users to good security behaviours instead of potentially triggering cognitive biases.

**Understand users’ interactions with technology, controls and data, and invest accordingly**

Gain a clear understanding of the current controls and technologies that employees interact with and consider their suitability and usability, locating single points of failure or specific touchpoints that may trigger cognitive biases. Consider the environment employees work in, the location and cultural setting, as interactions may differ across cultures and geographies. This should enable controls and technologies to be designed or enhanced around the users.

**Cognitive biases that may be mitigated:**
- Anchoring
- Bounded rationality
- Choice overload

---

**Restructure processes and policies**

Explore restructuring processes and policies to allow the workforce to easily and quickly determine a logical route to an appropriate behaviour. Identify areas of weakness or ‘pain points’ that have historically led to human error or circumvention of processes and policies.

**Cognitive biases that may be mitigated:**
- Bounded rationality
- Choice overload
- Herd behaviour

---

**Review successful external attacks**

Analyse real-world attacks, highlighting what techniques the attackers used, and which cognitive biases were subsequently triggered. Industry sector or peer group comparisons should be made, and lessons learned should be incorporated into awareness material.

**Cognitive biases that may be mitigated:**
- Affect heuristic
- Availability heuristic
- Optimism bias

---

Automate but maintain oversight

Automate some aspects of information security to reduce the impact of human errors being made, e.g. spam email protection. Automation should never be done in isolation and an oversight procedure should be in place to enable human review.

The ISF report *From Promoting Awareness to Embedding Behaviours: Secure by choice, not by chance* provides an overview of how to embed secure information security behaviours in organisations by reviewing and restructuring processes and policies.

**CREATE A HUMAN-CENTRED WORKING ENVIRONMENT**

Organisations with successful human-centred security programmes often have significant overlap between information security and human resource functions. The promotion of a strong mentoring network between senior and junior employees, coupled with the improvement of the structure of working days and the work environment, should help to reduce unnecessary stress that leads to the triggering of cognitive biases affecting decision making.

**Promote mentoring schemes**

Develop meaningful relationships between a mentor and mentee to create an equilibrium of knowledge and understanding. These relationships also serve as a support system that has the potential to reduce risky behaviour by providing multiple perspectives on a problem or task, which significantly helps reduce the impact of cognitive biases.38 Organisations can also create the role of ‘security champions’ – individuals who are not in the information security function but champion positive security behaviours in their teams or to their peers within the organisation.

**Re-evaluate and change the structure of the working day**

Create a working environment and work-life balance that reduces stress, exhaustion, burnout and poor time management, which all significantly increase the likelihood of errors being made. Restructuring the working day to include periods of meaningful rest, such as naps,39 has significant potential to reduce cognitive biases that affect information security behaviours, such as decision fatigue or ego depletion. Changes may also include cultural or environmental considerations.

---

38 “5 ways to reduce unconscious bias in the workplace”, The Insurance Institute, 3 May 2017, https://blog.iii.ie/inside-track/5-ways-to-reduce-unconscious-bias-in-the-workplace
**Improve workspace and environment**

Consider how the improvement or enhancement of workspaces and environments can reduce stress or pressure on the workforce. Consider what is the most appropriate work environment for the workforce as there may be varying options, e.g., working from home, remote working, or modernising office spaces, factories or outdoor locations.

The ISF briefing paper *Building Tomorrow’s Security Workforce* provides guidance on how to create an engaging working environment and how to develop a progressive working culture.

Member interviews identified that no organisation will have the same approach to managing human vulnerabilities, with organisations experiencing different risks and threats, and working to different budgets and requirements. Interviews also identified that some organisations proactively consider psychology in their training and awareness, with understanding and managing cognitive biases playing a key role.

Organisations that are already taking a human-centred approach to information security typically spend extended periods of time observing human interaction with technology, controls and data, to identify which specific cognitive biases are triggered, and understanding why this is the case. This has enabled effective and targeted investment in human-centred security improvement programmes which prioritise the highest risk areas.

There is, however, insufficient good practice in order to identify which solutions merit more investment than others, so it will depend on the organisation, the specific human vulnerabilities that lead to errors in decision making, and the most common types of attacks experienced.

**Cognitive biases that may be mitigated:**

- Decision fatigue
- Ego depletion
- Herd behaviour

> The ISF briefing paper *Building Tomorrow’s Security Workforce* provides guidance on how to create an engaging working environment and how to develop a progressive working culture.
Underlying psychological vulnerabilities mean that humans are prone to both making errors, and to manipulative and coercive attacks. Errors and manipulation now account for the majority of security incidents, so the risk is profound. By helping staff understand how these vulnerabilities can lead to poor decision making and errors, organisations can manage the risk of the accidental insider. To make this happen, a fresh approach to information security is required.

A human-centred approach to security can help organisations to significantly reduce the influence of cognitive biases that cause errors. By discovering the cognitive biases, behavioural triggers and attack techniques that are most common, tailored psychological training can be introduced into an organisation’s awareness campaigns. Technology, controls and data can be calibrated to account for human behaviour, while enhancement of the working environment can reduce stress and pressure.

Once information security is understood through the lens of psychology, organisations will be better prepared to manage and mitigate the risks posed by human vulnerabilities. Human-centred security might just help organisations transform their weakest link into their strongest asset.

WHERE NEXT?
The ISF encourages collaboration on its research and tools. ISF Members are invited to join the Human-Centred Security community on ISF Live to share experiences.
Human-Centred Security: Addressing psychological vulnerabilities
July 2019

AUTHOR
Dan Norman

CONTRIBUTOR
Jordon Kelly

ABOUT THE ISF
Founded in 1989, the Information Security Forum (ISF) is an independent, not-for-profit association of leading organisations from around the world. It is dedicated to investigating, clarifying and resolving key issues in cyber, information security and risk management and developing best practice methodologies, processes and solutions that meet the business needs of its Members.

WARNING
This document is confidential and is intended for the attention of and use by either organisations that are Members of the Information Security Forum (ISF) or by persons who have purchased it from the ISF direct. If you are not a Member of the ISF or have received this document in error, please destroy it or contact the ISF on info@securityforum.org. Any storage or use of this document by organisations which are not Members of the ISF or who have not validly acquired the report directly from the ISF is not permitted and strictly prohibited. This document has been produced with care and to the best of our ability. However, both the Information Security Forum and the Information Security Forum Limited accept no responsibility for any problems or incidents arising from its use.

CLASSIFICATION
Restricted to ISF Members, ISF Service Providers and non-Members who have acquired the report from the ISF.