The psychology of social engineering

Received (in revised form): 11th November, 2022

Barry Coatesworth
Director, Guidehouse, UK

Barry Coatesworth is a Director with Guidehouse’s Global Energy, Sustainability and Infrastructure (ES&I) Practice. He has over 30 years’ experience in energy, finance and public sector working within risk, compliance and security. An internationally recognised cyber security expert and adviser, Barry is part of the Industry Advisory Group for the Cyber Essential Scheme in conjunction with UK’s Department of Business, Energy and Industry Strategy, as well as a research adviser to the Parliamentary Office of Science and Technology (POST) on big data, ethics and privacy.

Guidehouse, Levels 7 & 8 Angel Court, 1 Angel Court, London, EC2R 7HJ, UK
Email: bcoatesworth@guidehouse.com

Abstract Social engineering is an ever-growing threat to organisations and people. This paper discusses the psychology behind social engineering and why it is still an effective strategy for criminals, nation states and hacktivists. The tactics, techniques and procedures (TTPs) described in this paper may help you identify threat actors/groups and aid in identifying emerging threats and developing appropriate countermeasures and awareness.

KEYWORDS: social engineering, information security, cyber security, psychology, cognitive bias

TACTICS, TECHNIQUES AND PROCEDURES USED BY ADVERSARIES

Perhaps the earliest record of a social engineering attack comes from the Greek myth of the Trojan horse. As the story goes, in 1184 BC, during the Trojan War, the Greeks departed the city of Troy in ships, leaving behind a large wooden horse as a victory offering.¹ When it was hauled inside the walls of Troy and Greek soldiers descended from the horse’s belly after dark to slay the guards and begin destroying the city, the ancient world witnessed one of its most famous perimeter breaches. This is an example of the social engineering technique of ‘reciprocity’, or gifting, in which the giver of a gift usually wants something more valuable — in this case, access to the city — in return.

As with the Trojan horse, this strategy can produce dire outcomes for the gift recipients. In more recent times, starting in the 20th century, there have been numerous wildly original social engineering scams. In the Eiffel Tower ‘sale’ of 1925, Victor Lustig, a charming con artist, travelled to Paris and chanced upon a newspaper article discussing the challenges of maintaining the Eiffel Tower. This gave him inspiration for a new con. Lustig invited a small group of scrap metal dealers to a confidential meeting, whereupon he identified himself to them as the deputy director-general of the Ministère de Postes et Télégraphes (Ministry of Posts and Telegraphs). In the meeting, he convinced the men that the upkeep of the Eiffel Tower was becoming too much for Paris and that the French government wished...
Coatesworth

to sell it for scrap in an auction. To con the unsuspecting scrap metal dealers back in 1925, Lustig hired a counterfeiter who created ‘official’ stationery for his imaginary role. Today, counterfeiting documents to misrepresent who you are or what you have done is much easier using digital media. Ultimately Lustig fled to Austria with the money he conned, and after his unsuspecting victim was too embarrassed to report it, he returned to Paris again to carry out the same stunt. This time his victim went to the police, and Lustig fled to the US.

Social engineering has clearly been around for ages, and in most cases it is not malicious. At its core, social engineering is simply the building and leveraging of influence to persuade others to act as you want them to. Most of us have used it at one time or another to influence those around us.

The rise of social media ‘influencers’ is a good example of social engineering in the digital age. But what is the significance of social engineering in the context of information security? The online resource for security professionals ‘Security Through Education’ defines social engineering as ‘Any act that influences a person to take an action that may or may not be in their best interest’, and lists things such as phishing and impersonation as examples.2 In the digital world, this type of manipulation ultimately works by compromising one of information security’s foundational pillars: trust. When we trust, we want to cooperate. This urge stems from our basic human instincts, because by cooperating, by establishing trust, we have a better chance of survival. Self-described ‘public-interest technologist’ Bruce Schneier describes this human trait well in one of his essays:

‘Humans are a trusting species. There were 120 people on my plane, almost all of them strangers to each other, and at no point did anyone jump up and attack the person sitting next to them. It’s absurd for me to even say it, but if we had been a planeload of chimpanzees, that would have been impossible. Trust is essential for society to function—our civilization would collapse completely without it—and the fact that we don’t think about it is a measure of how well that trust works.’3

It is the same with computer systems, where we have trust relationships — secure interactions in which information is passed between systems. We trust user accounts to access sensitive data, and have trusted certificates to protect our online banking, for example.

Many businesses, however, are moving away from an overreliance on trust and adopting zero trust models. The security and risk news website CSO defines zero trust as ‘a strategic approach to cybersecurity that secures an organization by eliminating implicit trust and continuously validating every stage of digital interaction’.4 It has been estimated that 98 per cent of cyberattacks involve some form of social engineering, and the average organisation is targeted by an astounding 700+ cyberattacks annually. On average, successful social engineering-driven attacks cost companies US$130,000 through money theft or data destruction.5

So why are these breaches of information security so successful? It is not just a technology problem, nor one that can be easily solved by technology alone. Technical solutions can only reduce the potential likelihood and impact of such crimes. A true analysis of what is behind the success of social engineering attacks must begin with a study of the human factor and psychological influences.

COGNITIVE BIAS AND HEURISTICS

Humans are complex creatures and our susceptibility to social engineering is not because of one factor — multiple psychological factors and biases influence our decision making; however, one common
The psychology of social engineering

trap that all humans are susceptible to on a psychological level is cognitive bias.

Cognitive bias can take various forms: confirmation bias, hindsight bias, self-serving bias, anchoring bias and availability bias are some of the most common examples that make people vulnerable to social engineering attacks.

The reasons for our poor decision making can be a consequence of heuristics and biases. In general, heuristics and biases describe a set of decision-making strategies and the way that we weigh certain types of information.

We are often presented with situations in life when we need to make a decision with insufficient information, and we subconsciously prejudice or bias that decision. Cognitive bias is a systematic or heuristic shortcut that occurs when we are interpreting information in a judgmental task and arises from problems related to memory, attention and other mental mistakes.

In addition, confirmation bias causes a person to seek out or interpret information that confirms the preconceptions that they have. People tend to make irrational decisions based on their past rational ones, and we see what we expect or want to see.

Anchoring bias is where people accept the first piece of information as truthful while arriving at a decision. An employee can easily fall into thinking that if an attacker presents ‘evidence’, this creates a sense of authenticity — for instance, if an attacker were to call claiming to be from the IT department and name-drop a familiar manager’s name. The name-drop can be enough for the employee to anchor on this information and give the attacker information or access to systems.

Cybercriminals can use these biases to manipulate their target’s perception to convince that person to engage in risky behaviours, such as clicking on a link they normally would not click on or entering sensitive information on a website.

Cialdini offered seven principles of persuasion that increase the likelihood that a social engineer will succeed: reciprocity, conformity, liking, scarcity, commitment, authority and unity. Within these principles, Cialdini shows that since many individuals are predisposed to trusting others they perceive as likable or those whom they have identified as authority figures, they are more likely to fall for social engineering attacks. Cialdini’s principles of influence using cognitive bias are the most common framework used in social engineering attacks, so we will look at these more closely.

THE PRINCIPLES OF SOCIAL ENGINEERING

In 1984, Robert Cialdini, a behavioural psychologist, proposed a concept called the ‘theory of influence’ in his book *Influence: The Psychology of Persuasion*. He found that influence is based on six key principles: reciprocity, commitment/consistency, consensus/social proof, authority, liking and scarcity. In 2016, he proposed a seventh principle: unity.

These seven principles have become integral to political and other social engineering efforts. To give just one example, in 2014, Edward Snowden’s leaks of US intelligence documents included a classified presentation from the UK’s Government Communications Headquarters (GCHQ) called ‘The Art of Deception: Training for Online Covert Operations’, which draws heavily on the psychology of influence and persuasion.

This presentation, which seems to have been put together by GCHQ’s Human Science Operations Cell, lists several of Cialdini’s six principles (reciprocity, social compliance/authority, and consistency). Most of the remaining principles (see Figure 1) are taken from Stajano and Wilson’s classic study ‘Understanding scam victims: Seven principles for systems security’, which describes six methods used by con artists. One item — authority — from Stajano and Wilson’s study overlaps with Cialdini’s principles, and the presentation additionally
10 Principles for Influence

Reciprocity
Reciprocity is also known as ‘gifting’, and it works because people do not like to feel indebted to others. When we receive a favour, we tend to try to repay it. Simply put, people are obliged to give back to others, either in the form of a behaviour, gift or service that they have received first, or its equivalent.

If a friend invites you to their party, there is an obligation for you to invite them to a future party you are hosting. If a colleague does you a favour, then you owe that colleague a favour. Furthermore, in the context of a social obligation, people are more likely to say yes to those they owe.

One of the best demonstrations of the principle of reciprocity comes from a series of studies conducted in restaurants. Researchers found that the custom of servers bringing a small gift, such as a fortune cookie or mint, with the bill effectively increased the chance of tipping.

Scarcity
Simply put, people want more of the things that they feel there may be less of.
When British Airways announced in 2003 that it would no longer be operating the twice-daily London–New York Concorde flight because it had become uneconomical to run, sales the very next day took off. Notice that nothing had changed about the Concorde itself. It certainly did not fly any faster, the service did not suddenly get better and the airfare did not drop. It had simply become a scarce resource. And as a result, people wanted it more.

Scarcity can also include a time constraint, for example a 24-hour sale, the use of limited-edition products and sometimes both (‘24-hour limited-edition sale!’). When people believe things are in short supply, they are more likely to feel the need to have them.\(^9\)

The scarcity principle limits the number of opportunities we have available to us. During the COVID-19 pandemic, phishing e-mails were sent around due to a shortage of personal protective equipment (PPE), eg face masks; as opportunities to purchase masks decreased, we were more inclined to want to purchase them and to click on the e-mails.

**Authority**

This is the idea that people follow the lead of credible, knowledgeable experts.

Physiotherapists, for example, can persuade more of their patients to comply with recommended exercise programmes if they display their medical diplomas on the walls of their consulting rooms. People are more likely to give change for a parking meter to a complete stranger if that requester wears a uniform rather than casual clothes.

Authority is probably the most plausible, obvious principle since most people have complied with authority at some point in their lives. We are more likely to follow what someone above us in the hierarchy — a chief executive officer (CEO) or manager, for instance — asks us to do, because doing otherwise creates the risk of repercussions, such as getting fired, losing a bonus, etc.

**Liking**

People prefer to say yes to those who they like. But what causes one person to like another? Persuasion science tells us that there are three important factors: we like people who are similar to us, we like people who pay us compliments, and we like people who cooperate with us toward mutual goals.

Author Lois McMaster Bujold’s phrase ‘If you make it plain you like people, it’s hard for them to resist liking you back’\(^10\) describes the liking principle perfectly. We prefer to comply with requests from people we know and like.\(^11\) It is a fundamental human motive to create and maintain relationships with others. This principle explains why giving a compliment can improve the odds of getting a favour.

**Commitment and consistency**

People like to be consistent with the things they have previously said or done and like to maintain consistent behaviour.\(^12\) As Cialdini states, ‘Once we have made a choice or taken a stand, we will encounter personal and interpersonal pressures to behave consistently with that commitment’.\(^13\)

When it comes to social engineering, if we accept a contact or connection request, we seldom will break the connection, unfollow or unfriend the person, and we are likely to continue to interact with them.

**Consensus or social proof**

Especially when they are uncertain, people will look to the actions and behaviours of others to determine their own.

People tend to do what they believe everyone around them is doing, particularly when they are unsure of what to do in the first place.\(^14\) If everyone in the office is wearing a security ID badge, we are more likely to wear one too. If we eat out with a group of people in an unfamiliar place or cannot read the menu, we are more likely to order what everybody else is eating or drinking.
Unity

We gravitate toward people whom we identify as being similar to us.15 According to Cialdini, the unity principle moves beyond surface-level similarities (which can still be influential but fall under the liking principle). Instead, he says, ‘It’s about shared identities’.16

In a way, the unity principle boils down to the third step on psychologist Abraham Maslow’s hierarchy of needs: the need to belong (see Figure 2).

Have you ever been at a party or conference and met someone who went to the same university as you? Or maybe you previously worked at the same company, or were both in the military or government service? You probably felt an instant connection.

When we belong, or feel we belong, to a group, we are likely to be more open to persuasion attempts.

In practice, these persuasion principles are often used in combination. For example, a marriage proposal entails commitment and reciprocity. Even a simple ‘I love you’ demonstrates commitment and liking. Similarly, attempted cyberattacks involving social engineering will typically involve a combination of the above principles. We will explore how these principles shape real-world examples of social engineering tactics, first by taking a closer look at the structural aspects of social engineering.

THE STRUCTURE OF SOCIAL ENGINEERING

There are four main steps in the social engineering life cycle (see Figure 3) that a would-be social engineer might use to manipulate and exfiltrate information from a target.
During the life cycle, the attacker will decide which cognitive bias and psychological approach will have the most success on their victim, whether it is an authoritarian influence of a senior manager or government official, to one that uses commitment or consensus, this normally start to take shape during the second phase when establishing a relationship with the victim.

**Information gathering/research**

To create a credible persona, the social engineer will develop a plan that involves extensive use of open-source intelligence (OSINT) to learn about the target’s likes/dislikes, social friends, subscriptions and personal information, etc., with the goal of combining these into a social engineering package to instil and build trust with the target.

An attacker could impersonate members of your IT team, like the hacker in Uber’s compromise successfully did to gain access to the company’s systems. They may also impersonate an employee, such as your CEO, or a supplier or friend. Doing this well requires patience, dedication and time, instilling trust credibility into the relationship.

Examples of credible behaviours that emerge from open-source intelligence include:

- Knowledge of your personal details, name, date of birth, address, etc.;
- Knowledge of who you work with in your department;
- Knowledge of your technology usage (Internet searches of partnerships, company news, etc.) (usually used to impersonate a supplier or vendor);
- Knowledge of personal and/or professional relationships from LinkedIn or other sources (usually used to impersonate mutual friends or colleagues).18

The attacker will begin to connect the dots into a social engineering package to establish a successful engagement with the target, which can sometimes just be for them to click on a link in an e-mail.

**Establishing a relationship/engagement**

When we first encounter someone new, there are typically a few key details that we pay attention to that help us form an opinion about whether we can trust them. We might ask ourselves:

- Why is this person contacting me/approaching me?;
Do we have any mutual friends or associates?
• Does this person appear trustworthy?
• Does this person have any authority?  

Social engineers ask themselves the same questions and more when they plan an attack; the answers will help them decide which principles of influence will work best on the target. Dedicated attackers will score each possible answer to their questions, sometimes using Bayesian analysis techniques to model their success factors. This can significantly improve a successful chance of engagement and render the target more vulnerable to manipulation and exploitation.

Exploitation
This stage of social engineering involves developing a trusted relationship with a target by using a social engineering package/persona curated from the first two stages.

For a social engineer, building rapport with a target requires the successful utilisation of the principles of influence. It is at this point that the attacker has developed several layers of trust with the victim, who is most vulnerable to exploitation.

Exploitation could take the form of:
• Phishing attacks;
• Spear phishing;
• Whaling;
• Smishing and vishing;
• Baiting;
• Pretexting;
• Quid pro quo (ie tech support scams);
• Honeytraps (romance scams);
• Watering holes.

Execution/escape
Following a successful attack, attackers will usually break off all communication with their target and start to cover their tracks. This final structural element of social engineering achieves closure for the criminal, while leaving targets with little recourse (see Figure 4).

As social engineering attacks have increased in number, they have also increased in variety. Some of the most common attack types seen today include phishing, watering holes, pretexting, baiting and quid pro quo attacks.

TYPES OF SOCIAL ENGINEERING ATTACKS
Phishing
Phishing is the most common type of social engineering attack. Attackers use e-mails, social media and instant messaging, etc. to manipulate victims into providing information or visiting malicious websites.

Phishing attacks usually have the following common characteristics:
• Messages are composed to stimulate curiosity and attract the user’s attention;

![Figure 4: The steps of a social engineering attack](image-url)
• Phishing messages can convey a sense of urgency (CEO fraud scams often target or impersonate CEOs or CFOs and require an urgent money transfer or payment);
• Attackers use shortened website links to direct victims to a malicious website that could host exploitative codes (waterhole attack);
• Attackers can spoof the e-mail address of the organisation or sender and incorporate logos, images, fonts and styles used on the legitimate website.

Watering hole
A watering hole — sometimes known as a strategic web compromise (SWC) attack — consists of a compromised website that contains malicious code or malware. When a victim visits the page on the compromised website, malware (usually a Trojan) is installed on the user’s computer. The attackers compromise websites, usually within a specific sector (e.g., defense or energy), that are likely to be visited by the target. The Dragonfly cyber espionage group successfully used waterhole attacks to compromise the Western energy sector and Ukraine’s power systems in 2015–16.

A watering hole method of attack is not commonly used by cybercriminals and hackers and is more commonly used for nation state-sponsored attacks.

Pretexting
Pretexting is a form of social engineering in which an attacker tries to persuade a victim to give up valuable information or access to a service or system. It is a form of impersonation that heavily relies on the use of authority to manipulate the victim.

Pretexting is what most often happens with data breaches from inside an organisation, when someone creates a fake persona or misuses their actual role. For example, a target might be asked by a superior or co-worker for their passwords when they go on holiday.

Edward Snowden infamously told his co-workers that he needed their passwords as their system administrator. Victims, respecting his authority and job title, willingly complied without giving it a second thought.

These attackers establish trust using their perceived authority, then persuade victims to give them sensitive data.

Pretexting has a fairly long history; in the UK, where it is also known as ‘blagging’, tabloid journalists have used this technique for years to gain access to information on celebrities and politicians. Pretexters tend to use e-mails, SMS and voice calls to manipulate and influence their victims.

Baiting
Baiting is a technique that stimulates and exploits curiosity. The most common form of baiting uses USB drives in what is called a USB drop attack. A drive is dropped either in a car park or other place that the target frequents. It is placed where the target will notice it and have an incentive curiosity to insert it into a personal or work system.

Quid pro quo attacks
A quid pro quo attack (aka ‘something for something’ attack) promises something to the victim in exchange for help; it is similar to baiting. In the most common quid pro quo technique, an attacker impersonates a member of an organisation’s IT staff. The attacker offers the target some kind of free upgrade or software, then might ask victims to facilitate the operation by disabling their antivirus software, clicking on a link or giving them a password.

Classic phishing e-mails
A closer look at the most popular phishing schemes reveals that aspects of Cialdini’s
principles of influence serve as the foundational psychology underlying most of them. The following is an analysis of how Cialdini’s principles are at work in different phishing schemes.20

**Reciprocity**
The influence of reciprocity used in phishing attacks. In this example (see Figure 5), an attacker uses a free coupon as a gift and then asks the user to sign up for an account.

**Scarcity**
The influence of scarcity is used in phishing attacks. In this example (see Figure 6), the attacker is using a limited number of bottles of champagne to get the victim to create an account and steal their password.

**Authority**
The influence of authority used in phishing attacks. In this example (see Figure 7), the attackers are masquerading as the CEO in order to initiate a money/wire transfer.

**Commitment and consistency**
The influence of commitment and consistency used in phishing attacks. In this example (see Figure 8), an e-mail with the Amazon logo says your order has been cancelled and that you need to reactivate your account. The email is consistent with Amazon’s email style.

**Liking**
The influence of liking used in phishing attacks. In this example (see Figure 9), the attacker has compromised a social media account.

---

![ASDA direct email](image)

**Figure 5:** Reciprocity in phishing21
Hi All

Subject: Free Cristal Champagne!

I am excited to announce that Louis Roederer have sponsored my new blog on fine wine, Champagne and Formula 1 news!

Please check out my blog [www.wordpress.com](http://www.wordpress.com) and let me know what you think, your feedback is invaluable and greatly appreciated 😊

As a thank you for the first 50 people that signup and register will receive a bottle of Cristal champagne!

LOUIS ROEDERER
CHAMPAGNE

Love
Eve
xxx

Figure 6: Scarcity in phishing
Source: Author

Figure 7: Authority in phishing

Figure 8: Commitment and consistency in phishing

Figure 9: Liking in phishing
account and is asking the parents to send them money.

HOW TO PROTECT AGAINST SOCIAL ENGINEERING

There is no silver bullet to stop social engineering attacks. An organisation’s greatest strength and greatest weakness is its people. So, what is the most effective social engineering defence?

There are numerous factors, and the best defence is a combination of both technical controls and informed and educated employees.

The US Cybersecurity & Infrastructure Security Agency (CISA) has some great common sense tips for avoiding social engineering and phishing scams:

• Be suspicious of unsolicited phone calls, visits or e-mail messages from individuals asking about employees or other internal information. If an unknown individual claims to be from a legitimate organisation, try to verify their identity directly with the company/agency;
• Do not provide personal information or information about your organisation, including its structure or networks, unless you are certain of a person’s authority to have the information;
• Do not reveal personal or financial information in e-mails, and do not respond to e-mail solicitations for this information. This includes following links sent in e-mails;
• Do not send sensitive information over the Internet before checking a website’s security;
• If you are unsure whether an e-mail request is legitimate, try to verify it by contacting the company directly. Do not use the contact information provided on a website connected to the request; instead, check previous account statements for contact information;
• Educate yourself and all employees on the types of attacks out there.25

Social engineering has been around for centuries, and we have all used it generally to influence others in a positive way, but even the most cautious and perceptive of us can get caught by social engineers.

Because most people are trusting by nature, it takes more than antivirus software and education to protect against phishing and other social media attacks. You also need to be vigilant, cautious, and aware of the psychological tactics that are being used against you.

References

9. Ibid.
11. Ellis, ref. 8 above.
12. Ibid.
13. Cialdini, ref. 6 above.
14. Ellis, ref. 8 above.
15. Ibid.
16. Cialdini, ref. 6 above.
19. Ibid.