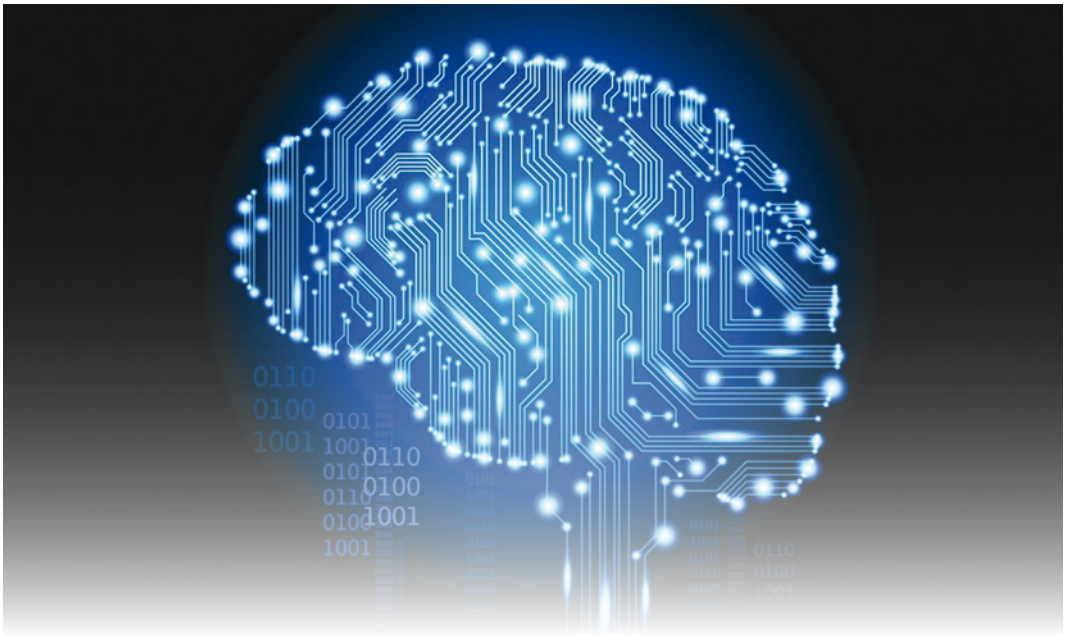


COGNITIVE BIASES IN INFORMATION SECURITY CAUSES, EXAMPLES AND MITIGATION

By Veselin Monev, information security and compliance practitioner

Abstract

This article makes a contribution to the theory of the human factor in the information security by exploring how errors in thinking distort the perceptions of InfoSec issues. Besides examples from the practice, the author proposes several ideas for mitigating the negative effects of the cognitive biases through training.

Keywords

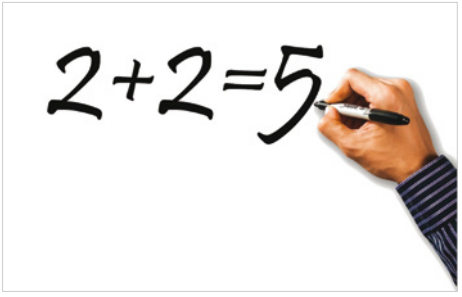
Information, security, bias, psychology, determinant, causes, mitigation, cognitive, training

INTRODUCTION

One of the components of a mature information security program is the human factor. Typically, the emphasis is on maintaining a security awareness program and mitigating risks caused by human mistakes and lack of knowledge of security.

Security awareness is necessary but also only one aspect of the human factor. Another challenge for security professionals is finding actionable arguments to support their analysis and recommendations on information security issues in their organisations. The key word here is “actionable”. Their experience shows that professional analysis, argumentation techniques and even supporting evidence combined may be insufficient for properly addressing some of the identified problems. Although a number of difficulties can be noted as causes for insufficient or inadequate actions on information security matters, like deficiency of budget, time or human resources,

management ignorance and so forth, the picture would be incomplete if the psychological phenomenon of cognitive biases are excluded.



The cognitive biases are inherent characteristics of the human nature and this way part of everyone's thinking. A bias is an error in thinking when people are processing and interpreting information and thus influencing the way they see and think about the world. Unfortunately, these biases lead to poor decisions and incorrect judgments. This article correlates researches on the biased thinking with examples from the InfoSec industry.

The first part of the article explains several important (and non-exhaustive) determinants for cognitive biases and then exemplifies them with realistic sample situations that an InfoSec specialist might encounter. The second part proposes several ideas on how organisations can deal with the biases so that their occurrences and impact are reduced. The author wants to emphasize the need for further exploration of the potency of these ideas in the real world and their role for a possible mitigation strategy. In addition, the reader is encouraged to learn about the types of cognitive biases - a topic not directly discussed here.

DETERMINANTS[1] FOR COGNITIVE BIASES AND EXAMPLES

The Misperception and Misinterpretation of Data or Events

People deal with data on an everyday basis. The common approach is to analyse the data by converting it into something more useful – information - and from there to continue the conversion into knowledge and then wisdom[2]. This complex processing chain may be impacted by the misperception or misinterpretation of random data or events. As an example, a data leakage prevention (DLP) analyst, tasked to inspect the DLP reports for irregularities, may suspect random events as real attacks on a network. In this instance, the "random" data could be misinterpreted. One should understand that human's nature is inclined to look for *patterns where such do not always exist*[3].

In a second example, a typical computer user could erroneously conclude that his computer troubles are caused by malware. However, an experienced IT support specialist could identify a different cause for the symptoms of the issue and quickly rule out the malware scenario as a cause.

Judgment by Representativeness[4]

Representativeness can be thought to have the reflexive tendency to assess the similarity of outcomes, instances, and categories on relatively salient and even superficial features, and then use these assessments of similarity as a basis of judgment.

Judgment by representativeness is often valid and helpful because objects, instances, and categories that go together usually do in fact share a resemblance. However, the overapplication of representativeness is what leads to biased conclusions. Many would likely recall personal experiences when a person, who belongs to a particular group, is attributed qualities, considered typical for that group. For instance, some IT experts perceive the members of their information security team as very strict security and compliance enforcers, but in reality not all of them may have this profile. The stereotypical over-generalisations

like "All the IT experts…", "All the auditors…", "All the consultants from that company…" often follow imprecise and even incorrect qualifications (negative or positive). The simplification can and in some instances will be misleading.

Misperceptions of Random Dispersions

If the information security professional analyses statistical data from a certain security tool, he may notice patterns, which could lead him to the conclusion that specific events occur more frequently at specific time frames[5]. For instance, if a particular type of security incident occurred for four consecutive months, each time in the last seven days of the month, this could indicate that there is a pattern. These incidents could be correlated with other known events and assumptions can be made about the underlying cause, but a definite conclusion should not be drawn without additional investigation.

Solidifying the Misperceptions with Causal Theories[6]

Once a person has (mis)identified a random pattern as a "real" phenomenon, it is likely going to be integrated into his *pre-existing beliefs*[7]. These beliefs, furthermore, serve to bias the person's evaluation of new information in such a way that the initial belief becomes solidly entrenched. For example, if a person participated as the auditee during an audit several years ago where he was supposed to provide to the auditor some of the IT security procedures, the same person could afterward develop false expectations about the requirements in other standards or for another type of organisations. This person could be convinced that he is well aware of all the auditing practices, but in reality, he could be lacking essential knowledge on the specifics of other security standards and types of audits (e.g., see the difference between SOC 2, type I and type II audits).

Misunderstanding instances of statistical regression

The statistics teach that when two variables are related, but imperfectly so, then extreme values on one of the variables tend to be matched by less extreme values on the other. For instance, a company's financially disastrous years tend to be followed by more profitable ones; Student's high scores on an exam (over 97%) tend to develop less regressive scores in the next exam.

If people are asked to predict the next result after an extreme value, they often tend not to consider the statistical regression and make non-regressive or only minimally regressive predictions (they predict a similar value).[8] A second problem is the tendency of people to fail to recognise statistical regression when it occurs and instead "explain" the observed phenomenon with complicated and even superfluous theories. This is called the regression fallacy. For example, a lesser performance that follows an exceptional one is attributed to slacking off; A slight improvement of the security incident rate is attributed to the latest policy update; Company's management may hold their IT Security Officer accountable for the decrease of the server compliance level after an excellent patching and hardening activity three months ago.

Misinterpretation of Incomplete and Unrepresentative Data (Assuming Too Much from Too Little)

The Excessive Impact of Confirmatory Information

The beliefs people hold are primarily supported by positive types of evidence. In addition, a lot of the evidence is *necessary* for the beliefs to be true but they are not always *sufficient* to warrant the same. If one fails to recognize that a particular belief rests on deficient

evidence, the belief becomes an *"illusion of validity*[9]*"* and is seen not as a matter of opinion or values but as a logical conclusion from the objective evidence that any rational person would take. The most likely reason for the excessive influence of confirmatory information is that it is easier to deal with it cognitively, compared to non-confirmatory information.

Information systems audits are good examples of *searching for confirmatory evidence*[10]*.* In an audit, unless a *statistical methodology*[11] is utilised for controls testing, the evidence for the effectiveness of the controls become open for interpretation and the auditor's intention to perform "reasonable assurance" on the controls becomes as ambiguous as it sounds. Auditors would usually ask about the existence of policies, procedures and mostly look for positive evidence. There may be even instances of auditors who ignore non-supportive evidence and ask the auditee for a supportive one. They shouldn't, but they might do so.

In another example, if the security specialist in a small company has a number of responsibilities for the entire information security management system (ISMS), there will probably be many opportunities for him to prove his skills but also to make mistakes. If the company's CEO favours the employee, he may look for achievements that indicate his professionalism. If the CEO doesn't favour him, the focus may be on the person's past mistakes, which considered alone, would indicate incompetence. In this last case, the past successes are often ignored.

The Problem of Hidden or Absent Data

In some cases, essential data could simply be absent. This makes it difficult to compare good and bad courses of action. In such situations, people could erroneously conclude that their evaluation criteria are adequate. For instance, the decision to increase the password complexity level and to lower the expiration period for the accounts of a particular business critical application is an accepted good security practice. However, if only this general best practice is taken into account, the expectations of the change could be overly optimistic. The reason for this is that a lot of missing information cannot be considered: it is nearly impossible to anticipate all the indirect consequences of such a change, like users starting to write down their passwords. If they do this, the risk for password compromise will most likely increase and the change will have the opposite effect.

In another example, the organisation's leadership decides to outsource certain IT security functions to a third-party provider instead of modernising the existing capabilities. This will likely improve the overall capabilities, but there will be very limited information if that course of action is the best decision because the other course of action will not be pursued and tested.

A third example can be given on the subject of risk assessment. People often think that if a certain risk has never materialized, then the likelihood for its occurrence in future is very low[12]. However, if a risk specialist thoroughly analyses the existing information on the risk, he may conclude that the likelihood is much higher.

Self-fulfilling Prophecies[13]

A peculiar case of the hidden data problem arises whenever our expectations lead us to act in ways that fundamentally change the world we observe. When this happens, we often accept what we see at face value, with little consideration of how things might have been different if we had acted differently. For example, if a senior manager believes that a member of the security team performs unsatisfactory, the last one will find it difficult to disprove him; If the CIO thinks the CISO behaves unfriendly, the last one could find it difficult to change his perception. Even the absence of friendliness could

be erroneously construed as unfriendliness. In such situations, the perceiver's expectations can cause the other person to behave in such a way that certain behaviours by the target person cannot be observed, making what is observed a biased and misleading indicator of what the person is like. Furthermore, if we do not like a person, we generally try to avoid him and give him little opportunity to change our expectations.

Seeing What We Expect to See[14]

The Biased Evaluation of Ambiguous and Inconsistent Data "I'll see it when I believe it."

People are inclined to see what they expect to see, and that is consistent with their pre-existing beliefs. Information that is consistent with our pre-existing beliefs is often accepted at face value, whereas evidence that contradicts it is critically scrutinised and discounted. Our beliefs may thus be less responsive than they should to the implications of new information.

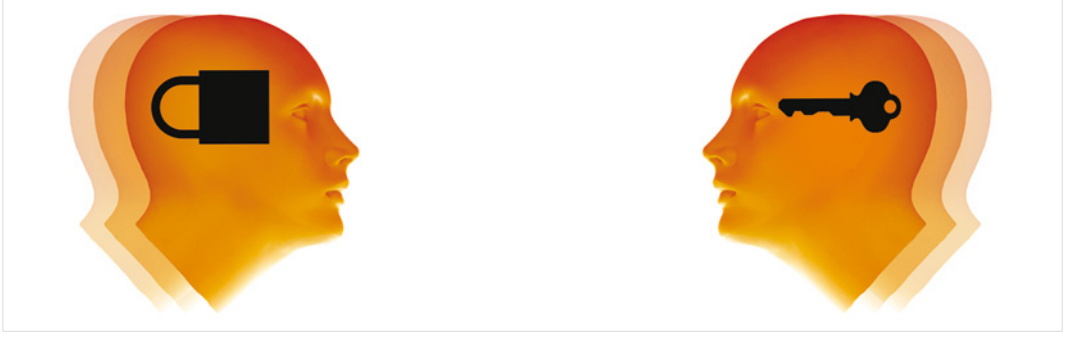
For instance, if a cybersecurity consultant is tasked to serve a client who is generally not satisfied with the IT services of the same company, the client may tend to scrutinise any piece of information the consultant provides to him and look for confirmations that the security consultancy services are at the same, unsatisfactory level as the IT services.

Ambiguous Information

If a decision is based on ambiguous information, we tend to perceive it in a way that fits our preconceptions. Why, for instance, would a newly hired Information Security Officer ask questions around in his organisation? Is he not aware of his duties or is he incapable of doing his job? Is he asking questions because there is a lack of pre-existing documentation left from his predecessor? Or is this what someone in this position is supposed to do? Or maybe because the ISMS can be effectively maintained only with the support and collaboration with the different roles in the organisation? The answer could be related to one of these questions, a combination of them or there could be a completely different explanation. Depending on the preconceptions of each employee interacting with the new Information Security Officer, they could make premature and erroneous conclusions about his capabilities.

Unambiguous Information

We tend to consider unambiguous information, which fits our beliefs, as true. However, we usually do not ignore it when it does not meet our expectations. Instead, we try to scrutinize it and look for additional information. To exemplify this, imagine a CIO who is convinced that the employees should not be occupied with information security training and instead technical controls should be preferred. Then, if he is confronted with studies, which provide evidence about the benefits of persistent security awareness training, he may tend to scrutinise them and challenge the significance of the results. He may also accept with much less scrutiny other studies, which point out the benefits of technical controls over security awareness.



MITIGATION OF COGNITIVE BIASES[15]

The list of determinants for cognitive biases can be extended. In any event, recognizing the problem is only the first issue. The second and more difficult challenge is to take adequate actions to mitigate the effects of the biases. As far as organisations are concerned, the author suggests the creation of an entire programme within the organisation, which aims to mitigate the effects of erroneous beliefs and improve employees' analytical capabilities. Depending on the characteristics of the organisation, the system could be integrated into the existing training/educational programme. The approach could focus on the following:

- Promoting the learning and self-improvement as a life-long process. People who embrace continuous learning and improvement will have more potential to detect their own cognitive biases and correct their erroneous beliefs. They will also be in a better position to respond on biased arguments of others.
- Promoting the benefits of scientific methods and techniques to create and test new theories with greater certainty. In addition to that, the knowledge on using scientific methods helps the people develop a mindset for structural thinking and distinguishes the critics from the closed-minded.
- Promoting and teaching argumentation techniques to improve the interpersonal skills of the employees.

Trained and motivated individuals should teach the actual techniques. The following ideas can be considered when creating such a programme.

- When evaluating something, the various outcomes should be specified in advance. This increases the likelihood to objectively assess the performance of processes, projects, systems and people.
- Differentiating between generating an idea and testing it. Often, people easily create ideas, but the process of proving if they work in practice is much more complicated.
- Organising training sessions to teach employees about logical constructs and avoiding biases.
- Distinguishing between secondhand and firsthand information and learning about the risks involved in relying on the first one.
- The benefits of using precise wording to describe and explain things and the perceived risks involved when using metaphors.
- The need to focus on both – the person and the individual situation, to limit distortions in the perception.
- The need to understand the false consensus effect that is defined as the tendency for people's own beliefs, values, and habits to bias their estimates of how widely others share such views and habits.
- The need to understand the distortions caused by the self-interest and how the organisation can refocus employees' attention to serve better its interest.
- Exploring the benefits of measurement methods.
- Learning about the benefits of focusing on both – the amount and kind of information.
- Learning about the tendency of positive self-assessments and the inclination of people to protect their beliefs.
- Promoting tolerance, which can be defined as the assumption that all people make mistakes. Learning about the tendency of people to remember their successes but forget their failures.
- Mastering learning techniques.
- Learning how to give and receive feedback. Often people hold back their own reservations and disbelief when they disagree with what someone is saying. Biased feedback leads to an inability to adequately evaluate alternative strategies.
- Learning how the human brain functions from a neurobiological perspective.

CONCLUSION

In a summary, this article first exemplified some determinants of cognitive biases in the context of information security and then provided some ideas on how to mitigate the implications of biased thinking in the organisations. The author believes that a better understanding and awareness of the cognitive biases will be novel for the concept of the "human factor" in the information security industry. Most importantly, the awareness of cognitive biases could provide a new perspective when designing security processes

and improve communication and decision-making of individuals. As a result of that, the already existing set of analytical and argumentation techniques of the information security professionals could be innovatively upgraded to an advanced level. Such an upgrade could improve the overall performance of the staff, especially if it encompasses the entire organisation. ■

REFERENCES

1. The determinants of cognitive biases and their definitions are discussed in the book of T. Gilovich, "How we know what isn't so", The Free Press, 1991.
2. This is known as DIKW. See L. Hayden, "IT Security Metrics", page 57-58, Mc. Graw-Hill, 2010.
3. The tendency of people to see patterns is discussed by M. Shermer, "How We Believe", 2nd edition, section "The pattern-seeking animal", Owl books, 2003.
4. This is related to the cognitive bias known as the *Representativeness Heuristic.* See A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases", pages. 1124-1131, Science, New Series, Vol. 185, No. 4157, 1974.
5. This phenomenon is also known as *Clustering Illusion*. It is well known among financial investors who could become overly confident when the price of a stock goes up for a couple of days in a row. See "Think again! Your guide to the cognitive biases that lead to bad investing behaviours and the ten things you can do about them".
6. The *Illusion of Causality* is a very well known phenomenon among scientific researchers. See "Illusions of causality: how they bias our everyday thinking and how they could be reduced", Front. Psychol., 02 July 2015.
7. It is also thought that pre-existing beliefs are the trigger for new beliefs. See "A cognitive account of belief: a tentative roadmap", Front. Psychol., 13 February 2015.
8. See D. Levitin, "Foundations Of Cognitive Psychology", pages 591-592, A Bradford Book, 2002.
9. The term is used by H. J. Einhorn & R. M. Hogarth in "Confidence in judgment: Persistence of the illusion of validity." Psychological Review, Vol 85 No 5 395-416, 1978.
10. See B. L. Luippold, S. Perreault and J. Wainberg, "AUDITORS' PITFALL: FIVE WAYS TO OVERCOME CONFIRMATION BIAS", 04.06.2015.
11. See "Practice Advisory 2320-3: Audit Sampling", The Institute of Internal Auditors, May 2013.
12. See section *Biases of imaginability* of reference 4.
13. See C. Ackerman, Self-Fulfilling Prophecy in Psychology: 10 Examples and Definition, May 2018.
14. See L. Yariv, "I'll See It When I Believe it? A Simple Model of Cognitive Consistency", Cowles Foundation Discussion Paper No. 1352, 2002.
15. The application of methods to remove or reduce bias from judgment and decision making is called *debiasing*. Multiple other techniques for mitigating the effects of cognitive biases are discussed in this article - "Debiasing", 2018

ABOUT THE AUTHOR

Veselin Monev is information security and compliance practitioner. He has over 5 years of information security experience in the academics and the private sector and more than 4 years of IT practice. In 2015 he received a master degree in Cybersecurity from the New Bulgarian University. He is author of several academic articles and co-author of an academic book for cybersecurity metrics.