Name of theory
Six weapons of influence

Key reference
Book: (Cialdini, 1993)
Article: (Cialdini & Goldstein, 2002)

Description
Robert Cialdini was interested in the various ways in which people are persuaded to comply with a request. To investigate what methods expert persuaders (e.g., salespeople) use to sell products to customers, he did extensive fieldwork observing these practices. He then posited his 'six weapons of influence', tactics that are used to make it more likely that someone will say yes to a request. These six weapons are: authority, liking, social proof, reciprocity, scarcity, and commitment & consistency.

To increase the probability that people will comply with a request, the person making the request can present the request as coming from an **authority**, be it an actual authority figure such as a CEO, scientist (e.g. in washing deterrent commercials) or other widely acknowledged authority. Another element of the person making the request is that they can attempt to be perceived as more **likeable**. Liking the person who makes the request will increase compliance. This liking can not only take the form of being nice to someone, but also be based on looks (attractive people are perceived as more likeable, see Halo Effect), or similarity between the requester and the person being asked to comply.

When making the request, someone can also point at the behaviour of others. The element of **social proof** suggests that people are likely to act in similar ways to the people around them. This is especially likely when people are not sure how to act, for example when the behaviour or situation is new to them. Pointing out what others are doing in the form of "XX% of people like you choose Y" can increase compliance. Minimising an initial request and later ask for a larger favour (foot in the door technique) is one of the ways in which **commitment and consistency** is used to persuade people to comply with requests. By getting people to say yes to a small request, they are more likely to say yes to a second, larger request as well.

If the request can be framed so that it is doing someone a favour who has earlier done you a favour, the request is also more likely to be successful. This is the weapon of influence called **reciprocity**. By first offering something, and only then asking for something in return, people are more likely to comply as they are eager to 'settle the score' and not be in someone's (moral) debt. A final method is by making the commodity you want to sell, or the option you want people to choose, scarce. **Scarcity**, ensuring that there are limits to when or how often people can choose something, makes that option more attractive.

Application within the field of cybersecurity
The field of social influence is used in cybersecurity from both the offender's as well as the defender's side. Below, the use of these six techniques are outlined for both attacking and defending purposes, as knowing how attacks work is useful when attempting to prevent these attacks from being successful.

Offensive use

The six weapons of influence are mostly used in scams and phishing messages (messages that attempt to steal a user's information, such as creditcard details, or log in credentials). **Authority** is used in scams by impersonating a trustworthy source, such as a bank or national agency. In direct face-to-face communication, authority can also be activated through clothes, looks and manners of the scammer, or by positing as a trustworthy service employee (plumbers, road workers etc.). **Social proof** is used in scams to suggest that many others have already gone before you, and therefore it is safe to act in a similar way. For example, fake advertisements on social media suggesting people should invest in crypto (e.g. Bitcoin), where they posit the number of people already making an investment, suggesting that a large group of people is in fact choosing this option. **Liking** is mostly used by social engineers, people who try to use social interactions to get access to systems, information, or physical locations that they should not have access to. For example, a social engineer might try to invoke empathy by stressing how overworked they are, or the (fake) predicament they are in. Furthermore, social engineers use liking techniques like 'talking shop' with someone, sharing vague stories that make them seem similar to their target. **Scarcity** is one of the most present and most used techniques to encourage people to go along with a scam or to favourably reply to a phishing email. Almost all phishing attempts and scams will try to limit the timeframe in which the target can respond, to ensure that the target does not take time to ask for help or advice of others or has the time to realise that it is a scam. This is usually achieved by instilling a sense of urgency (e.g. "if you don't act now, your account will be blocked" or "we see money being transferred, if you don't act now, we won't be able to stop that from happening".) **Commitment and consistency** can be found in scams that are happening over time. For instance, in romance scams, the target is persuaded by a scammer that the scammer is in love with them. They will have prolonged chats over the phone or via messaging services, and can decide to meet up, only for some incident to occur why the scammer cannot be there, unless some money is transferred (no money for the train journey, additional costs at border control etc.). After the money is transferred, the scammer makes up new reasons why they cannot visit, citing higher costs, asking for more money, resulting in large sums of money being transferred over time. The final factor, **reciprocity**, is used by social engineers who do something for their target before making their request. For example, a social engineer might open an unlocked door for their target with the hope that the target will then open the next door for the social engineer, usually a door that might require an access code or badge.

Defensive use

On the defence side, the various techniques can be employed to persuade end-users to behave more security. For instance, in communication about cyber-risks, threats, and incidents, it is helpful to deliver the message through people and organisations who are considered an **authority**. This can be a National Cybersecurity Centre for the larger incidents, but also individuals within teams who are considered to have a natural authority within that team and can make changes to procedures and practices. The effect of sharing of information can be further enhanced by ensuring that the stories or incidents are framed so that people can empathise with the victims. This increases the **liking** of these victims and might motivate users to behave more secure, so that these incidents do not happen again. **Social proof** can be used in communication around common security behaviours and practices, in the form of "X of your colleagues lock their screen when they leave their desk, ensuring that no unwanted eyes see company secrets", or the example of Facebook where a social norms message can

be used to persuade users to check their privacy settings (see slide deck). **Commitment and consistency** can be used in a wider communication strategy. To avoid overload of information, and end-users avoiding cybersecurity policies altogether, the use of small changes in the beginning, and then moving to larger changes is recommended. (e.g. starting by committing to locking screens, then doing it at least once a day, then doing it from now on every time you leave your desk). A further way to enhance the communication strategy is by using **reciprocity**. In many organisations, the IT department is frustrated by the lack of attention to cybersecurity from end-users, while end-users on the other hand see the IT department is being a hinderance. Adjusting the communication strategy to one where the IT department can show all the steps they take to reduce the cyber-risk, before asking end-users to do their part as well, could be a way to persuade end-users to better comply with security policies. **Scarcity** is not commonly used as a defence mechanism in cybersecurity, but perhaps methods around restricting training possibilities at first, can increase the demand for training activities by end-users.

Scarcity/commitment and consistency/reciprocity

Annotated bibliography

Cialdini (1993). This is the book that Robert Cialdini wrote on the six weapons of influence. It clearly outlines the various methods and the research underlying these factors using examples and stories from own experimental studies. It is an accessible read but might be too lengthy for some classes depending on the other literature on the list.

Cialdini & Goldstein (2002). This article covers the same six principles as Cialdini outlined in his book, but in a condensed article format.

Cialdini (2016). This is another book by Cialdini, covering the pre-suasion context, but more importantly offering a seventh weapon of influence: unity. Unity is the notion that people can achieve more when they are working in a team, or other group that has a unified goal. This factor has also been incorporated in the revised edition of Cialdini (1993).

Dove (2021). A short, easy to read book on the psychology of fraud and scams. This book outlines the various ways in which psychological methods are used to perform scams, including the six weapons of influence (see chapter 6).

References

Cialdini RB. (1993) *Influence: Science and practice*. 3rd ed. New York: Harper Collins.

Cialdini, R. (2016). *Pre-suasion: A revolutionary way to influence and persuade*. Simon and Schuster.

Cialdini, R. B., & Goldstein, N. J. (2002). The science and practice of persuasion. *Cornell Hotel and Restaurant Administration Quarterly*, *43*(2), 40-50.

Dove, M. (2020). *The Psychology of Fraud, Persuasion and Scam Techniques: Understanding what Makes Us Vulnerable*. Routledge.